AUG 28 1964

LA-DC-6526

MASTER

82576

This paper was submitted for publication in the open literature at least 6 months prior to the issuance date of this Microcard. Since the U.S.A.E.C. has no evidence that it has been published, the paper is being distributed in Microcard form as a preprint.

The Theory and Application of the Padé
Approximant Method*

George A. Baker, Jr.

University of California, Los Alamos Scientific Laboratory
Los Alamos, New Mexico

~~NOT FOR PUBLIC RELEASE - OFFICIAL
DISTRIBUTION MAY BE MADE. OFFICIAL
REQUESTS MAY BE FILLED. REPORT
CONTAINS NOTHING OF PATENT INTEREST.
PROCEDURES ON FILE IN RECEIVING
SECTION.~~

Prepared for "Advances in Theoretical Physics"

* Work performed under the auspices of the U. S. Atomic Energy Commission.

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:
A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or
B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.
As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

TABLE OF CONTENTS

	<u>Page</u>
I. Introduction	3
II. Formal Convergence Theory	7
A. General Theorems	
B. Series of Stieltjes	
C. Reduction of Bounded Analytic Functions to Series of Stieltjes	
D. The Padé Conjecture and Some Consequences	
III. Special Function Examples	47
IV. Some Applications to Physical Problems	57
A. An Optical Theorem and the Mandelstam Representation	
B. The Nature of the Forward Scattering Amplitude Series in the Coupling Constant	
C. The Ising Model and the Lattice Gas	
D. Thermodynamic Functions in Lattice Dynamics	
E. The Heisenberg Model	
F. The Many Body Problem	
G. Regge Poles	
V. Generalization of the Padé Approximant Method to Include Information at Two or More Points	87

CHAPTER I

Introduction

It is frequently the case in the solution of physical problems that one is unable to obtain a closed form for the answer. Yet one may, in principle, have completely solved the problem. This situation commonly occurs in problems of physics where one knows how, with enough effort, to carry a perturbation series expansion (commonly referred to as the "brute-force approach") to any desired order in the perturbation parameter. The motivation for the Padé approximant method is to obtain an efficient way to extract quantitative information about the solution sought from its perturbation expansion and qualitative or physical knowledge about the nature of the solution.

From the theory of functions of a complex variable (see, for instance, Copson, 1948) we know that a knowledge of every term of the power series (convergent) is equivalent by analytic continuation to a knowledge of the function everywhere (as long as there are no "natural boundaries"). However, practically speaking the power series may converge so slowly as to be a most difficult way to evaluate the function. In fact, the power series may not even converge at all at the point of interest, even though the function considered is perfectly smooth and well behaved. [Consider, for example, $\tanh(10.0)$.] From the standard theory of analytic continuation, what one must do to continue the func-

tion beyond the circle of convergence of its power series is clear. We simply compute the value of the function and as many derivatives as necessary to as high a degree of accuracy as desired at some new point inside the circle of convergence but closer to the point of interest than our original origin. We obtain a new convergent series expansion for the function. By repeating this process sufficiently often we may finally obtain the value of the function at the point of interest (presumed non-singular).

In certain cases, all this work is not necessary. We may automatically obtain the analytic continuation by appropriate manipulations on the power series. Consider, for example, a function $f(z)$ which has a singularity at $z = -1$ and all its other singularities within a circle of radius unity about $z = -2$. Suppose we wish to know $f(2)$ and $f(\infty)$. The change of variables $z = 3w/(1 - 2w)$ maps the exterior of the circle of radius unity about $z = -2$ in the z plane into the unit circle in the w plane. The points $z = 2$ and $z = \infty$ go into $w = 2/7$ and $1/2$, respectively. Since a power series converges out to the nearest singular point of the function it defines, the power series of $g(w) = f(3w/(1 - 2w))$ converges for all $|w| < 1$ and thus we may easily compute $f(2)$ and $f(\infty)$ from the power series expansion of $g(w)$.

The trouble is that one does not always a priori know the location of all the singular points of the function being studied. This trouble could be avoided if we could introduce a sequence of approximants to the function which is invariant under the group of homographic transformations, $z = Aw/(1 + Bw)$. (We shall actually consider only this subgroup of the full homographic group $z = (A + Bw)/(C + Dw)$.) One would expect such a

sequence to converge at least as well as the best power series obtainable by any trick of the type described above. (Such transformations are called Euler transformations in this connection.) Such a sequence should automatically effect the analytic continuation to any point, not directly blocked off from the origin by singularities of the function under consideration.

The sequence of $[N,N]$ Padé approximants has the property that it is invariant under the above mentioned group of homographic transformations. In general a Padé approximant [a good reference for the classical material is Wall (1948) whose notation we follow] is of the form of one polynomial divided by another polynomial. In the $[N,M]$ Padé approximant the numerator has degree M and the denominator degree N . The coefficients are determined by equating like powers of z in the following equations:

$$\begin{aligned} f(z)Q(z) - P(z) &= Az^{M+N+1} + Bz^{M+N+2} + \dots, \\ Q(0) &= 1.0. \end{aligned} \tag{I.1}$$

where $P(z)/Q(z)$ is the $[N,M]$ Padé approximant to $f(z)$. The full range of convergence of the sequence of $[N,N]$ Padé approximants is not known.

In subsequent chapters of this article we will detail, as far as is known, the convergence properties of sequences of Padé approximants. We shall also state a conjecture and give some examples and counterexamples relevant to its validity and possible extension. We will give some examples which illustrate that the Padé approximant method can be a highly efficient method of series summation.

The primary applications of Padé approximants to date fall into two general classes (a) providing efficient rational approximations to special mathematical functions and (b) to obtain quantitative information about functions which are known only from their power series and qualitative behavior. We shall not discuss the first class of applications but confine our attention to the second. We will review a number of sample applications which have been made to various physical problems. Most of these applications have been made to one aspect or another of the general field of "many-body" physics.

Finally, we will discuss the generalization of the Padé approximant method to include power series coefficients at more than one point.

CHAPTER II

Formal Convergence Theory

In this chapter we assemble various relevant theorems concerning the convergence properties of the sequence of $[N,N]$ Padé approximants. Ideally, we would like to have a theorem similar to that available for Taylor series, i.e., "the sequence of $[N,N]$ Padé approximants converges if and only if ...," where the condition is one which is simply checked and depends only on some analytic properties of the function. Unfortunately such a theorem is not now known. However a number of results of reasonable scope are available. Reduction to special cases can also usually be used to enable one to obtain rigorous upper and lower bounds which converge to the function sought in most cases of physical interest.

II A. General Theorems

The exact solution for the $[N,M]$ Padé approximant to $A(x)$ may be exhibited explicitly in terms of the series coefficients a_j . It is

$$[N, M](x) = \frac{\begin{vmatrix} a_{M-N+1} & a_{M-N+2} & \cdots & a_{M+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_M & a_{M+1} & \cdots & a_{M+N} \\ \sum_{j=N}^M a_{j-N} x^j, & \sum_{j=N-1}^M a_{j-N+1} x^j, & \cdots & \sum_{j=0}^M a_j x^j \end{vmatrix}}{\begin{vmatrix} a_{M-N+1} & a_{M-N+2} & \cdots & a_{M+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_M & a_{M+1} & \cdots & a_{M+N} \\ x^N & x^{N-1} & \cdots & 1 \end{vmatrix}} \quad (\text{II.1})$$

where $a_j \equiv 0$ if $j < 0$, and sums for which the initial point is larger than the terminal point are to be omitted. This formula is easily verified by substitution in the equation which it must satisfy, Eq. (I.1). The numerator and denominator of (II.1) have, for convenience, been multiplied by a common factor. The expression (II.1) is also unique (Wall, 1948). For suppose that there were another solution of Eq. (I.1) $Y(x)/Z(x)$ where Y and Z are of degree less than or equal to M and N respectively. Then if we multiply $\{[N, M](x) - Y(x)/Z(x)\}$ times $Z(x)$ times the denominator of II.1 we get a polynomial of degree $M + N$; however, by Eq. (I.1) all the coefficients through x^{N+M} are zero, hence, if we agree to cancel all

common factors between numerator and denominator so as to always present the results as a ratio of polynomials of lowest degree, then there is at most one solution to Eq. (I.1).

There are basically two types of sequences of Padé approximants which one can consider. In the first type only the degree of the numerator or denominator becomes infinite while the other remains finite. In the second type the degree of both numerator and denominator tend to infinity together. The second type seems to be the more powerful, while the first is the simplest one to obtain rigorous theorems for. R. de Montessus de Balloire (1902) has shown that if $P(z)$ is a power series representing a function which is regular for $|z| \leq R$, except for m poles within this circle, then the $[m, N]$ Padé converge uniformly, as N tends to infinity, to $P(z)$ in the domain obtained from $|z| \leq R$ by removing the interiors of small circles with centers at these poles. Wilson (1930) has extended this result to the case where the function defined by $P(z)$ has certain nonpolar types of singularities on the circle $|z| = R$. He proves that if $P(z)$ is a power series representing a function which is regular for $|z| \leq R$ except for m poles within this circle and one nonpolar singularity on the boundary such that $P(z)$ is continuous at the singularity, then the $[m + \tau, N]$ Padé approximants, $\tau \geq 0$, converge uniformly to $P(z)$ as $N \rightarrow \infty$ in the domain obtained from $|z| \leq R$ by removing the interiors of small circles with centers at these poles. The same type of results also hold when the degree of the denominator goes to infinity while that of the numerator remains finite. There are, however, examples (Peron, 1957) for which $P(z)$ has a finite or even infinite

radius of convergence for which $[1, N]$ diverge on an everywhere dense set within the circle of convergence of the power series.

We shall devote the remainder of this chapter to a discussion of $[N, N+k]$ sequences of Padé approximants with k finite. We remark that the study of any sequence of $[N, N+k]$ Padé approximants can be reduced to the study of the $[N, N]$ Padé approximants by simply considering

$$g(z) = z^{-k} f(z) \quad (\text{II.2})$$

if $k \leq 0$, or

$$g(z) = \left[f(z) + \sum_{j=0}^{k-1} f_j z^j \right] z^{-k} \quad (\text{II.3})$$

if $k > 0$, where f_j are the coefficients in the power series expansion of $f(z)$.

For one class of functions, exact expressions have been obtained for the $[p, p]$ Padé approximants, the error of the p th one, etc. by Luke (1958, 1960). He has proved that if z is not zero, a singularity, or a branch cut of $E(z) = {}_2F_1(1, w; \lambda+1; -z^{-1})$, then

$$\lim_{p \rightarrow \infty} \{ [p, p] - E(z) \} = 0. \quad (\text{II.4})$$

This class is a special case of the continued fraction of Gauss [Chap. 18, Wall (1948)], ${}_2F_1(a, b+1; c+1; -z) / {}_2F_1(a, b; c; -z)$, which encompasses many elementary functions and for which similar results can be developed. These are series of Stieltjes the theory for which we treat in the next section.

The following two theorems show that the "natural" domain of convergence for the sequence of $[N, N]$ Padé approximants is not a circle centered at the origin, but rather a region determined in some way by location of the nonpolar singularities of the function considered.

Theorem 1. If $P_N(z)/Q_N(z)$ is the $[N, N]$ Padé approximant to $f(z)$, then $P_N[Aw/(1 + Bw)]/Q_N[Aw/(1 + Bw)]$ is the $[N, N]$ Padé approximant to $f[Aw/(1 + Bw)]$.

Proof: If the numerator and denominator of $P_N[Aw/(1 + Bw)]/Q_N[Aw/(1 + Bw)]$ are multiplied by $(1 + Bw)^N$ then it becomes of the form $p_N(w)/q_N(w)$. If we expand this as a power series in w by expanding P_N/Q_N and then expanding $(1 + Bw)^{-1}$, we readily check that $p_N(w)/q_N(w)$ agrees with the expansion of $f[Aw/(1 + Bw)]$ at least through the terms in w^{2N} . Since, as we pointed out above, the $[N, N]$ Padé approximant is unique, Theorem 1 follows. We remark that for the $[N, N+j]$ sequence, it is $(1 + Bw)^j f[Aw/(1 + Bw)]$ to which this invariance property applies.

Theorem 2. If $P_N(z)/Q_N(z)$ is the $[N, N]$ Padé approximant to $f(z)$ and $C + Df(0) \neq 0$, then $\{A + B[P_N(z)/Q_N(z)]\} / \{C + D[P_N(z)/Q_N(z)]\}$ is the $[N, N]$ Padé approximant to

$$\{A + B f(z)\} / \{C + D f(z)\}.$$

Proof: If the numerator and denominator of

$$\{A + B[P_N(z)/Q_N(z)]\} / \{C + D[P_N(z)/Q_N(z)]\}$$

are multiplied by $Q_N(z)$, then it becomes of the form $p_N(z)/q_N(z)$ as long

as $q_N(0) = C + Df(0) \neq 0$. Thus Theorem 2 follows from the uniqueness of the Padé approximants.

Clearly, Theorems 1 and 2 can be applied successively so that both the value and the argument can be transformed simultaneously without effecting the sequence of Padé approximants.

Theorem 3. Let $P_k(z)$ be any infinite sequence of $[N, M]$ Padé approximants to a formal power series where $N + M$ tend to infinity with k . If the absolute value of the P_k is uniformly bounded for $|z| \leq R$, then the P_k converge uniformly for $|z| \leq r < R$ to an analytic function $f(z)$ whose power series has a radius of convergence of at least R .

Proof. If any $|P_k(z)|$ is bounded by W for $|z| \leq R$, then by its nature as a rational function of z it must be analytic in $|z| \leq R$ and its power series must have a radius of convergence of at least R . Now consider the region $|z| \leq r < R$. As the power series coefficients, $p_k(n)$, of z^n must, by Cauchy's inequalities, [see, for instance, Sec. 4.33 of Copson (1948)] be bounded by W/R^n uniformly in k , we obtain

$$\begin{aligned} \left| P_k(z) - \sum_{n=0}^J p_k(n) z^n \right| &\leq \sum_{n=J+1}^{\infty} |p_k(n) z^n| \\ &\leq \sum_{n=J+1}^{\infty} W(r/R)^n = W(r/R)^{J+1} / [1 - (r/R)] \end{aligned} \quad (\text{II.5})$$

by selecting J large enough, as $(r/R) < 1$, we may make the difference between $P_k(z)$ and the first $J + 1$ terms of its power series smaller than any preassigned error $\frac{1}{2} \epsilon > 0$ uniformly in k . As $N + M$ tends to infinity with k , there exists a K such that for $k \geq K$, $N + M \geq J$, and hence, by

Eq. I.1, the first $J + 1$ power series terms are identical for all $k \geq K$.

Therefore, if k , and $j \geq K$ and $|z| \leq r$

$$|P_k(z) - P_j(z)| \leq \left| P_k(z) - \sum_{n=0}^J p_k(n) z^n \right| + \left| P_j(z) - \sum_{n=0}^J p_j(n) z^n \right|$$

(II.6)

$$\leq \frac{1}{2} \epsilon + \frac{1}{2} \epsilon = \epsilon$$

which is just the Cauchy criterion for the convergence of the sequence.

By the theorem on the uniform convergence of a sequence of analytic functions [See, for instance, Sec. 5.13 of Copson (1948).] the limiting function $f(z)$ is analytic for $|z| < R$. Thus by applying Taylor's Theorem to $f(z)$ we complete the proof of Theorem 3. This proof is a generalization of one due to Van Vleck (1901) for continued fractions.

Theorem 4. Let $P_k(z)$ be any infinite sequence of $[N, M]$ Padé approximants to a formal power series where $N + M$ tend to infinity with k . If the $|P_k(z)|$ is uniformly bounded in any closed, simply-connected domain D_1 which contains the origin as an interior point and $|P_k(z)|^{-1}$ is uniformly bounded in any closed, simply-connected domain D_2 , which contains the origin as an interior point, then the P_k converge to an meromorphic function $f(z)$ in the interior of the union of D_1 and D_2 .

Proof: First we will consider D_1 . Since the origin is an interior point, there exists an R_1 such that $|z| \leq R_1$ is completely in D_1 . As the P_k are uniformly bounded rational functions in D_1 , they are analytic in D_1 and in particular in $|z| \leq R_1$. We can therefore establish by Theorem 3,

the conclusions of this Theorem in $|z| \leq R_1$. We can now extend, by repeated applications of Theorem 3, the results throughout D_1 . First we note that by the Theorem on the uniform convergence of a sequence of analytic functions that the sequence of nth derivatives also converge uniformly. Hence we may expand the $P_k(z)$ about any desired point a , $|a| < R_1$, apply Theorem 3 and thereby extend the conclusions of the Theorem to the union of $|z| < R_1$, $|z-a| < R_2$. If we continue to repeat this process, we can thus prove that the limit $f(z)$ is analytic at any interior point of D_1 . Since this procedure is nothing more than a justification of the application of the standard method of analytic continuation to this particular case, it follows that the P_k converge uniformly in any closed domain interior to D_1 , to that function which is the analytic continuation of the function $f(z)$ defined by the power series.

For D_2 we note that $[P_k(z)]^{-1} = [M, N]$ Padé approximant and thus the $[P_k(z)]^{-1}$ form a suitable sequence. This result follows from the uniqueness of the Padé approximant pointed out above (Eq. II.1). Thus by repeating the proof given for D_1 , we prove that $[P_k(z)]^{-1}$ converges uniformly in any closed domain interior to D_1 to that function which is the analytic continuation of the function $[f(z)]^{-1}$ defined by the power series. As the reciprocal of an analytic function is at worst meromorphic, Theorem 4 is proved.

This last Theorem gives a rule of procedure in practice. Select a region of interest in the complex plane in which it is known on physical grounds that the function is meromorphic (we will discuss a slight relaxation of this condition in the next section). This region might, for

instance, be a neighborhood of a portion of the real axis. From an infinite sequence of $[N, M]$ Padé approximants select an infinite subsequence which obeys the conditions of Theorem 4 in the region of interest. Then they must, by virtue of Theorem 4, converge to the desired answer. The main mathematical problem associated with the convergence problem is to establish the existence of such an infinite bounded subsequence when M and N tend to infinity together. As we have pointed out above when M or N remain finite, quite a number of results are known. The problem basically is not the value of the bound as if there exists an infinite subsequence less than W then there must, by Theorem 4, also exist an infinite subsequence less than $\max \left[\max_{D_1} |f(z)|, \max_{D_2} |f(z)|^{-1} \right] + \delta$ for any $\delta > 0$. The problem is to prove that there is an infinite subsequence for which the poles of the approximant are excluded from the region of interest. There are at least two partial proofs but neither were able to prove this point. Baker and Gammel (1961) considered meromorphic functions and argued that the $[N, N]$ Padé approximants should reproduce the closest $M \ll N$ poles very well. Chisholm (1963) showed that the exact solution of a nonhomogeneous linear integral equation with a kernel of rank n is given by the $[n, n]$ Padé approximant in the expansion parameter. The denominator of this approximant is a function only of the expansion parameter and vanishes at the reciprocals of the eigenvalues of the kernel. Chisholm (1963) has attempted to extend this result to convergence of the $[n, n]$ Padé approximants for compact kernels, but was not able to pin down the location of the poles.

II B. Series of Stieltjes

Although some functions of interest in physical problems can be shown to be represented by series of Stieltjes, e.g. the ladder approximation to the many-fermion energy for some potentials, our main interest in them lies in the fact that as we will show in the next section of this Chapter, the study of a large number of cases of physical interest can be reduced to their study.

By a series of Stieltjes [most of the basic results in this section are due to Stieltjes (1894) although we do not always follow his methods] we mean

$$f(z) = \sum_{j=0}^{\infty} f_j (-z)^j \quad (\text{II.7})$$

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

$$f_j = \int_0^{\infty} u^j d\phi(u). \quad (\text{II.8})$$

(Note that we do not necessarily assume that the series is convergent.) If we introduce the determinants

$$D(m,n) = \det \begin{vmatrix} f_m & f_{m+1} & \cdots & f_{m+n} \\ f_{m+1} & f_{m+2} & \cdots & f_{m+n+1} \\ \vdots & \vdots & \ddots & \vdots \\ f_{m+n} & f_{m+n+1} & \cdots & f_{m+2n} \end{vmatrix} \quad (\text{II.9})$$

then this definition is equivalent to the conditions

$$D(0,n) > 0, \quad D(1,n) > 0, \quad n = 0,1,2,\dots \quad (\text{II.10})$$

To see that (II.10) follows from (II.8) consider the quadratic form

$$\sum_{p,q=0}^n f_{p+q+m} x_p x_q = \int_0^\infty u^m (x_0 + x_1 u + \dots + x_n u^n)^2 d\varphi(n) \quad (\text{II.11})$$

> 0

as φ takes on infinitely many values and n is finite; but, by the variational principle, the minimum of (II.11) subject to $\sum x_j^2 = 1$ is the smallest eigenvalue of the matrix $D(m,n)$. Thus all the eigenvalues are positive and hence the determinants (II.9), being the products of their eigenvalue are necessarily positive. That (II.8) also follows from (II.10) is more difficult and we will indicate its proof following

Theorem 7.

From the arguments of (II.11) it clearly follows that $D(m,n) > 0$ for all m and n . If the lower limit in (II.8) and (II.11) is extended to $-\infty$, then for m even (II.10) still holds but not necessarily for m odd. One can show by the same methods that $D(0,m) > 0$ or $D(1,n) > 0$ alone imply the results of this section for the $[N, N+j]$ sequence where the whole real axis plays the role of the negative real axis in the following Theorems. Theorem 7, however, requires more than an obvious modification and Theorem 6 holds only to the right of zero and left of all poles.

We remark that if

$$(-1)^n D(1,n) < 0, \quad (-1)^n D(2,n) < 0$$

then, if $\bar{D}$ are the D's formed by the reciprocal series, by a Theorem of Hadamard (1892)

$$\bar{D}(1,n) = (-1)^{n+1} D(1,n)/f_0^{2n+2} > 0$$

$$\bar{D}(0,n) = (-1)^n D(2,n-1)/f_0^{2n+1} > 0,$$

and so $1/f(z)$ is a series of Stieltjes.

The importance of these determinantal inequalities is clearly seen from Eq. (II.1). For the determinant for the $[N,M]$ Padé approximant is clearly just $D(M-N+1, N-1) > 0$, for $M \geq N-1 \geq 0$ for a series of Stieltjes. We shall now show that these determinantal inequalities also determine the location of all poles of any $[N, N+j]$, $j \geq -1$, Padé approximant.

Theorem 5. If $\sum f_j(-z)^j$ is a series of Stieltjes, then the poles of the $[N, N+j]$, $j \geq -1$, Padé approximants are 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.

Proof: If we start with the 2nd column of the denominator in (II.1) and multiply by x and then subtract it from the first column, and then repeat the procedure for the 2nd and 3rd columns, etc. until we reach the last two columns, we can reduce the denominator to

$$(-1)^{N(N+j)} \begin{vmatrix} f_{1+j} + xf_{2+j}, & f_{2+j} + xf_{3+j}, & \cdots & f_{N+j} + xf_{N+j+1} \\ f_{2+j} + xf_{3+j}, & f_{3+j} + xf_{4+j}, & \cdots & f_{N+j+1} + xf_{N+j+2} \\ \vdots & \vdots & \ddots & \vdots \\ f_{N+j} + xf_{N+j+1} & f_{N+j+1} + xf_{N+j+2}, & \cdots & f_{2N+j-1} + xf_{2N+j} \end{vmatrix} \quad (\text{II.12})$$

If we now denote this determinant (without signfactor) by Δ_N and the co-axial minor formed by striking off the last r rows and columns by Δ_{N-r} (denominator for the $[N-r, N+j-r]$) then the sequence $\Delta_N, \Delta_{N-1}, \dots, \Delta_1, \Delta_0 = 1$, is a Sturm sequence, that is to say, if $\Delta_k = 0$, then Δ_{k-1} and Δ_{k+1} have opposite signs. To see this result we use the determinant Theorem [pg. 372, Muir (1960)] that if A is axisymmetric, then

$$A_{11}A_{22} - A_{12}^2 = A \cdot A_{12,12} \quad (\text{II.13})$$

where A_{11} , etc. are the various minors. Hence if A_{11} vanishes, A and $A_{12,12}$ most necessarily have opposite signs.

To locate the roots we first note that if $(-1)^{N(N+j)} Q_N^{(j)}(z)$ is the denominator given by (II.1) for the $[N, N+j]$ approximant $Q_N^{(j)}(0) = D(1+j, N-1) > 0$ for all N . Now

$$Q_1^{(j)}(z) = f_{1+j} + xf_{2+j}, \quad (\text{II.14})$$

and since all the $f_k = D(k, 0) > 0$, the root of $Q_1^{(j)}(z)$ is real and

negative. Since $Q_2^{(j)}(0)$ is positive and by (II.13) is negative at the negative real root of $Q_i^{(j)}(z)$, and a polynomial of degree two, it must vanish between zero and the root of $Q_1^{(j)}(z)$. The coefficient of x^2 is $D(j+2,1)$ by (II.1) and as N is even the sign factor of (II.12) is +1. Therefore as $x \rightarrow -\infty$ $Q_2^{(j)}(z)$ must become positive. Therefore there are two real negative roots. In general we see by induction that the first N roots of $Q_{N+1}^{(j)}(z)$ lie in the N intervals formed by zero and the N roots of $Q_N^{(j)}(z)$. As the coefficient of x^{N+1} is $(-1)^{N+1} D(j+2,N)$, the single remaining root must be real and negative.

To demonstrate the positiveness of the residues it is convenient to show first that if $(-1)^{N(N+j)} P_N^{(j)}(z)$ is the numerator given by (II.1) for the $[N, N+j]$ approximant then

$$P_{N+1}^{(j)}(z) Q_N^{(j)}(z) - P_N^{(j)}(z) Q_{N+1}^{(j)}(z) = (-z)^{2N+1+j} [D(1+j, N)]^2 \quad (\text{II.15})$$

Now by Eq. (I.1)

$$Q_N^{(j)}(z) \frac{P_{N+1}^{(j)}(z)}{Q_{N+1}^{(j)}(z)} - P_N^{(j)}(z) = O(z^{2N+1+j}); \quad (\text{II.16})$$

thus, multiplying by $Q_{N+1}^{(j)}(z)$ we obtain a polynomial of degree $2N+1+j$ for which only the highest coefficient differs from zero. To establish (II.15) it remains only to evaluate that coefficient. The $2N+1+j$ th coefficient in $[N+1, N+1+j]$ must be $(-1)^{1+j} f_{2N+1+j}$ and the $2N+1+j$ th coefficient in $[N, N+j]$ must be

$$-(-1)^{1+j} \begin{vmatrix} f_{1+j} & \cdot & \cdot & \cdot & f_{N+j} & f_{N+j+1} \\ \vdots & & & & \vdots & \\ f_{N+j} & \cdot & \cdot & \cdot & f_{2N+j-1} & f_{2N+j} \\ f_{N+j+1} & \cdot & \cdot & \cdot & f_{2N+j} & 0 \end{vmatrix} \div D(1+j, N-1) \quad (\text{II.17})$$

as this is coefficient which makes (I.1) hold to one higher order. Subtracting $[N+1, N+1+j] - [N, N+j]$ the coefficient of z^{2N+1+j} becomes $(-1)^{1+j} D(1+j, N) \div D(1+j, N-1)$. Thus using the values of $Q(0)$ we obtain (II.15) which it is convenient to rewrite as

$$\frac{P_{N+1}^{(j)}(z)}{Q_{N+1}^{(j)}(z)} - \frac{P_N^{(j)}(z)}{Q_N^{(j)}(z)} = \frac{(-z)^{2N+1+j} [D(1+j, N)]^2}{Q_{N+1}^{(j)}(z) Q_N^{(j)}(z)} \quad (\text{II.18})$$

By the interlacing property the residue at the first (counting in the negative direction from the origin) root of $Q_{N+1}^{(j)}$ must be positive on the right and so also on the left of (II.18). We now pass the first and only the first root of $Q_N^{(j)}$ before we come to the second root of $Q_{N+1}^{(j)}$. As both $Q_N^{(j)}$ and $Q_{N+1}^{(j)}$ are now negative, the residue at the second root is also positive. We can repeat this process until we have shown that all the residues are positive. This result, together with the sign oscillations of the denominator, force the roots of the numerator to interlace those of the denominator.

The following Theorem is one of the most valuable in the study of series of Stieltjes for it provides upper and lower bounds for the exact sum of the series, whether convergent or not.

Theorem 6. The Padé approximants for series of Stieltjes obey the following inequalities where $f(z)$ is the sum of the series $\sum f_j(-z)^j$, and z is real and non-negative.

$$(-1)^{1+j} \left\{ [N+1, N+1+j] - [N, N+j] \right\} \geq 0 \quad (\text{II.19a})$$

$$(-1)^{1+j} \left\{ [N, N+j] - [N-1, N+j+1] \right\} \geq 0 \quad (\text{II.19b})$$

$$[N, N] \geq f(z) \geq [N, N-1] \quad (\text{II.19c})$$

$$[N, N]' \geq f'(z) \geq [N, N-1]' \quad (\text{II.19d})$$

where $j \geq -1$. These inequalities have the consequence that the $[N, N]$ and $[N, N-1]$ sequences form the best upper and lower bounds obtainable from the $[N, N+j]$ approximants with a given number of coefficients and that the use of additional coefficients (higher N) improves the bounds. Equation (II.19a,b) are valid when differentiated, provided $j \geq 0$ for (II.19a).

Proof: Inequality (II.19a) follows directly from Theorem 5 and Eq. (II.18) as for $z \geq 0$ the Q 's are positive and so $(-1)^{1+j}$ times the right hand side of (II.18) is positive or zero in this range. To prove (II.19b) we proceed exactly as we did to derive Eq. (II.18) and we obtain

$$\frac{P_N^{(j)}(z)}{Q_N^{(j)}(z)} - \frac{P_{N-1}^{(j+2)}(z)}{Q_{N-1}^{(j+2)}(z)} = \frac{(-z)^{2N+1+j} [D(1+j, N-1) D(3+j, N-1) - D(1+j, N) D(3+j, N-2)]}{Q_N^{(j)}(z) Q_{N-1}^{(j+2)}(z)} \quad (\text{II.20})$$

The quantity in the square bracket is always positive by Eq. (II.13). We identify the axisymmetric determinant $D(1+j, N)$ with A , $D(3+j, N-1)$ with A_{11} , $D(1+j, N-1)$ with A_{22} , and $D(3+j, N-2)$ with $A_{12,12}$. Hence the square bracket

is $A_{12}^2 > 0$. As the Q 's are again positive for $z \geq 0$, $(-1)^{1+j}$ times the right hand side is non-negative. Repeated application of (II.19b) allows one to bound $[N, N+j]$ by the first $2N+j$ power series terms. To derive (II.19c) we again follow the procedure of (II.18) and obtain

$$\frac{P_N^{(0)}(z)}{Q_N^{(2)}(z)} - \frac{P_N^{(-1)}(z)}{Q_N^{(-1)}(z)} = \frac{z^{2N} [D(0, N)] [D(1, N-1)]}{Q_N^{(0)}(z) Q_N^{(-1)}(z)} \quad (\text{II.21})$$

from which (II.19c) follows at once. Since the denominator of (II.21) is of degree $2N$, (II.19d) follows at once by differentiating (II.21). The derivative inequalities corresponding to (II.19a,b) follow by differentiation of (II.18) and (II.20) as the degree of the numerator is greater than or equal to that of the denominator.

Theorem 7. Any sequence of $[N, N+j]$ Padé approximants for series of Stieltjes converges to an analytic function in the cut complex plane $(-\infty \leq z \leq 0)$. If, in addition, $\sum_{p=1}^{\infty} (f_p)^{-1/(2p+1)}$ diverges, then all the sequences tend to a common limit. (This condition is roughly equivalent to $|f_p| \leq (2p)!)$ 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 $(-\infty \leq z \leq -R)$ to the analytic function defined by the power series.

Proof: The first step is to note that the sequence of $[N, N+j]$ is strictly monotonic for z real and positive by (II.19c) of Theorem 6. For j odd they are monotonic increasing and bounded from above by, say, the $[1, 1]$ approximant. For j even they are monotonically decreasing and bounded from below by, say, the $[1, 0]$ approximant. Thus by the

Bolzano-Weierstrass Theorem [see, for example, Copson (1948)] these sequences converge at every real point.

The second step is to note that $|[N, N+j]| < \left| \sum_{p=0}^j f_p(-z)^p \right| + |z|^{1+j} f_{1+j}$ for $\text{Re}(z) \geq 0$. This result follows from Theorem 5, as follows. We may rewrite

$$[N, N+j] = \sum_{p=0}^j f_p(-z)^p + (-z)^{1+j} \sum_{p=1}^N \frac{\beta_p}{1 + \gamma_p z}, \quad \begin{matrix} \beta_p > 0 \\ \gamma_p > 0 \end{matrix} \quad (\text{II.22})$$

according to Theorem 5. Thus

$$\begin{aligned} |[N, N+j]| &\leq \left| \sum_{p=0}^j f_p(-z)^p \right| + |z|^{1+j} \sum_{p=1}^N \frac{\beta_p}{|1 + \gamma_p z|} \\ &\leq \left| \sum_{p=0}^j f_p(-z)^p \right| + |z|^{1+j} \sum_{p=1}^N \beta_p \\ &= \left| \sum_{p=0}^j f_p(-z)^p \right| + |z|^{1+j} f_{1+j} \end{aligned} \quad (\text{II.23})$$

as $\text{Re}(z) \geq 0$. When $\text{Re}(z) < 0$, we note that the absolute value of the denominators in (II.23) are greater than $|\gamma_p \text{Im}(z)|$ hence

$$|[N, N+j]| \leq \left| \sum_{p=0}^j f_p(-z)^p \right| + \frac{|z|^{1+j}}{|\text{Im}(z)|} \sum_{p=1}^N \frac{\beta_p}{\gamma_p} \quad (\text{II.24})$$

however, by (II.1) for the coefficient of the largest power of z ,

$$f_j - \sum_{p=1}^N \frac{\beta_p}{\gamma_p} = D(j, N)/D(1+j, N-1) > 0 \quad (\text{II.25})$$

as the D's are positive. Thus if $\text{Re}(z) < 0$

$$|[N, N+j]| \leq \left| \sum_{p=0}^j f_p (-z)^p \right| + f_j |z|^{1+j} / |\text{Im}(z)|.$$

(For $j = -1$, we set instead $|[N, N-1]| < f_0 / |\text{Im}(z)|$ here.) Thus, if we are outside the region $|\text{Im}(z)| \leq \delta$, $\text{Re}(z) < 0$, the $|[N, N+j]|$ is uniformly bounded independent of N . If we now fix our attention on a real point $0 < x_0 < \infty$, then as we have just seen in the circle $|z - x_0| \leq x_0$, every $[N, N+j]$ approximant is uniformly bound by, say W . Thus we may (see proof of Theorem 3) represent each approximant by a power series of $M(\epsilon)$ terms to within an error $1/6 \epsilon > 0$ uniformly in N . Suppose we now expand, by Taylor's Theorem with remainder [see, for example, Sec. 89 of Franklin (1940)], the $[N, N+j]$ approximant at x_0

$$\begin{aligned} g(x) = [N, N+j](x) &= [N, N+j](x_0) + [N, N+j]'(x_0) \frac{(x - x_0)}{1!} \\ &+ [N, N+j]''(x_0) \frac{(x - x_0)^2}{2!} + \dots + [N, N+j]^{(n)}(x_0) \frac{(x - x_0)^n}{n!} \\ &+ [N, N+j]^{(n+1)}(\xi) \frac{(x - x_0)^{n+1}}{(n+1)!} \end{aligned} \quad (\text{II.26})$$

where ξ lies between x and x_0 . Thus

$$[N, N+j](x_0) = g(x_0)$$

$$[N, N+j]'(x_0) = \frac{g(x_0 + \Delta) - g(x_0)}{\Delta} = [N, N+j]''(\xi_{11}) \frac{\Delta}{2!}$$

$$\begin{aligned} & \vdots \\ & [N, N+j]^{(n)}(x_0) = \sum_{r=0}^n \binom{n}{r} (-1)^r g[x_0 + (n-r)\Delta] \Delta^{-n} \\ & \quad - \frac{\Delta}{(n+1)!} \sum_{r=0}^n \binom{n}{r} (-1)^r [N, N+j]^{(n+1)}(\xi_{nr}) \end{aligned} \quad (\text{II.27})$$

where the ξ_{nr} are between x_0 and $x_0 + (n-r)\Delta$. If we pick Δ small enough so that all the points ξ_{nr} satisfy $|\xi_{nr} - x_0| \leq x_0$ for $r \leq n \leq M(\epsilon)$, then

$|[N, N+j]^{(n+1)}(\xi_{nr})| \leq (n+1)! W/x_0^{n+1}$ by Cauchy's inequalities. Hence

$$\begin{aligned} & \left| [N, N+j]^{(n)}(x_0) - \sum_{r=0}^n \binom{n}{r} (-1)^r g[x_0 + (n-r)\Delta] \Delta^{-n} \right| \\ & \leq W \Delta \cdot 2^n / x_0^{n+1}. \end{aligned} \quad (\text{II.28})$$

Thus, at a distance $r = |z - x_0| < x_0$ from x_0

$$\begin{aligned} & \left| \sum_{n=0}^{M(\epsilon)} [N, N+j]^{(n)}(x_0) \frac{(z - x_0)^n}{n!} - \sum_{n=0}^{M(\epsilon)} \sum_{r=0}^n \binom{n}{r} (-1)^r g[x_0 + (n-r)\Delta] \left(\frac{z - x_0}{\Delta} \right)^n \right| \\ & \leq \frac{W\Delta}{x_0} \sum_{n=0}^{M(\epsilon)} 2^n \left| \frac{z - x_0}{x_0} \right|^n \leq \frac{2^{M(\epsilon)} W\Delta}{x_0 - |z - x_0|} \leq \epsilon/6 \end{aligned} \quad (\text{II.29})$$

for $\Delta(\epsilon)$ taken small enough. Thus if we fix ϵ we may determine $M(\epsilon)$ by (II.5)

and then Δ by (II.29) such that

$$\left| [N, N+j](z) - \sum_{n=0}^{M(\epsilon)} \sum_{r=0}^n \binom{n}{r} (-1)^r g[x_0 + (n-r)\Delta] \left(\frac{z-x_0}{\Delta} \right)^n \right| \leq 1/3 \epsilon \quad (\text{II.30})$$

If now we select an $N(\epsilon)$ large enough so that the $g(x_0 + j\Delta)$, $j = 0, 1, \dots, n(\epsilon)$, for any N_1 and $N_2 > N(\epsilon)$ differ from each other by less than

$$\frac{1}{3} \epsilon [2|r/\Delta| - 1] / [2|r/\Delta|^{n+1} - 1]$$

then for all N_1 and $N_2 > N(\epsilon)$, and $|z - x_0| \leq r < x_0$

$$|[N_1, N_1 + j](z) - [N_2, N_2 + j]| \leq \epsilon \quad (\text{II.31})$$

which is just the Cauchy criterion for convergence. Hence the sequence of $[N, N+j]$ Padé approximants converges for $|z - x_0| < x_0$ to a regular function of z . We may now follow the method of proof of Theorem 4 and extend the region of convergence to any point in the cut plane; thereby obtaining the first statement of Theorem 7.

Suppose in addition to being a series of Stieltjes, the series is also convergent with radius of convergence R . From the property

$$|[N, N+j]^{(p)} / (p!)| \leq f_p \quad (\text{II.32})$$

we see at once that the nearest singularity of $[N, N+j]$ must be outside or on the circle $|z| \leq R$ for all N . Thus the γ_p of (II.22) must satisfy $\gamma_p \geq R^{-1} > 0$. Hence from (II.23), if $|z| \leq R$

$$|[N, N+j](z)| \leq \left| \sum_{p=0}^j f_p (-z)^p \right| + \frac{|z|^{1+j} f_{1+j}}{|1 + z/R|} \quad (\text{II.33})$$

By Theorem 4, the last statement of Theorem 7 follows immediately. Property (II.32) may be seen by expanding the right hand side of (II.18).

First, if $p \leq 2N+j$ then the equality holds. If $p > 2N+j$

$$f_p = (-1)^p [N, N+j]^{(p)} / (p!) = (-1)^p \left\{ [N+1, N+j+1]^{(p)} / p! - [N, N+j]^{(p)} / (p!) \right\} \\ + \dots + (-1)^p \left\{ (-1)^p f_p - [N_\alpha, N_\alpha + j]^{(p)} / (p!) \right\} > 0 \quad (\text{II.34})$$

where N_α is large enough so the last $\left\{ \right\}$ vanishes. The $\left\{ \right\}$ are nothing but single terms in the $(-1)^p$ times the right hand sides of (II.18) for successive values of N . The sign of z^p in (II.18) is $(-1)^p$ as we may rewrite the right hand side as

$$(-z)^{2N+1+j} K \prod_{k=1}^{2N+1} (1 + \alpha_k z)^{-1} = K \sum_{v=2N+1+j}^{\infty} \left[\sum_{\substack{v_1+v_2+\dots+v_{2N+1} \\ =v-2N-1-j}} \prod_{k=1}^{2N+1} \alpha_k^{v_k} \right] (-z)^v \quad (\text{II.35})$$

where K and the α_k are all positive. Thus using (II.35) in (II.34) we prove the inequality. As $(-1)^p [N, N+j]^{(p)} / (p!)$ by (II.22) we have

$$f_p \geq (-1)^p [N, N+j]^{(p)} / (p!) > 0 \quad (\text{II.36})$$

which proves (II.32).

We shall now establish the uniqueness of the limit even for certain types of divergent series. For series of Stieltjes any diagonal sequence of Padé approximants sums to the unique analytic function implied by the formal power series when it implies a unique function. The condition in Theorem 7 is, by a Theorem of Carleman [Hardy (1956)], necessary and

sufficient that there be at most one analytic function which is regular in some region which permits the approach of the origin on all rays and which has the derivatives implied by the power series as we approach along any ray. These conditions are clearly satisfied, except for the negative real axis. Here it suffices to consider a circular arc with center on the imaginary axis. Now suppose that there is a pole of an approximant at a small distance, d , from the origin on the negative real axis. This pole will be a distance of order d^2 from the circular arc; however, by considering the third derivative and property (II.32) the residue will be less than $d^3 f_3$; hence the contribution to the value on the circular arc will be of order d and hence we may uniformly find the value on the circular arc and hence obtain convergence there and therefore apply Carleman's Theorem. It is, of course, none the less true that there may be a cut extending to the origin from minus infinity. To establish the proof of this uniqueness we will show only that the $\lim \{ [N, N] - [N, N-1] \} = 0$. The proof for the other sequence is similar. By Theorem 6, the $[N, N]$ and $[N, N-1]$ provide the best bounds and so this result is the most interesting one. First, by the same method as we obtained (II.21) we may obtain

$$\frac{P_{N+1}^{(-1)}(z)}{Q_{N+1}^{(-1)}(z)} - \frac{P_N^{(0)}(z)}{Q_N^{(0)}(z)} = \frac{(-z)^{2N+1} D(1, N) D(0, N)}{Q_N^{(0)}(z) Q_{N+1}^{(-1)}(z)} \quad (\text{II.37})$$

If we now add (II.21) for $N = N+1$ to (II.37) and subtract (II.18) for $j = 0$ we obtain, after some manipulation

$$\frac{Q_{N+1}^{(0)}(z)}{D(0,N+1)} = z \frac{Q_N^{(0)}(z)}{D(0,N)} + \frac{[D(1,N)]^2}{D(0,N+1)D(0,N)} \frac{Q_{N+1}^{(-1)}(z)}{D(1,N)} \quad (\text{II.38})$$

Similarly using (II.18) for $j = -1$, we can obtain

$$\frac{Q_{N+1}^{(-1)}(z)}{D(1,N)} = z \frac{Q_N^{(-1)}(z)}{D(1,N-1)} + \frac{[D(0,N)]^{-2}}{D(1,N-1)D(1,N)} \frac{Q_N^{(0)}(z)}{D(0,N)} \quad (\text{II.39})$$

If we now introduce the notation

$$\begin{aligned} \Omega_{2N-1} &= Q_N^{(-1)}(z) / [D(1,N-1) z^{N-\frac{1}{2}}] \\ \Omega_{2N} &= Q_N^{(0)}(z) / [D(0,N) z^N] \\ \omega_{2N} &= [D(0,N)]^2 / [z^{\frac{1}{2}} D(1,N-1) D(1,N)] \\ \omega_{2N+1} &= [D(1,N)]^2 / [z^{\frac{1}{2}} D(0,N) D(0,N+1)] \end{aligned} \quad (\text{II.40})$$

then (II.38) and (II.39) become [see, Frobenius (1881) for a more general discussion of these recursion relations]

$$\Omega_{p+1} = \Omega_{p-1} + \omega_p \Omega_p \quad (\text{II.41})$$

and the right hand sides of (II.21) and (II.37) become $\pm z^{\frac{1}{2}} / (\Omega_p \Omega_{p+1})$. We seek now to show that for real positive z the $\Omega_p \rightarrow \infty$ and thus the difference between the $[N,N]$ and the $[N,N-1]$ goes to zero. Equation (II.41) can be used to give

$$\Omega_{p+3} = (\omega_p + \omega_{p+2} + \omega_p \omega_{p+1} \omega_{p+2}) \Omega_{p+1} / \omega_p - \omega_{p+2} \Omega_{p-1} / \omega_p \quad (\text{II.42})$$

Thus, as all the ω_p are positive by throwing away the term $\omega_p \omega_{p+1} \omega_{p+2}$ and taking absolute values we obtain

$$|\Omega_{p+3}| - |\Omega_{p+1}| \geq \left| \frac{\omega_{p+2}}{\omega_p} \right| (|\Omega_{p+1}| - |\Omega_{p-1}|) \quad (\text{II.43})$$

If we compute $\Omega_0, \Omega_1, \Omega_2$, and Ω_3 from (II.40) and (II.1) and pick $k = \min \left[f_0^{-1}, z^{\frac{1}{2}}, f_2 / (f_0 f_2 - f_1^2) \right]$, then we have

$$|\Omega_2| - |\Omega_0| \geq k \omega_1, \quad |\Omega_3| - |\Omega_1| \geq k \omega_2 \quad (\text{II.44})$$

$$|\Omega_0| \geq k, \quad |\Omega_1| \geq k$$

Applying (II.43) recessively to (II.44) we obtain,

$$|\Omega_{2p}| \geq k(1 + \omega_1 + \omega_3 + \dots + \omega_{2p-1}) \quad (\text{II.45})$$

$$|\Omega_{2p+1}| \geq k(1 + \omega_2 + \omega_4 + \dots + \omega_{2p})$$

It is clearly sufficient for either the even or the odd series to diverge, hence it is sufficient for $\sum_{i=1}^{\infty} \omega_i$ to diverge. Now by Carleman's inequality [Carleman (1926)]

$$e \sum_{i=1}^{\infty} \omega_i \geq \sum_{p=1}^{\infty} (\pi_p)^{1/p}, \quad \pi_p = \prod_{i=1}^p \omega_i \quad (\text{II.46})$$

where e is the base of the natural logarithms. From (II.40) we find

$$\pi_{2p+1} = \frac{f_1 D(1, p)}{f_0 D(0, p+1)} z^{-p-\frac{1}{2}}, \quad \pi_{2p} = \frac{f_1 D(0, p)}{f_0 D(1, p)} z^{-p} \quad (\text{II.47})$$

Now the numerator of π_{2p+1} is the cofactor of f_{p+1} in the denominator.

We can make all other elements in the first row zero beside f_{p+1} by multiplying the other columns by the coefficients of the denominator of the

$[p+1, p]$ Padé approximant, where the signs of the coefficients are alternated and we normalize the coefficient of z^p to be +1. If $(-1)^{p+1} A$ is the multiplier of the last column, then for p odd

$$\begin{aligned} \pi_{2p+1} &= (f_1/f_0) \left\{ z^{p+\frac{1}{2}} \left[f_{p+1} - \left(f_{p+1} - A (f_{2p+2} - [p+1, p]^{(2p+2)}/(2p+2)!) \right) \right] \right\}^{-1} \\ &\geq \left\{ z^{p+\frac{1}{2}} f_{p+1} \right\}^{-1} \end{aligned} \quad (\text{II.48})$$

by property (II.32). Thus omitting all but the odd p terms

$$\sum_{i=1}^{\infty} \omega_i \geq e^{-1} \sum_{r=1}^{\infty} \left(\frac{f_1}{f_0 f_{2r} z^{2r-\frac{1}{2}}} \right)^{1/(4r-1)} \quad (\text{II.49})$$

Now as any f_r is given by some $[N, N-1]$ Padé approximant with the properties implied by (II.22), we may apply the Hölder inequality [see, for instance, Courant (1936)] to show $(f_0/f_s)^{1/s} \leq (f_0/f_r)^{1/r}$ for $r \leq s$. Thus the divergence of (II.49) is equivalent to the divergence of

$$\sum_{p=1}^{\infty} (f_0/f_p)^{1/(2p-1)} \quad (\text{II.50})$$

But, as we have seen that divergence implies the divergence of (II.45) which implies the coincidence via (II.22) of the $[N, N]$ and $[N, N-1]$ limits, which together with a similar proof for the other sequences completes the proof of Theorem 7.

We remark that if $f(z)$ is the limit of the $[N, N-1]$ sequence, then $\text{Im}[f(z)] < 0$ if $\text{Im}(z) > 0$, by (II.22). This observation, together with

Cauchy's Theorem, and a little knowledge of Stieltjes integration suffices to show that the determinantal conditions (II.10) alone suffice to assure the existence of a $\varphi(u)$ in (II.8) for

$$f(z) = \int_0^\infty \frac{d\varphi(u)}{1+uz},$$

$$\varphi(u) - \varphi(w) = \frac{1}{\pi} \lim_{\epsilon \rightarrow 0} \int_w^u \operatorname{Im}[f(-(v-i\epsilon)^{-1})] \frac{dv}{v} \quad (\text{II.51})$$

Theorem 7 gives a sufficient condition that φ be unique. The theory developed in this section gives, plainly, the solution to the moment problem in the interval 0 to ∞ , which problem served as the original motivation for its development.

We further remark that for any series which can be made a series of Stieltjes through the use of the transformations of Theorems 1 and 2, there must clearly be related results to those of this section. We note that if we transform according to Theorem 1 with A and B complex, then [pg. 419, Muir (1960)] if $\bar{D}$ are the determinants formed from the new coefficients, they are related to the original D's by

$$A^{-\frac{1}{2}n(n+1)} \bar{D}(1,n) = D(1,n) > 0 \quad (\text{II.52})$$

Thus, by the remark before Theorem 5, we can extend many of the results of this section for $[N,N]$ approximants to any series for which (II.52) holds with $A = f_1$. The principle results obtaining are (a) all the poles of all the $[N,N]$ approximants lie on the arc of a certain circle through the origin, are simple, and interlace, and (b) the Padé approximants converge uniformly for any positive distance from this circle.

II.C Reduction of Bounded Analytic Functions to Series of Stieltjes

The bulk of the material in this section is taken from Baker (1964). By and large there are two main types of questions which physicists seek to answer from perturbation theory calculations. The first is the value of some real function at a real point. If the series converges, then the results of this section provide a method of obtaining (rapidly) converging upper and lower bounds for the answer. In addition, these results are available for certain types of asymptotic series. The second type is the location of some real singularity of physical significance in a real function, e.g., critical point for a thermodynamic system. The results here are frequently sufficient to provide bounds on the location of the singularity, and some bounds on quantitative properties of the function in the neighborhood of the singular point.

The first step is to transform functions which are bounded from below into series of Stieltjes. To do this we need the following theorem due to Riesz (1911) and Herglotz (1911).

Theorem 8. The most general expression for a analytic function $f(z)$ which is real and has positive real part for $|z| < 1$, and which is real for real values of z is

$$f(z) = \frac{1+z}{1-z} \int_0^{2\pi} \frac{d\varphi(t)}{1 + \frac{4z}{(1-z)^2} \sin^2 \frac{1}{2}t} \quad (\text{II.53})$$

where φ is monotonic non-decreasing.

Proof: Suppose $f(z)$ is regular inside $|z| < 1$, then if we pick for

a contour the circle of radius $r < 1$, then if $|z| < r$, we have by Cauchy's Theorem

$$f(z) = \frac{1}{2\pi} \oint \frac{f(w) dw}{w - z} \quad (\text{II.54})$$

$$f^{(n)}(0)/n! = \frac{1}{2\pi i} \oint f(w) w^{-1-n} dw$$

If $w = re^{i\theta}$ and $f(re^{i\theta}) = h(\theta) + ig(\theta)$, then (II.54) becomes

$$f^{(n)}(0)/n! = \frac{r^{-n}}{2\pi} \int_{-\pi}^{\pi} [h(\theta) \cos n\theta + g(\theta) \sin n\theta] d\theta \quad (\text{II.55})$$

as $f^{(n)}$ is assumed real. We note that for n negative the integral in (II.55) vanishes, so if we multiply the $-n$ case by r^{-2n} and add it to (II.55) we get

$$f(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(\theta) d\theta, \quad f^{(n)}(0)/n! = \frac{r^{-n}}{\pi} \int_{-\pi}^{\pi} h(\theta) \cos n\theta d\theta, \quad n \geq 1 \quad (\text{II.56})$$

However, by assumption $h(\theta) > 0$. Thus if we define

$$\varphi(t) = \lim_{r \rightarrow 1} \frac{1}{\pi} \int_{-\pi}^t h(\theta) d\theta, \quad (\text{II.57})$$

it will be monotonic non-decreasing and if we multiply (II.56) by z^n and sum $n = 0$ to ∞ , we obtain

$$\begin{aligned} f(z) &= \frac{1}{2} \int_{-\pi}^{+\pi} \left(1 + 2 \sum_{n=1}^{\infty} \cos nt z^n \right) d\varphi(t) \\ &= \frac{1+z}{1-z} \int_{-\pi}^{\pi} \frac{d\varphi(t)}{1 + \frac{4z}{(1-z)^2} \sin^2 \frac{1}{2}t} \end{aligned} \quad (\text{II.58})$$

which complete the proof of Theorem 8.

Suppose now we have a function $f(z)$ which is regular $|z| < 1$ and bounded from below there. We may, therefore, add a constant M to give it a positive real part. Let us further make the change of variables

$$z = [(1+w)^{\frac{1}{2}} - 1] / [(1+w)^{\frac{1}{2}} + 1] \quad (\text{II.59})$$

then by Theorem 8 we find

$$\begin{aligned} F(w) &\equiv \left\{ f\left(\frac{(1+w)^{\frac{1}{2}} - 1}{(1+w)^{\frac{1}{2}} + 1}\right) + M \right\} (1+w)^{-\frac{1}{2}} \\ &= \int_{-\pi}^{-\pi} \frac{d\varphi(t)}{1 + w \sin^2 \frac{1}{2} t} \end{aligned} \quad (\text{II.60})$$

If we use $\sin^2 \frac{1}{2} t$ as a variable, then (II.60) implies (II.8) with an upper limit of 1 which in turn implies that for any $f(z) + M$ satisfying the conditions of Theorem 8, $F(w)$ is a series of Stieltjes with radius of convergence of at least unity. The transformation (II.59) has the effect of mapping the interior of the unit circle in the z -plane onto the cut w -plane $(-\infty, -1)$. The coefficients of $F(w)$ are readily given in terms of those of $f(z)$ by the following formula due to Gronwall (1932).

$$f(z) = \sum_{j=0}^{\infty} f_j (-z)^j, \quad F(w) = \sum_{j=0}^{\infty} F_j (-w)^j \quad (\text{II.61})$$

$$F_j = 4^{-j} \sum_{k=0}^j \binom{2j}{j-k} (f_k + \delta_{k,0} M)$$

where $\delta_{k,0}$ is the Kronecker delta function.

If we now form $[N, N+j]$ Padé approximants to $F(w)$, then by Theorem 6 for w real and positive the j even approximants and their derivatives all

lie above the correct values and the j odd approximants and their derivatives all lie below the correct value. Convergence from above and below is assured by Theorem 7 and it is monotonic in N (exception: $[N, N-1]$ slope which is even so a better lower bound than any other odd j slope).

We will now expand the class of real functions to which these results are applicable. Suppose that there is a branch point at $0 < +R_1 < 1$ and another one at $-1 < -R_2 < 0$. Let us distort the contour of the proof of Theorem 8 to run as follows: begin at R_1 , procede along the real axis (branch of function continued from upper half plane) to $+r < 1$, continue in upper half plane along arc $|z| = r$ to $z = -r$, then along negative real axis to $z = -R_2$ (branch of function continued from upper half plane), and close the contour by following the complex conjugate path to the one just described. Now as $f(z)$ is a real function for real z , it follows from Schwarz's principle of symmetry [Sec. 8.4, Copson, (1948)] that the discontinuity across the cuts is purely imaginary as $f(\rho+i\epsilon) = f^*(\rho-i\epsilon)$.

Equation (II.56) becomes

$$\begin{aligned}
 f(0) &= \frac{1}{2\pi} \int_{-\pi}^{+\pi} h(\theta) d\theta + \frac{1}{\pi} \int_{-r}^{-R_2} \text{Im}[f(\rho+i\epsilon)] \frac{d\rho}{\rho} + \frac{1}{\pi} \int_{R_1}^r \text{Im}[f(\rho+i\epsilon)] \frac{d\rho}{\rho} \\
 f^{(n)}(0)/n! &= \frac{r^{-n}}{\pi} \int_{-\pi}^{+\pi} h(\theta) \cos n\theta d\theta + \frac{1}{\pi} \int_{-r}^{-R_2} \text{Im}[f(\rho+i\epsilon)] [\rho^{-1-n} + \rho^{n-1} r^{-2n}] d\rho \\
 &\quad + \frac{1}{\pi} \int_{R_1}^r \text{Im}[f(\rho+i\epsilon)] [\rho^{-1-n} + \rho^{n-1} r^{-2n}] d\rho \quad n \geq 1
 \end{aligned}
 \tag{II.62}$$

If we multiply (II.62) by z^n and sum we get

$$\begin{aligned}
f(z) &= \frac{1+z}{1-z} \int_{-\pi}^{+\pi} \frac{d\phi(t)}{1 + \frac{4z}{(1-z)^2} \sin^2 \frac{1}{2}t} \\
&+ \frac{1}{\pi} \frac{(1+z)}{(1-z)} \int_{-1}^{-R_2} \frac{\rho^{-1} \operatorname{Im}[f(\rho+i\epsilon)] d\rho}{1 - \frac{(1-\rho)^2}{4\rho} \frac{4z}{(1-z)^2}} \\
&+ \frac{1}{\pi} \frac{(1+z)}{(1-z)} \int_{R_1}^1 \frac{\rho^{-1} \operatorname{Im}[f(\rho+i\epsilon)] d\rho}{1 - \frac{(1-\rho)^2}{4\rho} \frac{4z}{(1-z)^2}} \quad (\text{II.63}) \\
&= \frac{1+z}{1-z} \int_{-T_1}^{T_2} \frac{d\Phi(T)}{1 + \frac{4z}{(1-z)^2} T}
\end{aligned}$$

$$T_1 = \frac{1}{4} (1-R_1)^2 / R_1, \quad T_2 = \frac{1}{4} (1+R_2)^2 / R_2$$

where $\Phi(T)$ is defined by combining the other three integrals. $\Phi(T)$ defined this way will be monotonic non-decreasing if and only if $\rho^{-1} \operatorname{Im}[f(\rho+i)] \geq 0$ and $h(\theta) > 0$. If we define $F(w)$ from $f(z)$ as in (II.60) then it follows that

$$\hat{F}(u) = (1+T_1 u)^{-1} F[u(1+T_1 u)^{-1}] \quad (\text{II.64})$$

is a series of Stieltjes which is regular in the cut u -plane

$[-\infty, -(T_1 + T_2)^{-1}]$. From Theorem 1, we note that the $[N, N]$ approximants have all the various properties in the cut plane $(-\infty, -T_2^{-1})$, $(T_1^{-1}, +\infty)$.

We see from form (II.63) that we may, with impunity, let $R_2 \rightarrow 0$

$(T_2 \rightarrow +\infty)$ and retain $\hat{F}(u)$ as a series of Stieltjes. We have now proved the following Theorem.

Theorem 9. Let $f(z)$ be a function possessing all real, right-hand derivatives at $z = 0$. Let $f(z)$ be analytic $|z| \leq 1$, $\text{Im}(z) \neq 0$, and $\text{Re } f(z)$ bounded from below by a constant, $-M$, there. Further let $f(z)$ be such that $\lim_{\epsilon \rightarrow 0} x^{-1} \ln[f(x+i\epsilon)] \geq 0$ for real $-1 < x < +1$. Then $\hat{F}(u)$ is a series of Stieltjes, where

$$\begin{aligned}\hat{F}(u) &= (1 + T_1 u) \cdot F[u(1 + T_1 u)^{-1}] \\ F(w) &= \left\{ f\left((1+w)^{\frac{1}{2}} - 1\right) / \left[(1+w)^{\frac{1}{2}} + 1\right] + M \right\} (1+w)^{-\frac{1}{2}} \\ T_1 &= \frac{1}{4} (1 - R_1)^2 / R_1\end{aligned}\tag{II.65}$$

where R_1 is the minimum of 1 and the least upper bound of real regular points of f (which bound we assume to be positive). Convergent bounds to $f(z)$ are then given via $\hat{F}(u)$ by Theorems 6 and 7.

A couple of examples will suffice to indicate the nature of the requirement on the imaginary part. Basically it means that the type of singularity which is acceptable is one which vanishes for x positive no faster than linearly and diverges for x negative no faster than linearly. Functions such as

$$\begin{aligned}A - z/\log(\beta - z), & \quad 0 < \beta < 1 \\ A + (\beta - z)^\alpha & \quad 2N \leq \alpha \leq 2N + 1\end{aligned}$$

where N is an integer (positive, negative, or zero) are allowed. For z negative,

$$A + (\beta + z)^\alpha \quad 2N - 1 \leq \alpha \leq 2N$$

is, for example, allowed. A few simple manipulations using some physical

knowledge frequently suffice to cast a function with a physical singularity into manageable form.

In case the region in which the conditions of Theorem 9 are satisfied is not the unit circle, we know by Riemann's Theorem on conformal mapping (see, for example, Chapter 5, Carathiodory [1932]) that we can analytically map any simply connected domain with a reasonable boundary curve into the unit circle. If the other domain is a circle then a linear fraction transformation is a very convenient one to use.

II D. The Padé Conjecture and Some Consequences

The Theorems given in the preceding sections of this Chapter do not by any means exhaust the range of convergence of the Padé approximant method, as can be shown by examples. The following conjecture was put forward by Baker, et al (1961). So far, there are no known counter examples nor any valid proof or it. Even it, which greatly extends the range of convergence, probably does not give the complete range of convergence for the Padé approximant.

Conjecture. If $P(z)$ 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 point the function is assumed continuous when only points $|z| \leq 1$ are considered, 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 centers at these poles.

This conjecture can be weakened to read as follows:

PADE CONJECTURE. If $P(z)$ 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 point the function is assumed continuous when only points $|z| \leq 1$ are considered, then at least a subsequence of the $[N, N]$ Padé approximants are uniformly bounded (in N) in the domain formed by removing the interiors of small circles with centers at these poles and uniformly continuous at $z = +1$ for $|z| \leq 1$.

That this conjecture implies the conjecture of Baker, et al (1961) may be seen as follows. As $f(z)$ is regular at every boundary point of the convergence domain save $z = +1$, we can by the Heine-Borel covering Theorem (see, for example, Apostol [1957]) cover the closed set, L , $\delta \leq \arg z \leq 2\pi - \delta$, $\delta > 0$, of boundary points with a finite number of interiors of circles in which $f(z)$ is regular and hence enclose L interior to a circle about the origin in which $f(z)$ is regular. We can map this circle into the unit circle by a linear fractional transformation under which, by Theorem 1, the $[N, N]$ Padé approximants are invariant. As any prescribed point in the domain of the conjecture of Baker, et al (1961) is interior to some such a circle, we have by virtue of the Padé conjecture and Theorem 4, that there exists an infinite subsequence of Padé approximants which converge to a limit function which is regular everywhere in the domain of the conjecture of Baker, et al (1961) except at $z = +1$ where we obtain convergence by the uniform continuity of the subsequence. This completes the proof that the Padé conjecture implies the conjecture of Baker, et al (1961).

By a Theorem of Lindelöf (1915) [See Chapter XIII, Dienes (1957) for a good discussion of this type of problem.] any function with the regularity properties (which imply boundedness) described by the Padé conjecture, is continuous at $z = +1$ and hence the continuity assumption is only an assumption on the Padé approximants and does not restrict the class of functions. We shall now list a few quasi-Theorems of Baker, et al (1961) which follow from the Padé conjecture. By "converge" we shall mean that at least a subsequence converges.

Definition: A function $f(z)$ will be said to be of type I if a straight line can be drawn in the z plane, such that $f(z)$ is regular, except for a finite number of poles, everywhere in the open half plane containing the origin and continuous in this half plane at infinity. This half plane will be called $R_1(f)$. The region $R_2(f)$ is defined as the closed half plane containing the origin and bounded by a line parallel to the one defining $R_1(f)$ and between it and the origin, except for the interior of certain small circles containing the finite number of poles.

Quasi-Theorem 1. If $f(z)$ is of type I, then the $[N,N]$ Padé approximants "converge" to it uniformly in $R_2(f)$.

Proof: By Theorem 1 we may study equivalently $g(w) = f[A^{-1}w/(1-w)]$. If the line bounding $R_2(f)$ passes through the point $-1/2A$, then the transformation $w = Az/(1 + Az)$ maps $R_2(f)$ onto the unit circle. The point at infinity is carried into $w = +1$ and the origin is left unchanged. Thus $g(w)$ is regular in the unit circle except for a finite number of poles in $|w| \leq 1$ and continuous at $w = +1$. Hence by the Padé conjecture the

$[N,N]$ Padé approximants for $g(w)$ "converge" uniformly to $g(w)$ in $|w| \leq 1$ except for the interior of small circles centered at the poles. Hence $[N,N]$ Padé approximants for $f(z)$ "converge" uniformly to $f(z)$ in $R_2(f)$.

Quasi-Corollary 1: The $[N,N]$ Padé approximants "converge" uniformly to any function $f(z)$ in any circle containing the origin as an interior point in which $f(z)$ is regular except for a finite number of poles and for one boundary point where, however, it is continuous.

The interiors of certain small circles about the poles are excluded from this region of uniform convergence.

Proof: By Theorem 1 and Quasi-Theorem 1 it suffices to show that any such circle can be mapped onto the unit circle by the transformation of Theorem 1. For a circle with center at C and of radius R , the transformation

$$w = e^{i\theta} z / [R(1 - |C/R|^2) + \bar{C}z/R] \quad (\text{II.66})$$

accomplishes this mapping, where θ is such that the non-regular boundary point, if any, is carried into $w = +1$. This corollary now follows from the arguments establishing Quasi-Theorem 1.

Quasi-Corollary 2: The $[N,N]$ Padé approximants "converge Quasi-Theorem 1 uniformly to $f(z)$ in the union $R_3(f)$ of any finite number of regions $R_2(f)$ and circles described in Quasi-Corollary 1. Note that $f(z)$ need not be of type I for either this or the previous Quasi-Corollary to hold.

Proof: By Quasi-Theorem 1 and Quasi-Corollary 1 $f(z)$ "converges Quasi-Theorem 1 at every point of this union. As the union of a finite number of closed sets is again a closed set, $f(z)$ must converge uniformly in the union.

These results prove that if $f(z)$ is of type I and its only non-polar singularity in the finite plane is a branch point, then the $[N,N]$ Padé approximants "converge" to $f(z)$, cut on a ray. If there are two branch points which may be connected by a cut to define a single valued function, then this cut lies on the arc of a circle through the two points and the origin. If there are n branch points which may be connected by cuts to define a single valued function, then these cuts lie inside or on a polygon whose sides are arcs of circles through the origin and pairs of branch points, and which contains all the branch points.

Definition: A function $g(w)$ will be said to be of type II if

$$\begin{aligned} T(w) &= Aw^n/p_n(w) \\ g(w) &= f[T(w)] \end{aligned} \tag{II.67}$$

where $f(z)$ is of type I and $p_n(w)$ is a polynomial of degree at most n .

Theorem 10. The $[nN + r, nN + s]$ Padé approximants, $r + s \leq n - 1$, to $g(w)$ are given by $P_N[T(w)]/Q_N[T(w)]$ where $P_N(z)/Q_N(z)$ is the $[N,N]$ Padé approximants to $f(z)$.

Proof: The proof is the same as for Theorem 1, which is seen to be a special case of this Theorem. Here, however, the approximant is good through terms of order $w^{n(2N+1)-1}$ so that the Padé table is not normal.

If the Padé table is normal then, in simplest terms, the numerator and denominator of the $[N,M]$ are of degree M and N respectively. For a fuller discussion on normality see Wall (1948) and Padé (1892).

Quasi-Corollary 3: The $[N,N]$ Padé approximants for $g(w)$ "converge"

uniformly to $g(w)$ in $T^{-1}[R_3(f)]$.

Proof: By Theorem 10, the Padé approximants to $g(w)$ are identically obtained from those of $f(z)$ which by quasi-corollary 2 "converge" uniformly in $R_3(f)$. Hence those for $g(w)$ must "converge" uniformly to $g(w)$ in the inverse mapping, T^{-1} , of $R_3(f)$ in the z plane onto the w plane. It should be noted that this inverse mapping is 1 to n . It should also be noted that this corollary exhibits a class of functions for which the $[N,N]$ Padé approximants "converge," although the Padé table is not normal.

We remark that for a regular function, we need only assume that there is an infinite subsequence of $[N,N]$ Padé approximants which is free of poles in a circular region of radius R to prove some convergence. For the largest the $[N,N]$ approximant can be is to have an N th order pole at the nearest boundary point. Then the denominator is larger than or equal to $(1 - r/R)^N$ for $|z| \leq r$. It follows from (II.1) that the numerator is bounded by $W(1 + r/R)^N$ where w is a bound for $f(z)$ and its partial sums in $|z| \leq r$. Hence the $[N,N]$ is bounded by

$$|[N,N]| \leq W \left(\frac{R+r}{R-r} \right)^N \quad \text{for } |z| \leq r \quad (\text{II.68})$$

The absolute value of the difference between $[N,N]$ and $f(z)$ is for $|z| = \rho \leq r$ bounded by (summing a geometric series and using Cauchy's inequalities for derivatives)

$$\left[\frac{\rho W}{r - \rho} \right] \left[\left(\frac{\rho}{r} \right)^2 \frac{(R+r)}{(R-r)} \right]^N \quad (\text{II.69})$$

which goes to zero or is bounded if we choose, for example, $r \geq 2\rho$,

$R = 5r/3$. Thus if the $|z| < R$ is free from poles, by Theorem 3 the $[N,N]$

Pade' approximants must converge at least in $|z| < 0.3 R$.

CHAPTER III

Special Function Examples

In this Chapter we give a number of examples of the results obtained by computing Padé approximants to special cases. The bulk of this Chapter is drawn from the work of Baker, et al (1961). The examples in Table I relate to the possibility of an extension of the Padé conjecture.

The functions listed in Table I are

$$\begin{aligned} a(x) &= (1 - e^{-x})/x \\ b(x) &= (1 - e^{-x}) \\ u(x) &= (1 + x^2)^{\frac{1}{2}}/(1 + x) \\ v(x) &= [(1 + x)/(1 + \frac{1}{4}x)]^{\frac{1}{2}} + \pi - 2 \tan^{-1} (5/12) \\ &\quad - 2 \tan^{-1} (2.4 - 2.6 x) \\ t(x) &= [((1 + x + x^2) (1 + 2x))^{\frac{1}{3}} - 1]/x. \end{aligned} \tag{III.1}$$

The function $a(x)$ is of type I but not a series of Stieltjes. Although the limit as z tends to infinity on any ray in the left half plane is infinite, one seems from Table I that the Padé approximants converge at infinity like $1/(N + 1)$. The function $b(x)$ is continuous at infinity on any ray in the right half plane, save the imaginary axis. As is evident from Table I, the $[N, N]$ Padé approximants for $b(x)$ do not converge at infinity (although a Cesaro mean of them does). From this pair of examples it seems

to us likely that continuity in $R_2(f)$ at infinity is necessary for convergence at infinity of the $[N,N]$ Padé approximants, and that it is also sufficient.

The function $u(x)$ has branch points at $x = \pm \sqrt{-1}$. This excludes $u(x)$ from Type I; $u(x)$ is also not of type II as the coefficient of x in the power series is not zero. Since the $[N,N]$ Padé approximants for $u(x)$ appear to converge to $\sqrt{2} \pm 1$ as N is even or odd, the sequence of $[N,N]$ Padé approximants do not converge at infinity. The function $v(x)$ has branch points at $x = -1$, -4 , and $(12 \pm 5\sqrt{-1})/13$, and is neither of type I nor type II. The Padé approximants for $v(x)$ appear to be converging at infinity. This pair of examples shows that the $[N,N]$ Padé approximants may or may not converge, if one is not inside the region $R_3(f)$.

The function $t(x)$ has three branch points on the line $\text{Re}(x) = -\frac{1}{2}$. The equations for the $[1,1]$, $[4,4]$, $\dots$, $[3N+1, 3N+1]$ $\dots$ Padé approximants are inconsistent and so those approximants do not exist. However, the other Padé approximants do exist and appear to converge to the correct limit as $N \rightarrow \infty$. This example proves that the entire sequence need not converge, and supports the conjecture that there is always at least a subsequence which does. In this regard we remark that in some cases studied by Baker, et al (1961), occasionally there appear poles in a region which had previously fairly well converged. For example the $[10,10]$ Padé approximant to

$$s(x) = \pi - 2 \tan^{-1} (2.4) - 2 \tan^{-1} [(5 - 13x)/12] + (1 - x)^{-\frac{1}{2}} \quad (\text{III.2})$$

has a pole at about $x = -24$, even though all the singularities of $s(x)$ are in the right half plane and all the poles of the previous approximants were too. Further, the $[16,16]$ has a pole at about -11.4 and the $[18,18]$ at about -64 . The residues at these poles are quite small. It is not known whether this behavior is eventually excluded for N large enough or whether a subsequence must be deleted for this case. The values of $[N,N]$ Padé approximants at infinity do appear to be converging in accordance with Theorem 4.

In Table II we give a number of examples of type I which illustrate how well low order $[N,N]$ Padé approximants can work. The function $w(x)$ in Table II is

$$w(x) = \pi - 2 \tan^{-1}(5/12) - 2 \tan^{-1}(2.4 - 2.6x). \quad (\text{III.3})$$

The functions $\text{sech} \sqrt{x}$ and $\tanh^2 \sqrt{x}$ have essential singularities at infinity, since infinity is a limit point (on the Riemann sphere) of poles. Neither is a series of Stieltjes. The Padé approximants to these functions are expected to converge uniformly in any $R_2(f)$. From Table II we see that the error at infinity is decreasing by about a factor of (-5) at each step. The closest poles are very accurately located by the low order Padé approximants. For instance, the $[3,3]$ Padé approximant for $\text{sech} \sqrt{x}$ locates the closest pole to an accuracy of 8 decimal places and the second nearest one to within 1 part in 200. The $\tanh^2 \sqrt{x}$ has double poles. The Padé approximants simulate them by two very close single poles. For instance, the $[4,4]$ approximant has its closest "double" pole at

$$x = -2.4674003 \pm 9.6396 \times 10^{-3} \sqrt{-1} \quad (\text{III.4})$$

which may be compared with

$$x = - (\pi/2)^2 = - 2.4674010. \quad (\text{III.5})$$

The function $\exp[- x/(1 + x)]$ is regular in the whole complex plane, except for an essential singularity at $x = - 1$. The increase in accuracy of the approximants at infinity is about 2 decimal places per unit increase in N . The value given by the $[4.4]$ approximant is off only one part in the eighth decimal place. The excellent results obtained in this case illustrate the fact that an essential singularity is, except in its neighborhood, almost identical with a pole of finite order. The poles of the Padé approximant tend, as one would expect, to cluster about the essential singularity.

The function $(\tan^{-1} \sqrt{x})^2$ is regular in the complex plane which is cut from $x = - 1$ to $-\infty$. The limit as $x \rightarrow \infty$ in $R_2(f)$ is $(\pi/2)^2$. As a cut represents a discontinuity in a function, and the Padé approximants are meromorphic functions, the only way they can represent a cut function is to cluster poles along the cut. That a cut may be thought of as a line of poles can be seen from a standard formula (No. 120, Dwight (1947)) for the $\tan^{-1} x$.

$$\tan^{-1} x = x \int_1^\infty \frac{dz}{x^2 + z^2} \quad (\text{III.6})$$

Now as is well known, the cut, and by Cauchy's Theorem the path of integration in Eq. (III.6), need not follow a ray to infinity. However, when there is only a single branch point in the finite plane, the $[N,N]$ Padé approximants "converge" to the function defined by cutting along a ray. This behavior follows from Quasi-Corrolary 2. (As well as, for this case, from Theorem 5, via (II.51) which implies $(\tan^{-1} \sqrt{x})^2/x$ is a series of Stieltjes.)

In this case all the poles of the Padé approximants lie on the negative real axis between -1 and $-\infty$. The convergence of the sequence of Padé approximants is assured by Theorem 7 and, at infinity is directly comparable to that of the well known series [No. 48.2, Dwight (1947)]

$$(\pi/2)^2 = \frac{3}{2} \left[1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \dots \right] \quad (\text{III.7})$$

for the limiting value as x tends to infinity.

For the function $[(1 + 2x)/(1 + x)]^{\frac{1}{2}}$ (which by (II.51) is the reciprocal of a series of Stieltjes), all the poles of the Padé approximants lie on the ray from the branch point at $x = -\frac{1}{2}$ between it and the branch point at $x = -1$. The convergence of the $[N, N]$ Padé approximants at infinity in this case is good. The error for $N = 5$ is only one part in the ninth place.

The function $w(x)$ has branch points at

$$x = (12 \pm 5\sqrt{-1})/13. \quad (\text{III.8})$$

All the poles of the Padé approximants lie on the arc of a circle through the origin and the two branch points. The convergence here is again good; the $[4, 4]$ approximant is off only 3 parts in sixth place. The function $w(x)$ is not a series of Stieltjes nor related to one by Theorems 1 or 2, since the $[N, N]$ alternate above and below the correct answer.

Whether a function is represented by a convergent power series or only by an asymptotic expansion, one may formally form Padé approximants to it. We shall do this for some examples and investigate the error in the Padé

approximants at infinity.

In Table III we have listed the values of the limit as $z \rightarrow \infty$ of the $[N, N]$ Padé approximants to several functions which tend to a constant on every ray except the negative real axis as $z \rightarrow \infty$. The first function has the asymptotic expansion

$$\int_0^\infty \frac{e^{-t} dt}{1 + xt} = 1 - 1! x + 2! x^2 - 3! x^3 + \dots \quad (\text{III.9})$$

which was first studied by Euler [see p. 26 ff, Hardy (1956)]. This function, and also the second one are series of Stieltjes and so the $[N, N]$ Padé approximants must converge. We choose the $[N, N]$ Padé approximants for the first two functions even though they tend to zero at infinity because the first goes like $\ln x/x$ and the second like $\sqrt{\pi/x}$. We note in passing that, using the $[6, 6]$ Padé approximant for the first function, which is derived from the first 13 terms of the asymptotic series, we calculate the value of Euler's divergent series (the function at $x = 1$) to be 0.5968 which may be compared with the exact value of 0.5963 and Lacroix's (Hardy, 1956) value of 0.5992 which he obtained from the first 13 terms through the use of three Euler transformations. The rate of convergence of the Padé approximants at $x = 1$ is essentially the same as at $x = \infty$; namely, the error is proportional to $1/(n + 1)$. In the case of the second function, one easily deduces from Stirling's approximation for the gamma function that the error at infinity tends to zero like $(N)^{-\frac{1}{2}}$. The poles of the Padé approximants to both the first two functions lie on the negative real axis, as they must.

The third example is the reciprocal of a series of Stieltjes and the fourth is x over a series of Stieltjes. The third function appears to converge faster than $(N)^{-\frac{1}{2}}$ but more slowly than N^{-1} and the fourth about like $(N)^{-\frac{1}{2}}$. For all the cases presented in Table III the coefficients of the power series ultimately diverge like $n!$ and the Padé approximants converge very slowly.

When a function does not tend to infinity like a constant times some integral power of z , then no Padé approximant can represent it in the neighborhood of infinity. Examples of such functions can be found among those with an essential singularity at infinity. We shall discuss several examples of functions with essential singularities at infinity from the point of view of the question of how far can one proceed toward infinity with a given order approximant, rather than from the viewpoint of what order approximant is required to give a specified accuracy at a given point.

The examples listed in Table IV have essential singularities of infinity but no singularities in the finite plane. The $[N,N]$ Padé approximants have poles in the finite plane. We view the location of the spurious poles of the Padé approximant as defining a neighborhood of infinity in which the Padé approximant breaks down. In Table IV we list the value of the $[N,N]$ Padé approximant at the magnitude of the smallest root of the denominator.

For the function (see also, Luke (1957, 1958)) $b(x)$ the results are quite good. The odd approximants give exactly the correct value

of the limit as $x \rightarrow +\infty$ and the even ones appear to converge rapidly. The smallest root of the denominator recedes to infinity in a manner almost linear with N , as observed first by Luke (1957).

The second function in Table IV manifests a more typical behavior. (We treat this function and the last one in Table IV as functions of x^2 .) Comparison with tabulated values of the sine integral reveal that starting with the $[3,3]$ approximant on, an accuracy of 0.2 is obtained and not improved upon. The magnitude of the smallest root in this case as in the first one appears to increase linearly with N . As the radius within which the power series yields the function value to within a given accuracy also increases linearly with N in these two cases (though with a smaller coefficient for the same accuracy), it appears to be the case that for these functions we may proceed toward infinity a distance which is proportional to N .

The third function in Table IV illustrates much the same behavior as the second. However it illustrates a phenomenon mentioned above (III.2) which occurs occasionally in practice; namely, the appearance of a pole in a region which had already started to converge. The $[4,4]$ Padé approximant very nearly fills a block in the Padé table of order 1. The smallest root of the $[5,5]$ approximant is very nearly cancelled by a root of the numerator.

$$\text{Root of numerator} = 2.8852000$$

$$\text{Root of denominator} = 2.8851989 \quad (\text{III.10})$$

If we do not go as far as the first pole, but only to $x = \frac{1}{2}(N + 1)$, we

find much improved convergence in these examples. In fact, the [3.3] approximant gives the function value with an error of less than 1 part in 10^4 , for at least the first two functions, which are tabulated.

In cases where the exact location of the cuts is not determined by the Theorems of Chapter II, it is of interest to investigate the behavior of the $[N,N]$ Padé approximants outside of any $R_3(f)$. We have formed Padé approximants to several functions with more than two branch points to investigate this question. We have studied among others, $v(x)$, $t(x)$, $s(x)$, $c(x)$, and $d(x)$ where

$$c(x) = [(1-x)/(1-\frac{1}{4}x)]^{\frac{1}{2}} + \pi - 2 \tan^{-1}(2.4) - 2 \tan^{-1} [(5-12x)/13]$$

$$d(x) = \left\{ [(1+x+x^2)(1+x)]^{\frac{1}{3}} - 1 \right\} / x \quad (\text{III.11})$$

and the others have been defined previously. In these cases the poles of the Padé approximants do not all lie on any curve, but rather if S is the set of all the limit points of the poles of the Padé approximants, S appears to form a simple curve. Our evidence for the following conclusions about S is somewhat inconclusive. It appears, however, that the Padé approximants always cut (for functions not of type II, $n > 1$) on the arc of a circle through the origin. When this circle is not completely determined by Quasi-Corollary 2, then the slope of the cut at the branch point at which it originates appears to be such that the cut lies in the interior of the angular region available to it. In the case of $f(x)$ and $d(x)$, we obtain ϵ shaped cuts, a trident shaped cut for the case of $c(x)$, and an infinitely long handled trident shaped cut for the case of $s(x)$.

It appears to us to be the case that the $[N,N]$ Padé approximants "converge" at every point of the complex plane which is not a singularity of $f(z)$ or on a cut, provided the cuts leave the complex plane a connected region. (In case the cut plane is not connected, then the Padé approximants "converge" in the connected portion which contains the origin. For instance the $[N,N]$ Padé approximants to a function with the unit circle as a natural boundary can only be expected to "converge" inside the unit circle.) It should be noted that for the function $u(x)$ given in Table I, the point at infinity necessarily lies on a cut. Also, the Padé approximants may "converge", as we have seen, at certain types of singularities.

CHAPTER IV

Some Applications to Physical Problems

In this chapter we describe a few of the applications to physical problems which have been made and suggested. None of these applications has yet been exhausted, and the first two are to date only prefaces to more work yet to be done. An additional application to the S matrix is given in the article by J. L. Gammel in this volume.

IV A. An Optical Theorem and the Mandelstam Representation

In the last few years the interpretation of elementary particles has been approached through the use of a few general principles drawn from experience with previous field theories and attempts a systematic study of the mathematical implications of these assumptions. In this approach, one does not consider the details of the interaction, or the divergences which may occur in going from "bare" to "dressed" particles, but merely the consequences of certain physical restrictions on the transformation from initial to final states. One of the main results which has arisen from this approach has been various dispersion relations [e.g. Klein (1961)], which follow primarily from Lorentz invariance and microscopic causality. The dispersion relations apparently do not exhaust the content of the analytic properties implied by the physical principles assumed. We shall show in this section that in fact the index of refraction for classical materials, and the scattering amplitude for quantum-field-theoretic processes are closely related to series of Stieltjes and hence all the results of the theory in section II B, on convergence, bounds, etc. are applicable to these problems. (Padé approximation can, of course, be thought of as a refined form of Polology.)

Let us now consider the index of refraction in a non-magnetic, dispersive medium. Referring to a text on Electromagnetism (See, for example, Slater and Frank (1947) chaps. 8 and 9) we note that the index of refraction, n (it may be complex), is given by

$$n^2(\omega) = 1 + \sum_k (N_k c^2 / m \epsilon_0) (\omega_k^2 - \omega^2 - 2i\omega/\tau_k)^{-1} \quad (\text{IV.1})$$

where the ω_k are associated frequencies and the τ_k the corresponding relaxation times. Let us introduce $\Omega = -i\omega$; $\Omega = \Psi + i\theta$; then (IV.1) becomes

$$n^2 = 1 + \sum_k \frac{(N_k e^2 / n \epsilon_0) [\omega_k^2 + \Psi^2 - \theta^2 + 2\Psi/\tau_k - 2i\theta(1/\tau_k + \Psi)]}{[\omega_k^2 + \Psi^2 - \theta^2 + 2\Psi/\tau_k]^2 + 4\theta^2 (1/\tau_k + \Psi)^2} \quad (\text{IV.2})$$

Aside from the properties (i) $\lim_{\omega \rightarrow \infty} n=1$, (ii) $n(\omega) = n(-\omega)$ for ω real,

and (iii) n is regular in $\text{Im}(\omega) > 0$ necessary for the derivation of dispersion relations (Kramers-Krönig relation here, see Kramers, [1927]), n^2 has by (IV.2) the following additional property

$$\begin{aligned} \text{Im}(n^2) < 0 & \quad \text{Im}(\Omega) > 0 \\ \text{Im}(n^2) > 0 & \quad \text{Im}(\Omega) < 0 \end{aligned} \quad (\text{IV.3})$$

when $\text{Re}(\Omega) > -1/\tau$, where τ is the largest $\tau_k < \infty$. If we now introduce a new variable $\hat{\Omega} = (2\Omega + \tau\Omega^2)$, which maps the cut $\hat{\Omega}$ -plane $(-\infty, 0)$ into the cut $(-\tau^{-1}, 0)$ $\hat{\Omega}$ -half-plane $\text{Re}(\Omega) > -\tau^{-1}$. Property (IV.3) coupled with (II.51) suffice to show that n^2 is a series of Stieltjes in terms of $\hat{\Omega}$, or, in terms of ω ,

$$n^2(\omega) = \int_0^\infty \frac{d\varphi(t)}{1 - \omega(2i + \tau\omega)t} \quad (\text{IV.4})$$

where φ is monotonic, non-decreasing. Theorem 7 assures us that the $[N, N+j]$ Padé approximants in $[-\omega(2i + \tau\omega)]$ certainly converge in the physical region, ω real.

Moving on to the field theoretic case,

the validity of the Mandelstam representation (Mandelstam, 1958) suffices to make the scattering amplitude in both the energy and momentum-

transfer variables differ from a series of Stieltjes by an integral transformation. If A is the scattering amplitude for, for example, pion-nucleon scattering, then A is of the form (Mandelstam, 1958)

$$\begin{aligned}
 A = & \frac{g^2}{m^2 - s} + \frac{g^2}{\mu^2 - s_c} + \frac{1}{\pi^2} \int_{(M+\mu)^2}^{\infty} ds' \int_{\mu^2}^{\infty} dt' \frac{A_{13}(s', t')}{(s' - s)(t' - t)} \\
 & + \frac{1}{\pi^2} \int_{(M+\mu)^2}^{\infty} ds'_c \int_{\mu^2}^{\infty} dt' \frac{A_{23}(s'_c, t')}{(s'_c - s_c)(t' - t)} \\
 & + \frac{1}{\pi^2} \int_{(M+\mu)^2}^{\infty} ds' \int_{(M+\mu)^2}^{\infty} ds'_c \frac{A_{12}(s', s'_c)}{(s' - s)(s'_c - s_c)}
 \end{aligned} \tag{IV.5}$$

in the energy and momentum transfer variables s , s_c , t . (The usual subtractions are implied, if necessary.) They are not independent but satisfy

$$s + s_c + t = 2(M^2 + \mu^2) \tag{IV.6}$$

Thus let us select any one, call it u , of them to study and any other as a parameter, call it v . In the chosen variable A becomes, through some obvious manipulations, of the usual dispersion relation form

$$A(u, v) = \frac{1}{\pi} \int_{X(v)}^{\infty} \frac{d\alpha(u', v)}{u' - u} - \frac{1}{\pi} \int_{\infty}^{-Y(v)} \frac{d\beta(u', v)}{u' - u} \tag{IV.7}$$

where the pole terms are included in the integrals. As $d\alpha$ and $d\beta$ are (the "absorptive parts") positive Mandelstam, (1958), [See also, Symanzik (1960), Weinberg (1961), Foldy (1964), and Balachandran (1964)] we can put (IV.7) in the form

$$A(u, v) = \int_{-(X(v))^{-1}}^{(Y(v))^{-1}} \frac{d\phi(w, v)}{1+wu} \quad (IV.8)$$

Form (IV.8) suffices to prove that the determinants $D(0, n)$ are positive (Eq. II.9) and hence that the poles of the $[N, N+j]$ Padé approximants are all real with positive (absorptive) residues. For the purposes of obtaining upper and lower bounds, it is convenient to convert (IV.8) into a series of Stieltjes. One can easily show that both

$$\frac{X(v) + Y(v)}{X(v) + Y(v)(1+\xi)} \leq A\left(\frac{X(v)Y(v)\xi}{X(v) + Y(v)(1+\xi)}, v\right) \quad (IV.9a)$$

$$\frac{X(v) + Y(v)}{Y(v) + X(v)(1+\xi)} \leq A\left(\frac{-X(v)Y(v)\xi}{Y(v) + X(v)(1+\xi)}, v\right) \quad (IV.9b)$$

are of the form of (II.51) with an upper limit of 1.0, and hence are series of Stieltjes. It follows from Theorem 5 that for $0 \leq u \leq X(v)$, ($0 \leq \xi \leq \infty$)

$$\begin{aligned} [N, N-1] &\leq (IV.9a) \leq [N, N] \\ [N, N-1]' &\leq (IV.9a) \leq [N, N]' \end{aligned} \quad (IV.10a)$$

and for $-Y(v) \leq u \leq 0$, ($0 \leq \xi \leq \infty$)

$$\begin{aligned} [N, N-1] &\leq (IV.9b) \leq [N, N] \\ [N, N-1]' &\leq (IV.9b) \leq [N, N]' \end{aligned} \quad (IV.10b)$$

where the Padé approximants are the ones for the function being bounded.

These results should be a more powerful method to use the analytic structure of the scattering amplitude than those used previously (e.g., Frazer, [1961]) as it depends explicitly on the "absorptive parts" being non-negative definite, which, for instance, the usual relation expressing the real part in terms of the imaginary part of the scattering amplitude

does not. The analytic continuation into the cut, complex u -plane is obtained as Theorem 7 proves convergence at every point of it. One does not, however, directly obtain a continuation onto other, non-physical Riemann sheets.

IV B. The Nature of the Forward Scattering Amplitude Series in the Coupling Constant

We point out in this section that, below all thresholds, the forward scattering amplitude is a series of Stieltjes (or the two-sided extension; see the remark after (II.11)) in the coupling constant. The reasoning here is similar to that of Baker, et al (1963). Let v be the two-body hermitian potential operator, and $1/b = P/(H_0 - E - i\epsilon)$ where H_0 is the two-body kinetic energy operator, E the unperturbed energy of the initial state, and P a projection operator which is zero for the initial state and one otherwise. The forward scattering amplitude is then, (Wu and Ohmura, 1962)

$$f(0) = - \frac{1}{4\pi} \langle \psi | (v\lambda - v \frac{1}{b} v\lambda^2 + v \frac{1}{b} v \frac{1}{b} v\lambda^3 - v \frac{1}{b} v \frac{1}{b} v \frac{1}{b} \lambda^4 + \dots) | \psi \rangle \quad (\text{IV.16})$$

where λ is regarded as an expansion parameter, ψ is an unperturbed eigenstate of H_0 with eigenvalue E . Below threshold (E negative) we may set $\epsilon=0$ as $1/b$ is bounded and non-negative definite. Hence in this energy range $1/b$ has a well defined real square root. Let us therefore rewrite (IV.16) as

$$f(0) = - \frac{1}{4\pi} \left[\langle \psi | v | \psi \rangle \lambda - \lambda^2 \langle \psi | \left\{ 1 - ((1/\sqrt{b})v(1/\sqrt{b})) \right\} \right. \\ \left. + \left(((1/\sqrt{b})v(1/\sqrt{b}))^2 \lambda^2 - \dots \right) | \psi \rangle \right] \quad (\text{IV.17})$$

where $w = (1/\sqrt{b})v\psi$. As the operator $w = (1/\sqrt{b}) v (1/\sqrt{b})$ is hermitian with real eigenvalues w_k , we may rewrite (IV.17) as

$$f(0) = -\frac{1}{4\pi} \left[\langle \Psi | v | \Psi \rangle \lambda - \lambda^2 \left\{ \sum_k (1 - w_k \lambda + w_k^2 \lambda^2 - w_k^3 \lambda^3 + \dots) |a_k|^2 \right\} \right] \quad (\text{IV.18})$$

By the normalization of Ψ and the assumed finiteness of the perturbation series terms, there exists a bounded, non-decreasing function $\varphi(u)$, such that, if we denote by $(-1)^p c_p$ the term from the $(p+2)$ nd power of λ

$$c_p = \int_{-\infty}^{+\infty} u^p d\varphi(u) \quad (\text{IV.19})$$

or

$$f(0) = -\frac{\lambda}{4\pi} \langle \Psi | v | \Psi \rangle + \int_{-\infty}^{+\infty} \frac{\lambda^2 d\varphi(u)}{1 + u\lambda} \quad (\text{IV.20})$$

If we assume that v is a bounded operator, or, less restrictively, that w (for E fixed) is bounded, then the limits of the integral in (IV.20) become finite and through manipulations similar to (IV.9) we may cast $[f(0) + \lambda \langle \Psi | v | \Psi \rangle / 4\pi] / \lambda^2$ into a series of Stieltjes. All the results of Section II B now apply to the subtracted forward scattering amplitude as a function of the coupling constant λ . In particular for E below threshold we obtain monotonic convergence from above and below by the j even and j odd $[N, N+j]$ Padé approximants. As we saw in the previous section, the forward scattering amplitude is also analytic in the cut E -plane so we may continue the E -below-threshold results to the outgoing or incoming wave solutions for physical E , accordingly as we use points in the upper or lower-half plane.

If v is a repulsive potential operator such that a real $\sqrt{v}$ may be defined, then we can rewrite (IV.16) as

$$f(0) = -\frac{\lambda}{4\pi} \langle \sqrt{v}\Psi | \left(1 - (\sqrt{v} \frac{1}{b} \sqrt{v})\lambda + (\sqrt{v} \frac{1}{b} \sqrt{v})^2 \lambda^2 - \dots \right) | \sqrt{v}\Psi \rangle \quad (\text{IV.21})$$

As $(\sqrt{v} \frac{1}{b} \sqrt{v})$ is a hermitian, non-negative definite operator, we may write

$$f(0) = -\lambda \int_0^\infty \frac{d\Phi(u)}{1+u\lambda} \quad (\text{IV.22})$$

where $\Phi(u)$ is a bounded monotonic non-decreasing function, and hence $-f(0)/\lambda$ is a series of Stieltjes. An example of this result was previously given by Baker and Gammel (1961). In the limit E tends to the elastic threshold ($E = 0$) the forward scattering amplitude tends to minus the scattering length. For a repulsive square-well we have

$$a = c \left[1 - \frac{\tanh \lambda^{\frac{1}{2}}}{\lambda^{\frac{1}{2}}} \right] \quad (\text{IV.23})$$

where c is the range and $\lambda = 2\mu Vc^2/\hbar^2$ is the coupling constant (V is height of the well). The $[N,N]$ Padé approximants ($[N,N-1]$ to $-f(0)/\lambda$) converge rapidly, even for the hard core limit. The result for $\lambda = \infty$ is

$$[N,N](\infty) = \{1 - [(2N+1)(N+1)]^{-1}\}c \quad (\text{IV.24})$$

IV C. The Ising Model and the Lattice Gas

The problem of the Ising model is probably the one which has been most thoroughly explored by means of the Padé approximant method. As a very complete catalogue of recent results has been given in a review article by Fisher (1963), we shall not give such a list here, but will describe a few of the principle results obtained. The reader is also referred to the review article by Domb (1960) for a very thorough account of the whole theory of cooperative phenomena in crystals prior to the application of the Padé approximant method. The Ising model problem is well suited to the use of the Padé approximant method because quite long perturbation series are available in the literature for most of the interesting thermodynamic quantities and because one has a good notion from the physical aspects of the problem what the analytic structure of the functions should be.

IV C.1. Properties of the Ising Model

Although information is available in the literature (See, Fisher and Sykes [1962ab] and also Danielian [1964]) for the study of the properties of the antiferro magnetic Ising model, almost all the applications have been made to the ferromagnetic Ising model and we shall review only these calculations here.

(a.) Magnetic Susceptibility

We will describe in this subsection how the Padé approximant method

was first applied to the investigation of the critical point of some two and three dimensional Ising model lattices (Baker, 1961).

The simplest assumption that one can make concerning the nature of the singularity of the magnetic susceptibility for a two- or three-dimensional Ising model lattice is that in the neighborhood of the critical point the magnetic susceptibility is proportional to $(1 - T_c/T)^{-1-g}$. The work of Sykes and Domb [Sykes (1961), Domb and Sykes (1961)] indicates that this assumption is very likely so. If it is so, then the logarithmic derivative of the magnetic susceptibility will have a simple pole at the critical point, and its residue will give the nature of the singularity. Since the Padé approximant is the ratio of two polynomials, a simple pole has the possibility of being exactly represented by the Padé approximant and hence one would expect convergence at this type of singularity to be particularly good. For this reason we have computed the $[N,N]$ Padé approximants to the logarithmic derivatives of the magnetic susceptibilities as functions of w where

$$w = \tanh J/kT, \quad J = \text{exchange integral} \quad (\text{IV.25})$$

and we will abbreviate the lattices by fcc, for face-centered cubic lattice; bcc, for body-centered cubic lattice; sc, for simple cubic lattice; t, for triangular lattice; sq for simple quadratic lattice; and h, for honeycomb lattice.

In Table V we have listed, except where otherwise noted, the location of and residue at the closest pole to the origin of the $[N,N]$ Padé approximant. By examination of the results listed in Table V it seems

reasonable that, within an error of the order of a few thousandths, all the three-dimensional lattices have $g = 0.25$ and all the two-dimensional lattices have $g = 0.75$. These results confirm the previous observations of Domb and Sykes (1961).

It should be remarked that the poles in Table V which lie nearer than the one which corresponds to the ferromagnetic critical point always have a zero very close by and represent a severe perturbation of the function value over only a very small range. The occurrence of these perturbations is not clearly understood, as we have discussed in Chap. III.

To test the consistency of these results and obtain more accurate estimates of the location of the critical point we have raised the three-dimensional susceptibilities to the $4/5$ power and the two-dimensional ones to the $4/7$ power. This operation has the effect of converting the ferromagnetic singularity into a simple pole. In Table VI we list the location of the nearest pole and its residue for the $[N,N]$ Padé approximants to $[\chi(w)]^m$.

From looking at the results listed in Table VI we see that the Padé approximants have converged to the known answers for the location of the critical point within a few parts in the sixth place for the sq and h lattices and within a part in the fifth place for the triangular lattice. The rate of convergence for the three-dimensional lattices is such that about the same degree of accuracy is obtainable for them as was obtained for the two-dimensional cases. Furthermore they seem to show an oscillatory pattern of convergence (starting with $N = 2$ for the loose packed lattices)

which enables one to estimate the error. Using the rate of convergence as a guide to interpolate the last two approximants for each lattice, we estimate for the critical point

$$\begin{aligned} \text{fcc } w_c &= 0.101767 & \text{Residue} &= 0.09923 \\ \text{bcc } w_c &= 0.1561789 & \text{Residue} &= 0.152773 \\ \text{sc } w_c &= 0.218156 & \text{Residue} &= 0.22138 \end{aligned} \quad (\text{IV.26})$$

where the last place quoted must be considered rather uncertain. It is to be noted that these estimates are also consistent with those of Table V.

As an additional consistency check on our procedures we may now remove the simple pole in the logarithmic derivative which corresponds to the ferromagnetic critical point by multiplying the logarithmic derivative by $(w - w_c)$. We may then use Padé approximants to evaluate $(w - w_c)d \ln \chi / dw$ at w_c . Actually, for the close-packed lattices this refinement is not strictly necessary as the power series converges reasonably well at w_c . For the loose packed lattices χ goes to zero for a value of w somewhat smaller (algebraically) than $-w_c$, the antiferromagnetic critical point. This zero causes a singularity in $d \ln \chi / dw$ which slows the convergence of its power series so that the use of the Padé approximant method is necessary to obtain a reasonably accurate value at w_c from the terms available. We have listed the values of the $[N,N]$ Padé approximants to $(w - w_c)d \ln \chi(w_c) / dw_c$ in Table VII. They show consistency with the assumed values for g within an error of at most 10^{-3} . The exact values of w_c were used for the two-dimensional cases.

In Figs. 1 and 2 we have plotted the reduced magnetic susceptibility vs w/w_c . For loose-packed lattices this plot is carried over the whole disordered range from -1 to +1. The close packed lattices are carried further in the negative direction as they have no antiferromagnetic state of order and hence no antiferromagnetic transition. We see that the magnetic susceptibility curves are differentiated almost entirely by the dimensionality of the lattice and the coordination number is relatively unimportant.

Essam and Sykes (1963) have subsequently given a number of coefficients for the diamond lattice which permit a similar study of its properties.

(b.) Spontaneous Magnetization.

We will describe in this section how the nature of the singularity in the spontaneous magnetization was determined by the Padé approximant method. We will draw here on the work of Essam and Fisher (1963), who extended and refined the work of Baker (1961). Essam and Fisher also studied the low temperature susceptibility and we will discuss that investigation in sub-section IV C.2.

Essam and Fisher (1963) first derive a large number of coefficients in the low temperature expansion in powers of $u = \exp(-2J/kT) = (1-w)/(1+w)$ where the variables are defined in IV C.1.(a). They obtained the coefficients through the 28th power, fcc, 21st power, bcc, and 18th power, sc. They then consider the sequences of $[N, N+j]$ for various values of j ($-5 \leq j \leq -1$) to

$$(u_c - u) \left(\frac{d}{du} \right) \ln (I_0/Nm), \quad (\text{IV.27})$$

which is basically the method used to construct Table VII of the previous subsection. The rationale for this procedure is that, if, at the critical point, the spontaneous magnetization vanishes like $(u_c - u)^\beta$, [which it does exactly with $\beta = \frac{1}{8}$ for two dimensional lattices, (Yang, 1952)], then the logarithmic derivative has a simple pole and (IV.27) goes to the constant value β at u_c . The value of u_c to be used is taken from the results of the analysis of the susceptibility series. They conclude that

$$\beta = 0.312^{+0.006}_{-0.009} \quad (\text{IV.28})$$

and conjecture that $\beta = 5/16$ exactly. They then compute the $[N, N]$ approximants to $(I_0/Nm)^{-1/\beta}$ using $\beta = 5/16$. If this hypothesis is correct, they should (and do) observe a single pole at $u = u_c$. They then tabulate the difference between the location of the pole and the value of u_c derived from the susceptibility series to check consistency.

(c) Specific Heat.

Baker (1963a) finds that the specific heat, at least for temperatures below the critical temperature has a logarithmic singularity for three dimensional lattices just as it did for two-dimensional ones (Onsager, 1944).

He first checked to see if the log derivative had a simple pole at the critical point as a function (low-temperature) of u (defined in IV C.1. (b)) and determined that it did not seem to have, thus arguing against a

power-law singularity. To pursue the question further, with particular emphasis on the possibility that C_v may be proportional to some power of $\log|1 - T/T_c|$, he expanded the neighborhood of the critical point into the neighborhood of infinity by means of the transformation $u = u_c(1 - e^{-x})$, where u_c is the critical value of u as determined from the susceptibility series. After making this transformation, he computed the $[N,N]$, $[N,N + 1]$, and $[N,N + 2]$ Padé approximants to $(kT/J)^2 C_v[u_c(1 - e^{-x})]$, which are appropriate to $C_v \sim x^0, x^1, x^2$ as $x \rightarrow +\infty$. All three series converged to graphing accuracy out as far as $x \approx 3.3$ although the $[N,N + 1]$ converged the best. In Fig. 3 we have graphed $(kT/J)^2 C_v(u)/[\log_e(1 - u/u_c)]$ based on these results. It is to be noted that the resultant plot is reasonably well behaved near the critical point. On this basis we adopt the tentative hypothesis that C_v is proportional to the first power of $\log(1 - u/u_c)$ in three dimensions, just as in two dimensions.

In order to confirm (or negate) this hypothesis, we have computed $[N,N]$ Padé approximants to $-(kT/J)^2 C_v(u)/[\log_e(1 - u/u_c)]$. The Padé approximants evaluated at u_c are given in Table VIII for the simple cubic lattice (sc), the body-centered cubic lattice (bcc), the face centered cubic lattice (fcc), the two-dimensional simple quadratic lattice (sq), and the function $-x/[\log_e(1 - x)]$. The values of u_c used in the preparation of this table are:

sq,	0.17157287525;	
sc,	0.411940;	
bcc,	0.5326607;	(IV.29)
fcc,	0.664658.	

The first thing one can note from Table VIII is that in the simple quadratic case where the tabulated quantity is known (Onsager, 1944) to tend to $8/\pi$ we see apparent rapid convergence to a value about one per cent smaller than the true value. Several remarks can be made about this. First, the difference between the $[5,5]$ and the $[4,4]$ approximants is abnormally small and is caused by the pole of very small residue between 0 and u_c . This phenomenon is quite common and has been discussed previously (Chap. III). It invariably slows the convergence by causing the very near equality of two successive approximants. Leaving the $[5,5]$ aside, it still appears that the Padé approximant is not converging to the right value. (That it should converge to the wrong value is precluded by theorem 3.) To investigate what sort of rate of convergence we would expect we note that there will be terms which tend to zero inversely as $\log(1 - u/u_c)$. This observation suggests considering the convergence of $-x/[\log(1 - x)]$ at $x = 1$. The convergence of the $[N,N]$ Padé approximants to this function has been shown by Luke (1958) to be exponential everywhere in the complex plane except on the cut $1 \leq x \leq \infty$. At $x = 1$ Luke's arguments break down. We have listed in Table VIII the values of the $[N,N]$ Padé approximants at $x = 1$ for $N = 1 - 12$. We find that $\log(N + 1)$ times this sequence forms a monotonically increasing sequence and $\log(N + 3)$ times this sequence forms a monotonically decreasing sequence. Thus the Padé approximants at $x = 1$ apparently converge inversely as $\log(N)$. This very slow rate of convergence must be expected for all the lattices listed in Table VIII. Thus we see why the value obtained for the simple quadratic

lattice is less accurate than might be expected in view of the closeness of successive approximants. We remark that the same slowness of convergence appears in the Taylor series expansion for

$$-\frac{\log_e(1-x)}{x} = 1 + \frac{1}{2}x + \frac{1}{3}x^2 + \dots \quad (\text{IV.30})$$

The partial sums here for $x = 1$ are only proportional to $\log(N)$.

The slow rate of convergence precludes the precise determination of the coefficient of $\log(1 - u/u_c)$. However, judging by the reasonably accurate results obtained for the simple quadratic lattice and the agreement with the graphical extrapolation for the simple cubic lattice we feel that the conclusion that C_v is proportional to $\log(1 - u/u_c)$ at u_c is confirmed and that the coefficients given by the last entry in Table VIII are more accurate than 10%.

The high temperature expansion was also investigated, but no firm conclusions were possible beyond the remark that the behavior is consistent with the hypothesis that there is a logarithmic singularity on the high-temperature side as well as the low-temperature side, and the coefficient of the logarithm is about 0.3 times that on the low temperature side.

IV C.2. Properties of the Lattice Gas.

The work reported in this subsection is due primarily to Essam and Fisher (1963). Lee and Yang (1952) showed that the Lattice gas problem is mathematically equivalent to the Ising model problem with a magnetic field. In the lattice gas atoms are restricted to occupy only

lattice sites. A downward spinning spin represents an occupied site and an upward spinning one an empty site. Only one atom may occupy a site at a time and the ferromagnetic Ising model corresponds to a lattice gas in which there are hard cores and a short-range (nearest neighbor) attraction. The low-temperature initial susceptibility is isomorphic to the compressibility of the corresponding lattice gas at its point of condensation. The magnetization-vs-T plot is also isomorphic to the vapor-liquid p -T coexistence curve.

The procedure for obtaining the nature of the singularity in the compressibility of the "liquid" phase (low-temperature susceptibility) used by Essam and Fisher (1963) was similar to that used [IV C.1.(b)] was similar to that used [IV C.1.(b)] to obtain the spontaneous magnetization. They first derived a large number of terms in the perturbation series in u [defined in IV C.1. (b)] -- through the 28th power, fcc, 21st power bcc, 18th power, sc, 13th power, t, 9th power, sq, and 23rd power, h. The index γ' of the assumed singularity [proportional to $(u_c - u)^{\gamma'}$] is then estimated from a study of sequences of $[N, N + j]$ Padé approximants to $(u_c - u) (d/du) \ln (KT \chi_0 / 4Nm^2)$. They find, using u_c from the high-temperature susceptibility results, that the index γ' = the index γ which was found from the high temperature susceptibility, i.e.,

$$\gamma' = 1.75 \pm .01, \text{ two dimensional case} \quad (IV.31)$$

$$\gamma' = 1.25^{+0.07}_{-0.02}, \text{ three dimensional case}$$

Essam and Fisher (1963) now point out that on the basis of a model calculation that one might expect that

$$\alpha' + 2\beta + \gamma' = 2 \quad (\text{IV.32})$$

where α' is the index of divergence of the specific heat below T_c , and β is the index associated with the spontaneous magnetization. The best values we have obtained are

two-dimensions

$$\alpha' = 0, \beta = 0.125, \gamma' = 1.75 \quad (\text{IV.33a})$$

three-dimensions

$$\alpha' = 0, \beta = 0.3125, \gamma' = 1.25 \quad (\text{IV.33b})$$

The proposed equation holds in two-dimensions but not in three. In addition to the remarks that $\gamma' = 1.375$ is not excluded by their results for bcc but seems a little high for the other lattices, and that α' or β may be slightly off, we are indebted to G. S. Rushbrooke (1963) for the following clarification. Eq. (IV.32) is really a thermodynamic relation related to, for example, eq.(20.23) of Epstein (1937)

$$C_H - C_M = T \left(\frac{\partial M}{\partial T} \right)_H^2 \left(\frac{\partial H}{\partial M} \right)_T \quad (\text{IV.34})$$

where C is specific heat, M , magnetization, T , temperature, and H magnetic field. As the singularities in C_H and C_M might cancel the best that can be derived by equating powers of $(1 - T/T_c)$ just below the critical point is an inequality where a $\geq$ sign replaces the $=$ sign in (IV.32). This change does not, however, help as we see that some of the numbers must be slightly off. The technique of sections II C. and IV E. might be profitably applied here in an effort to resolve this difficulty.

IV D. Thermodynamic Functions in Lattice Dynamics

We review in this section the work of Domb and Isenberg [Domb and Isenberg (1962), Isenberg (1963)]. The problem in lattice dynamics is to determine the various the various thermodynamic properties, for example specific heat, from models of intermolecular forces. As all of these properties represent averages over the energy spectrum for the lattice under the assumed forces, attention has been concentrated on this spectrum. One easily shows that power series, in reciprocal temperature for the various thermodynamic quantities are directly developable from the moments of the spectral function. In addition a few terms in the low temperature expansion can be obtained. As an example of the power of the method, Isenberg (1963), has calculated the first 34 even moments of the spectrum for a face centered cubic lattice with nearest neighbor interaction and thence, via an [18,15] Padé approximant to $(C_v/3N^3k)^2$ the specific heat. The $[N+3, N]$ approximant to C_v^2 was used as C_v is a function of T^{-2} which goes like T^3 as T goes to zero. Below $T/\Theta = 0.038$ a three term low temperature expansion was used. To check the accuracy the [17, 16] Padé approximant to $(C_v/3N^3k)$ was computed for the linear chain where the exact answer may be computed at all temperatures. The maximum fractional error of the Padé approximant together with a three term low temperature expansion was only about 2×10^{-4} and usually much smaller. Certainly the methods of Domb and Isenberg could be extended to similar models with great precision.

If the results obtained were not already so accurate, there are a couple of additional approaches which might be fruitfully investigated. Almost surely the use of the two-point Padé approximant (see, Chap. V) would be helpful in avoiding the very small interpolated region. Also, as the spectral function, $g(x)$, is non-negative definite, the moment series is a series of Stieltjes

$$F(t) = \int_0^1 \frac{g(x)dx}{1+tx^2} \quad (\text{IV.35})$$

The Padé approximants (see, Sec. II B.) make a well defined approximation to $g(x)$. It is

$$g(x) \approx \sum_{i=1}^N \alpha_i \delta(x-\beta_i) \quad (\text{IV.36})$$

where $\delta(x)$ is the Dirac delta function, $\beta_i = 1/\sqrt{-t_i}$, $\alpha_i = R_i/(-t_i)$; t_i are the locations of the N poles of the $[N, N+j]$ Padé approximant ($t_i < -1$) and R_i are the residues of those poles ($R_i > 0$). One could use this $g(x)$ to calculate the various thermodynamic averages. One might even try to combine this procedure with the two-point method (Chap. V).

IV E. The Heisenberg Model

The applications which have been made to the Heisenberg model [Gammel and Marshall (1963); Baker (1963a), (1964)] have been confined to the magnetic susceptibility, χ , and are similar in general method to those made to the Ising model in IV C.1.(a), so we will not go into such detail here. Baker (1964) has added the further refinement of applying the method of section II C. to obtain upper and lower bounds for the values and the nature of the singularity.

His procedure for extracting relevant information from the power series is as follows. First for some function, $f(z)$ (related to χ) he hypothesizes that in and on a circle with diameter $(-a, b)$, $f(z) + A$ is in the class of functions described just before eq. (II.62). Second, he maps that circle onto the unit circle so that

$$g(z) = f\left(\frac{2z}{(a^{-1} + b^{-1}) - (a^{-1} - b^{-1})z}\right) + A \quad (\text{IV.37})$$

is of a form to which he may apply theorem 9. It then follows that

$$h(w) = g\left[\frac{(1+w)^{\frac{1}{2}} - 1}{(1+w)^{\frac{1}{2}} + 1}\right] (1+w)^{-\frac{1}{2}} \quad (\text{IV.38})$$

is of the form

$$h(w) = \int_{-T_1}^{T_2} \frac{d\Phi(T)}{1+wT} \quad (\text{IV.39})$$

where $\Phi(T)$ is monotonic non-decreasing. In practice, one makes his hypothesis in such a way that $T_1 = 0$. The coefficients for h are given

in terms of those for g by (II.61). He then computes the approximation to $f(x)$ as

$$\{(1+w)^{\frac{1}{2}} [N,M](w) - A\} \quad (\text{IV.40})$$

where

$$w = \frac{2(a^{-1} + b^{-1})x(1 + \frac{1}{2}(a^{-1} - b^{-1})x)}{[1 - b^{-1}x]^2} \quad (\text{IV.41})$$

The properties of upper and lower bounds to the function are preserved, as are those for the slopes. For the slope one must check the values of the slope of the approximants in order to tell whether the slope at an unknown but bounded point is bounded by the slope at the upper or lower bound.

By an investigation of the Stieltjezed logarithmic derivative and χ^{-m} Baker (1964) concludes that the most likely values of the location of the critical points are

Simple cubic	$K_c = 0.588$
body-centered cubic	$K_c = 0.3923$
face-centered cubic	$K_c = 0.2460$

where the last figure is to be considered uncertain, and that $\chi(K_c - K)^{-4/3}$ near $K = K_c$.

As a check on the hypothesis that $\gamma = 4/3$, he computed, using Padé approximants, $(K_c - K)(d \ln \chi / dK)$ at $K = K_c$. This value should be $4/3$ if the hypothesis is correct. We list his results in Table IX. We consider them to be good confirmation, at least, for the body-centered cubic and face-centered cubic lattices. The series for the simple cubic lattice

seems to be more irregular and the results converge more slowly, but we know of no reason why it should be qualitatively different from the body-centered cubic.

IV F. The Many Body Problem.

IV F.1. The Many Fermion System.

The material in this sub-section is taken from the work of Baker et al (1963), (1964b). Their procedure is to compute the perturbation series expansion through fourth order in the potential strength and then use the Padé approximant method to sum the perturbation series and compare the complete series with the Brueckner and ladder approximations to it. They prove, in a manner quite parallel to section IV B., that the ladder approximation to the many fermion energy with a static, purely repulsive potential divided by V , the expansion parameter, is a series of Stieltjes. By a slightly more elaborate argument they show that for a velocity-dependent force which exactly simulates a hard-core for $\lambda = 1$ (expansion parameter) that the energy shift is of the form

$$\Delta E = C_{-1}\lambda - \lambda \int_{-1}^{\infty} \frac{\lambda d\varphi(u)}{1 + u\lambda} \quad (\text{IV.42})$$

where $\varphi(u)$ is monotonic, non-decreasing. This form may be easily converted into a series of Stieltjes. Baker (1963b) has shown that the ladder, Brueckner and complete series diverge only like $(n!)$ and, at least for the repulsive square-well potential, are asymptotic in the full right-half plane. Thus by Theorem 7 we are dealing with the case of a unique sum. The results of their calculations are summarized in Fig. 4. The ratio of the $[1,2]$ and $[2,2]$ to the $[1,1]$ (the $[1,1]$ approximant is the same for the ladder approximation, the Brueckner approximation, and the complete perturbation series) is plotted.

The points represent the values of the approximants. The $[2,2]$ are joined by lines and the different theories are slightly offset for clarity. The flags represent statistical uncertainty in the values of the approximants due to inaccuracies in the calculation of the coefficients.

In fig. 5 we show their results for a hard-core simulating, velocity-dependent force. We mention that the standard K matrix method to calculate the Brueckner approximation fails above $k_F \beta = 1.0$ and one must depend entirely on the Padé method in this range.

Baker et al (1964b) have also computed "mock He^3 " which take a velocity-dependent hard-core simulating force plus a longer range attractive force. They find that the Padé method converges well for this case as well.

Throughout all the calculations reported in this section Baker et al (1963) (1964b) find that the primary difference at low-density between the complete series and Brueckner's approximation to it is his neglect of the ring diagrams.

IV F.2. The Many Boson Problem.

The material in this sub-section is taken from the work of Emery et al (1963). They study a many boson system which interacts via a finite repulsive square-well potential. They expand the many-body energy to fourth order in the potential strength. Then they form the $[2,2]$ and let the strength become infinite. The result is an approximation to the energy of a hard core boson gas. They find that their resultant energy

risks sharply in the neighborhood of the density of closest packing of hard spheres. Their theory has two adjustable parameters, c , the range of the force, and ρ_0 the density of the zero-momentum state. By adjusting these to $c = 2.16 \text{ \AA}$ and $\rho_0 = \frac{1}{2}\rho$ they are able to fit to within a few percent the mean sound velocity, u , and mean of $du/d\rho$ at 1.25° K over the range of pressures -- vapor pressure to 25 atm.

IV G. Regge Poles.

Lovelace and Masson (1962) have shown how the Padé approximant method may be used to calculate the trajectories of Regge Poles. They have illustrated their procedure for a number of Yukawa potentials. They first show how to obtain an asymptotic expansion in k^{-1} (k is the magnitude of the relative momentum) for the function $\alpha(k)$ which tells how the angular momentum α of a generalized bound state varies with the energy. On the basis of an Ansatz, they obtain the asymptotic expansion through a recursion relation in terms of the power series expansion of $rV(r)$, where V is the potential energy. They then use the $[N,N]$ Padé approximants in k^{-1} to sum the series for $\alpha(k)$. They note that for strongly attractive Yukawa potentials that the $\omega_{2N}, \omega_{2N+1}$ of (II.40) seem to tend to limits. As constant ω 's is appropriate to a function of the form $\omega(K = ik)$

$$\frac{1}{2}[(b + K) - [(b + K)^2 - 4a]^{\frac{1}{2}}] \quad (\text{IV.43})$$

they calculate the last denominator by replacing the coefficient of $\Omega_p + 1$ in (II.42) by an appropriately determined factor of the form

$$\frac{1}{2}[(b + K) + [(b + K)^2 - 4a]^{\frac{1}{2}}] \quad (\text{IV.44})$$

with a similar modification for the numerator. This modification improves their convergence for a single Yukawa well, but the regularity of the ω 's does not obtain for the sum of two or more Yukawa wells and hence can not be used for that case. For weak potentials they used the standard $[N,N]$ Padé approximants.

They were able by this method to compute from the [11,11] approximant (or modified approximant) entire Regge trajectories with high accuracy. Their results enable them to obtain 6 figure accuracy for the binding energies.

CHAPTER V

Generalization of the Padé Approximant Method to include Information at Two or More Points

The generalization of the Padé approximant method to two or more points was introduced explicitly by Baker, et al (1964a) although it had been used implicitly before [e.g., pg. 1384 of Baker et al (1963) and possibly also by others]. Their idea was quite simple. Sometimes it happens that one has information about a function at two (or more) points. They suggest that this additional information be taken into account by requiring the Padé approximant to satisfy exactly the conditions at the other points as well as those conditions at the origin or first point.

In the examples discussed by Baker et al (1964a) they imposed the value of the function at infinity on the $[N,N]$ Padé approximants. The required modification in the linear equations which determine the coefficients of P and Q is slight. One replaces the equation, which makes the last power series term agree, by one which makes the Padé approximant equal to a given value at infinity. They ran a few cases to illustrate the nature of the results which might be obtained and to indicate what may and may not reasonably be expected of this method. They considered:

$$a(x) = 1 - \exp(-x)$$

$$u(x) = (1 + x^2)^{\frac{1}{2}} / (1 + x) \quad (V.1)$$

$$c(x) = [(1 + x + x^2)^{\frac{1}{2}} - 1] / x$$

The standard $[N,N]$ Padé approximants do not converge to $a(\infty)$, (see, Chap. III, Table I) because infinity is an essential singularity, but oscillate 0, 2, 0, 2, 0, ... indefinitely. When the value $a(\infty) = 1.0$ is fixed then the $[N,N]$ Padé approximants converge rapidly. The $[7,7]$ gives $a(5)$ to one part in 10^5 and more accurately for smaller real arguments. The most inaccurate region is for large but not infinite arguments where there can occur errors of about one percent. This represents a substantial improvement in the large argument region over the standard Padé approximant as pointed out above. In this example the two point method was used to specify the value at an essential singularity.

For the function $u(x)$, the point at infinity lies on the branch cut (see, Chap. III) connecting the two branch points at $\pm i$, when we use the cut convention defined by the $[N,N]$ Padé approximants. Continuing from positive or negative real values of x , we get $+1$ and -1 , respectively, for $u(\infty)$. When $+1$ was specified as the asymptotic value, all the poles and zeros by which the $[N,N]$ Padé approximants simulate a branch cut lie in the left half plane. $u(1)$ is given by the $[8,8]$ approximant to better than one part in 10^6 and for smaller real positive arguments the accuracy is better. For larger arguments the accuracy decreases to about two percent in the range $10 \leq x \leq 30$ and, of course, then increases to zero error at $x = \infty$.

When they specified $u(\infty) = -1$, then all the poles and zeros simulating the branch cut were in the right-half plane. A consideration of the negative real values of x shows that the accuracy picture is much the same as it was for the positive case of $u(x)$. The pole at $x = -1$ is located to within 3 parts in 10^6 by the $[8,8]$ approximant. The maximum error is about 3 percent here. For the function $u(x)$ the two point method was used to specify which lip of a branch cut for the approximant to converge to.

The function $c(x)$ has branch points at $x = -\frac{1}{2} \pm i\sqrt{3}/2$. The standard Padé approximant would cut (Chap. III) this function so as to connect these branch points and form a single-valued function. The branch cut would cross the real axis at $x = -2$. If one continues through this cut along the negative real axis, one obtains $c(-\infty) = -1$. When one continues along the positive real axis, one obtains $c(+\infty) = +1$, the branch given by the standard $[N,N]$ Padé approximants. When one fixes the value $c(\infty) = 1$, one specifies the value at a regular point and does not much perturb the structure of the approximants, except to greatly accelerate convergence in the neighborhood of that regular point. The $[5,5]$ for this case is good to better than one part in 10^4 for all positive real x .

When, however, they specified $c(\infty) = -1$ in an effort to cause the $[N,N]$ Padé approximants to choose the other branch in the region of large negative x , they were not especially successful. Although through the $[2,2]$ Padé approximant, it appeared as though the desired branch was being taken, as they went to higher approximants the branch cut was made in the

same general way as before and, in fact, the $[5,5]$ Padé approximant does not become negative until $|x| > 170$. We see from this example that specification at a regular point works very well, but that specifying the value alone, does not cause the $[N,N]$ Padé approximants to change their "natural" Riemann sheet. It may be that with a more nearly balanced ratio between coefficients given at zero and infinity one could force a change of what is the "natural" Riemann sheet, but no investigation of this point has been made. Presumably a number of results could be obtained for the many point Padé approximant which are analogous to those of the one-point method, but as yet no detailed investigation has, to my knowledge, been made.

REFERENCES

- Apostol, T. M. (1957) "Mathematical Analysis" Addison-Wesley, Reading, Mass.
- Baker, G. A. Jr. (1961) Phys. Rev. 124, 768.
- Baker, G. A. Jr. (1963a) Phys. Rev. 129, 99.
- Baker, G. A. Jr. (1963b) Phys. Rev. 131, 1869.
- Baker, G. A. Jr. (1964) submitted to Phys. Rev.
- Baker, G. A. Jr. and J. L. Gammel (1961) Jour. Math. Anal. and Appl. 2, 21.
- Baker, G. A. Jr., J. L. Gammel, and J. G. Wills (1961) Jour. Math. Anal. and Appl. 2, 405.
- Baker, G. A. Jr., J. L. Gammel, and B. J. Hill (1963). Phys. Rev. 132, 1373.
- Baker, G. A. Jr., G. S. Rushbrooke, and H. E. Gilbert (1964a). Phys. Rev.
(to be published)
- Baker, G. A. Jr., B. J. Hill, and R. J. McKee, Jr., Phys. Rev. (1964b), (to be published)

- Balachandran, A. P. (1964). Phys. Rev. 134, B 197.
- Carathéodory, C. (1932) "Conformal Representation," Cambridge, London.
- Carleman, T. (1926) "Les fonctions quasi analytiques," p.78, Gauthier, Paris.
- Chisholm, J. S. R. (1963) Jour. Math. Phys. 4, 1506.
- Copson, E. T. (1948), "Theory of Functions of a Complex Variable," Oxford University Press, London.
- Courant, R. (1936) "Differential and Integral Calculus," Vol. II, pg. 201, Interscience, New York.
- Danielian, A. (1964) Phys. Rev. 133, A1344.
- de Balloire, R. de Montessus (1902) Bull. Soc. Math. de France 30, 28.
- Dienes, P. (1957) "The Taylor Series," Dover, New York.
- Domb, C. (1960) Phil. Mag. Supp. 9, 149.
- Domb, C. and C. Isenberg (1962) Proc. Phys. Soc. 79, 659.
- Domb, C. and M. F. Sykes (1961) Jour. Math. Phys 2, 63.
- Dwight, H. B. (1947) "Tables of Integrals" Macmillan, New York.
- Emery, V. J., J. L. Gammel, and F. R. A. Hopgood (1963): Phys. Rev. 132, 10.

- Epstein, P. S. (1937) "Textbook of Thermodynamics," Wiley, New York.
- Essam, J. W. and M. E. Fisher (1963) Jour. Chem. Phys. 38, 802.
- Essam, J. W. and M. F. Sykes (1963). Physica 29, 378.
- Fisher, M. E. (1963). Jour. Math. Phys. 4, 278.
- Fisher, M. E., and M. F. Sykes (1962a) Physica 28, 919.
- Fisher, M. E. and M. F. Sykes (1962b) Physica 28, 939.
- Foldy, L. L. (1964) Phys. Rev. 133, B1221.
- Franklin, P. (1940) "A Treatise on Advanced Calculus," Wiley, New York.
- Frazer, W. R. (1961) Phys. Rev. 123, 2180.
- Frobenius, G. (1881) Jour. für Math. 90, 1.
- Gammel, J. L. and W. Marshall (1963) Proc. Roy. Soc. (Lond.) A275, 257.
- Gronwall, T. H. (1932) Annals of Math. 33, 101.
- Hadamard, J. (1892) Jour. de Math. (4) 8, 101.
- Hardy, G. H. (1956) "Divergent Series" Oxford, London.
- Herglotz, G. (1911) Leipzig Ber. 63, 501.
- Isenberg, C. (1963) Phys. Rev. 132, 2427.

- Klein, L. (Ed.) (1961) "Dispersion Relations and the Abstract Approach to Field Theory," Gordon and Breach, New York.
- Kramers, H. A. (1927) Atti. Congr. Intern. Fisici. Como. 2, 545.
- Lee, T. D., and C. N. Yang (1952) Phys. Rev. 87, 410.
- Lindelöf, E. (1915) Act. Soc. Sci. Fennica (46), 4, 1.
- Lovelace, C. and D. Masson (1962) Nuovo Cimento 26, 472.
- Luke, Y. L. (1957) J. Assoc. Comp. Mach. 4, 24.
- Luke, Y. L. (1958) J. Math. Phys. 37, 110.
- Luke, Y. L. (1960) J. Math. Phys. 38, 279.
- Mandelstam, S. (1958) Phys. Rev. 112, 1344.
- Muir, T. (1960) "A Treatise on the Theory of Determinants" (revised by W. H. Metzler) p. 372, Dover, 1960.
- Onsager, L. (1944) Phys. Rev. 65, 117.
- Padé, H. (1892) Thesis, Ann. de l'Éc. Nor. 9, 1, supplement.
- Perron, O. (1957) "Die Lehr von den Kettenbrücken" Vol. II, B. G. Teubner, Verlagsgesellschaft, Stuttgart.
- Riesz, F. (1911) Ann. de l'Éc. Ncr. 28, 34.
- Rushbrooke, G. S. (1963) J. Chem. Phys. 39, 842.

- Slater, J. C. and N. H. Frank (1947) "Electromagnetism" McGraw-Hill, New York.
- Stieltjes, T. J. (1894) Ann. Fac. Sci. Toulouse 8, 9, 1.
- Sykes, M. F. (1961) Jour. Math. Phys. 2, 52.
- Symanzik, K. (1960) Jour. Math. Phys. 1, 249.
- Van Vleck, E. V. (1901) Annals of Math. 3, 1.
- Wall, H. S. (1948) "Analytic Theory of Continued Fractions". D. Van Nostrand Company, Inc., Princeton, New Jersey.
- Weinberg, S. (1961) Phys. Rev. 124, 2049.
- Wilson, R. (1930) Proc. London Math. Soc. 30, 38.
- Wu, T. Y. and T. Ohmura (1962) "Quantum Theory of Scattering" Prentice-Hall, Englewood Cliffs, New Jersey.
- Yang, C. N. (1952) Phys. Rev. 85, 808.

Table I

N	a(x)	b(x)	w(x)	v(x)	t(x)
1	-1/2	2	0.333333	-1.39793	-
2	1/3	0	2.333333	0.48732	4/3
3	-1/4	2	0.411764	0.63483	5/4
4	1/5	0	2.411764	1.28677	-
5	-1/6	2	0.414141	0.98855	53/42
6	1/7	0	2.414141	1.14742	131/104
7	-1/8	2	0.414211	1.16828	-
8	1/9	0	2.414211	1.22103	1.259978
Limit	0	1	1	1.21042	1.259921

Table II

N	$\operatorname{sech} \sqrt{x}$	$\tanh^2 \sqrt{x}$	$\exp\left(-\frac{x}{1+x}\right)$	$(\tan^{-1} \sqrt{x})^2$	$\sqrt{\frac{1+2x}{1+x}}$	$w(x)$
1	-1/5	3/2	1/3	1.5	7/5	-5/6
2	12/313	150/163	7/19	1.8677	41/29	-0.7877462
3	-127/14615	7371/7244	71/193	2.0331	239/169	-0.7896576
4	17/9326	0.99636	1001/2721	2.1271	1393/985	-0.7895792
5	-3.8254 $\times 10^{-4}$			2.1876	1.414213552	
6				2.2299		
7				2.2611		
8				2.2850		
Limit	0	1	0.367879441	2.4674	1.414213562	-0.7895822

Table III

N	$\int_0^{\infty} \frac{e^{-t}}{1+xt} dt$	$\frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-t^2}}{1+xt^2} dt$	$\int_0^{\infty} e^{-t} \left(\frac{1+2xt}{1+xt} \right)^{1/2} dt$	$4e^{2/x} \times [\text{Erfc}(x^{-1/2})]^2$
1	1/2	2/3	6/5	1.0
2	1/3	2·4/3·5	376/297	1.39130
3	1/4	2·4·6/3·5·7	1.29935	1.61598
4	1/5	$\frac{2 \cdot 4 \cdot 6 \cdot 8}{3 \cdot 5 \cdot 7 \cdot 9}$	1.31982	1.76764
5	1/6	$\frac{2 \cdot 4 \cdot 6 \cdot 8 \cdot 10}{3 \cdot 5 \cdot 7 \cdot 9 \cdot 11}$	1.33377	1.87936
Limit	0	0	1.41421	3.14159

Table IV

N	$\int_0^x e^{-y} dy$	R	$\left[\int_0^x \frac{\sin y}{y} dy \right]^2$	R	$\left[\int_0^x \left(\frac{\sin y}{y^3} - \frac{\cos y}{y^2} \right) dy \right]^2$	R
1	1.0	2.0	4.5	3.0	0.0462963	2.73861
2	0.928203	3.46410	2.01456	4.50884	0.0402306	3.70084
3	1.0	4.64437	2.40654	5.53833	0.0437683	4.40956
4	0.994911	6.04653	1.80747	6.35405	0.0432710	4.99528
5	1.0	7.29348	2.50818	7.35452	0.0432724	4.95526 ^a
6	0.999640	8.67206	2.31515	8.59101	0.0432865	5.66079
7	1.0	9.94356	-	-	-	-
Limit 1.0	-	-	2.46740	-	0.0438649	-

^a Second smallest root.

TABLE V

Results Derived From $d \ln \chi / d w$

N	fcc		bcc		sc	
	location	residue	location	residue	location	residue
1	0.10101010	-1.2243648	0.13953488	-0.9345592	0.19047619	-0.8707483
2	0.10187683	-1.2564708	0.15593068	-1.242589	0.2150662	-1.2048243
3	0.10171078	-1.2451572	0.15618195	-1.2500231	0.21896751	-1.2808656
4			0.15601592	-1.2392262	0.21815114	-1.2505286
5					0.21818264	-1.2518014
					(-0.18387029	0.00068696)

TABLE V (cont'd.)

N	location ^t	residue	location ^{sq}	residue	location ^h	residue
1	0.2666667	-1.7066667	0.28571428	-0.6530612	0.33333333	-0.3333333
2	0.26705389	-1.7124366	0.41118648	-1.6545587	0.50000000	-1.0000000
	(-0.12912285	0.00218936)				
3	0.26671401	-1.7055748	0.40926772	-1.6257290	no positive real pole	
4	0.26392285	-1.7516520	0.41644866	-1.7973526	0.56797947	-1.5308261
	(0.22948115	0.06341672)				
5	0.26795019	-1.7495212	0.41216606	-1.6823402	0.56793836	-1.5302682
			(-0.20919322	-0.00003223)	(-0.07395963	0.8×10^{-9}
6			0.41412464	-1.7458396	0.57301019	-1.6176357
7			0.41421058	-1.7496448	0.57266504	-1.6104612
					(-0.09651947	-0.35×10^{-8})
8					0.57737042	-1.7512127
9					0.57739267	-1.7521167
					(-0.24445536	0.8×10^{-8})
10					0.57736770	-1.7510165
11					0.57737006	-1.7511293
					(-0.49713508	0.738×10^{-3})
∞	0.267949193		0.414213562		0.5773502692	

TABLE VI

Results Derived From $(\chi)^m$

N	fcc		bcc		sc	
	location	residue	location	residue	location	residue
1	0.10204082	-0.09958349	0.16129032	-0.16549323	0.22727272	-0.24793388
2	0.10174095	-0.099113327	0.15627479	-0.15314145	0.21871837	-0.22349263
3	0.10177220	-0.099263130	0.15617854	-0.15277110	0.21804071	-0.22055930
4	0.10176345	-0.099202123	0.15617908	-0.15277365	0.21819754	-0.22169306
5					0.21814435	-0.22121910

TABLE VI (cont'd.)

N	location	t residue	location	sq residue	location	h residue
1	0.26923077	-0.24852071	0.46666667	-0.49777778	0.73684211	-0.93074791
2	0.26837648	-0.24526609	0.41823005	-0.37414691	0.56866762	-0.42032321
3	0.27030309	-0.24664458	0.41404192	-0.35510985	0.56346843	-0.40380449
	(0.21287916	-0.0023842)			(-0.14787515	-0.00000302)
4	0.26784605	-0.24297979	0.41281162	-0.34657517	0.58226219	-0.47764838
			(-0.23886399	-0.00001166)		
5	0.26795882	-0.24383274	0.41421872	-0.35728803	0.57767633	-0.45254064
6	0.26795969	-0.24383884	0.41420612	-0.35719057	0.57719100	-0.44915654
	(0.12757395	-5.65×10 ⁻⁹)	(-0.19136145	-0.9×10 ⁻⁸)		
7			0.41421714	-0.35728743	0.57734619	-0.45052022
8					0.57734607	-0.45051897
					(-0.01404287	-1.15×10 ⁻¹¹)
9					0.57737434	-0.45082173
10					0.57734262	-0.45045917
					(-0.54474770	-0.00015766)
11					0.57734823	-0.45052893
12					0.57735732	-0.45070361
∞	0.267949193		0.414213562		0.5773502692	

TABLE VII

The $[N,N]$ Padé approximants to $(w - w_c) d \ln \chi / dw$ at w_c

N	fcc	bcc	sc	t	sq	h
1	-1.2591654	-1.2211120	-1.2253865	-1.7320495	-1.5672233	-1.3660254
2	-1.2497759	-1.2507893	-1.3044557	-1.7320495 ^{a/}	-1.7058387	0.0
3	-1.2497511	-1.2498902	-1.2501891	-1.7479548	-2.0000000	-1.6096621
4		-1.2493802	-1.2507423	-1.7508444	-1.7279384	-1.7179725
5			-1.2505479	-1.7494566	-1.7515982	-1.8376120
6					-1.7498932	-1.7409649
7					-1.7497973	-1.7542884
8						-1.7503526
9						-1.7501918
10						-1.7501958 ^{b/}
11						-1.7500932

^a Pole at -0.562×10^{-6} with residue of -8.4×10^{-14} .^b Pole at 0.1186061 with residue of -2.2×10^{-9} .

TABLE VIII

Results Derived From $(kT/J)^2 C_v(u) / [\log_e(1 - u/u_c)]$

N	$-x/[\log_e(1-x)]$	sq	sc	bcc	fcc
1	0.4	2.58817	...	...	...
2	0.3	2.51469	2.4012	...	...
3	0.2553	2.51857	12.0669	8.8835	...
4	0.2290	2.51576	12.1906	14.7564	...
5	0.2113	2.51567 ^a	11.6193	20.2049	-299.28
6	0.1983		11.6207	21.6414	33.4795
7	0.1883			20.7962	33.2534
8	0.1803			20.3615	44.6989
9	0.1737				42.2747
10	0.1681				43.0475
11	0.1633				43.0834
12	0.1592				
∞	0.0	2.54648			

^aPole at 0.10405212 with residue of -2.09×10^{-8} .

TABLE IX

 γ from $(K_c - K)(\ln \chi/aK)$

lattice	lower bound	upper bound	a	b	A
S.C.	1.254	1.347	0.5	2.0	4.0
B.C.C.	1.326	1.343	0.5	1.5	1.0
F.C.C.	1.332	1.337	1.0	0.8	2.0

FIGURE CAPTIONS

FIG. 1 Reduced Magnetic susceptibility vs. w/w_c for loose-packed lattices.

The discontinuity in the vertical scale at 1.0 should be noted.

FIG. 2 Reduced magnetic susceptibility vs. w/w_c for the close-packed

lattices. The discontinuity in the vertical and horizontal scales should be noted.

FIG. 3 $C_v/\ln(1 - u/u_c)$ as a function of u/u_c taken from the [7,7] Padé approximant.

FIG. 4 Comparison of the ladder approximation, the Brueckner approximation, and the complete perturbation theory for $V = 2.5$ as represented by the [1,2] and [2,2] Padé approximants. For convenient presentation the results have been normalized by dividing by the [1,1] Padé approximant.

FIG. 5 Sum of complete perturbation series, ladder approximation series, and Brueckner approximation series as given by the [2,2] Padé approximant in units of the first order coefficient. The tic mark on the left border is the exact zero density value.