ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 21, Number 1, Winter 1991

CONVERGENCE THEOREMS FOR ROWS OF HERMITE-PADÉ INTEGRAL APPROXIMANTS

GEORGE A. BAKER, JR. AND P.R. GRAVES-MORRIS

ABSTRACT. The object of this paper is the study of the convergence of integral approximants, which are a special case of Hermite-Padé approximants of Latin type, to functions which are analytic in a disk except for one interior singular point. We give detailed estimates of the rate of convergence of the sequence of approximants of type $[L/M;1]$ for fixed M, as $L \to \infty$, in a model case study. We also give estimates of the rate of convergence of approximants of type $[L/M;1;2]$ for fixed M, as $L \to \infty$, for a model exhibiting a confluent singularity. We prove that integral approximants of these types converge uniformly on compact subsets of the disk which is centered on the origin and has the singular point of the given function on its boundary. We further prove convergence on additional Riemann sheets beyond the principal one in a lune near the singular point.

1. Introduction. Functions which are defined on a multiply-connected Riemann surface can be approximated accurately only by functions having a similar Riemann surface. To this end, Hermite-Padé approximants (of Latin type) have been used successfully to approximate functions having branch cuts; these approximants were introduced by Padé [**23, 24**] and contemporaneously by Hermite [**14**]. Let $f(z)$ be a function which is analytic except for a finite number of branch points of square root type. Shafer studied the special case of quadratic approximants [**25**], which are suitable for $f(z)$. He showed how polynomials $P(z)$, $Q(z)$ and $R(z)$ can be found from a knowledge of $f(z)$ so that

$$P(z)y(z)^2 + Q(z)y(z) + R(z) = 0 \tag{1.1}$$

has a solution $y(z)$ which approximates $f(z)$ near its branch points. Provided that $P(z), Q(z)$ and $R(z)$ have been suitably chosen, the branch points of $y(z)$ will be located close to those of $f(z)$, and the Riemann surfaces of $f(z)$ and $y(z)$ will be similar.

Hermite-Padé approximation is the preferred method of approximation of a function $f(z)$ when $f(z)$ has known analytic properties (ideally, the topology of the Riemann surface of $f(z)$ should be known) and when the Maclaurin expansion of $f(z)$ is given. Using this expansion, one can, for example, find nontrivial polynomials $P(z), Q(z)$ and $R(z)$ for which

$$P(z)f(z)^2 + Q(z)f(z) + R(z) = O(z^{\partial p+\partial Q+\partial R+2}), \tag{1.2}$$

where ∂P denotes the maximum allowed degree of $P(z)$, etc. Clearly, $\partial P + \partial Q + \partial R + 2$ coefficients of the Maclaurin expansion of $f(z)$ are needed in order to derive $P(z), Q(z)$ and $R(z)$ (up to an irrelevant constant common factor). A quadratic, or Shafer, approximant of $f(z)$ then follows as the corresponding solution of (1.1). Further explanation is given by Baker and Graves-Morris [**5**].

In this paper, we are primarily concerned with approximation techniques which allow extrapolation further onto the Riemann surface. The problem of finding Padé approximants for a function which is analytic in the cut z-plane has received a great deal of attention [**1, 5, 22, 26**, and references therein]. In ideal circumstances, a suitably chosen sequence of Padé approximants for $f(z)$ converges to $f(z)$ except that certain poles and zeros of the approximants accumulate on lines which, in this sense, are the natural locations of the cuts of $f(z)$. Commonly, these lines are arcs of circles [**4, 20**], and, as such, they are not components of a Mittag-Leffler star [**13**]. Because each Padé approximant is single-valued, there is no natural method of continuation through the natural cuts. Certain Hermite-Padé approximants, including the quadratic approximants, have the potential to approximate accurately on a multi-sheeted surface and to reveal features which are thought to exist there.

In many physical applications, functions are known (either by value or by their power series expansions) on what is called the physical sheet, because physical quantities are the values (possibly boundary values) of the function on this sheet. For one class of problems of current interest, namely the reconstruction of thermodynamic variables from their power series expansions, e.g., [**27**], we have an incomplete knowledge of the analytic structure of $f(z)$, and we have the values of a finite number (say the first 20 odd) of its Maclaurin coefficients. For example, Baker, Rushbrooke and Gilbert [**8**] discuss the specific heat $C_H(K)$ of the linear ferromagnetic Heisenberg model, where $K = J/kT$ in conventional notation. Near diagonal sequences of Padé approximants of $C_H(K)$, place natural cuts of this function across the negative K axis and prevent extrapolation of the specific heat on the negative axis, which corresponds to the antiferromagnetic model at lower temperatures. Approximants which are singularity-free on the real K axis are to be preferred. A similar difficulty, in which near diagonal sequences of Padé approximants place a natural cut across the positive real time axis occurs in the "bubble" problem [**21**]. Scattering amplitudes are usually constructed as boundary values of a function with cut-plane analyticity, and one of these cuts is the elastic scattering cut. Resonances of the scattering particles are represented by poles on unphysical sheets, and the more significant poles are close to the physical boundary [**11, 19**].

Undoubtedly, much contemporary research on Hermite-Padé approximants is motivated by interest in evaluating the critical exponents of certain thermo-dynamic variables [**12, 15**] and on their impact on the theory of universality.

We are concerned here with *integral approximants* which belong to the class of Hermite-Padé approximants of Latin type. Integral approximants are defined in the following way. Given a function defined by its Maclaurin series as

$$f(z) = \sum_{j=0}^{\infty} f_j z^j, \tag{1.3}$$

we can always find nontrivial polynomials $P(z), Q(z), \ldots, S(z), T(z)$, of degrees $p, q, \ldots, s, t$ at most, such that

$$P(z)\frac{d^m f}{dz^m} + Q(z)\frac{d^{m-1} f}{dz^{m-1}} + \cdots + S(z)f + T(z) = O(z^{n+1}), \tag{1.4}$$

where $n+1 = m+p+q+\cdots+s+t$ [**6**]. Except in degenerate cases, we can impose the normalization $P(0) = 1$. Corresponding to these polynomials, we define the integral approximant

$$y(z) := [t/s\,;\ldots;q\,;p] \tag{1.5}$$

as the solution, or integral, of the differential equation

$$P(z)\frac{d^m y}{dz^m} + Q(z)\frac{d^{m-1} y}{dz^{m-1}} + \cdots + S(z)y + T(z) = 0 \tag{1.6}$$

initialized by (1.3). A review of current progress is given in [**3**].

We concentrate on a special aspect of integral approximants in this paper. Specifically, we will assume that $f(z)$ is analytic in the disk $|z| < \rho$, for some $\rho > 1$, except at the single point $z = 1$. Suppose that $f(z)$ has a singular point of finite order at $z = 1$, and that the continuations of $f(z)$ provide precisely $m + 1$ linearly independent coverings of the disk $|z| < \rho$, one of which is the analytic (background) term. Then, Baker, Oitmaa and Velgakis [**7**] have show that there exists a polynomial $P(z)$ and also functions $Q(z), \ldots, S(z), T(z)$ which are analytic in $|z| < \rho$ such that $f(z)$ is a solution of the differential equation

$$P(z)f^{(m)}(z) + Q(z)f^{(m-1)}(z) + \cdots + S(z)f(z) + T(z) = 0. \tag{1.7}$$

The separation property [**7**] allows the class of functions so defined to be split into two sub-classes. Roughly speaking, $f(z)$ has the separation property if it can be decomposed into the sum of two functions, one of which has no singularities inside $|z| < \rho$, and the other having no singularities in $\rho < |z| < \infty$. If $f(z)$ has the separation property in $|z| < \rho$, then the functions $Q(z), \ldots, S(z)$ are, in fact, polynomials, as is $P(z)$, and $T(z)$ is analytic in $|z| < \rho$. Otherwise, at least one of $Q(z), \ldots, S(z)$ is not a polynomial but merely a function which is analytic in $|z| < \rho$.

Baker, Oitmaa and Velgakis [**7**] have not treated the theory of the case where the separation property fails, although they do give some numerical examples. It is this case with which we shall be concerned here. We will investigate the convergence of sequences where $t \to \infty$ and $p, q, \ldots, s$ are held fixed and finite. In the cases we have treated, we find that this limit reproduces the behavior of $P, Q, \ldots, S$, as expanded about $z = 1$ with explicit results on the error. The approximation $y(z)$ converges to $f(z)$ for $|z| < 1$, and a detailed description of $f(z)$ for z near 1 is desirable. As is well known from the theory of ordinary differential equations, the singularity at $z = 1$ corresponds to a zero (possibly multiple) of $P(z)$ at $z = 1$. We find this zero, together with estimates of its rate of approach to unity, in the sequence of approximants which we study.

In Section 2, we consider Hermite-Padé approximation of functions expressible as

$$f(z) = A(z)(1 - z)^{-\gamma} + B(z),$$

where $A(z)$ and $B(z)$ are analytic in $|z| < \rho$ and $\rho > 1$ is given. Such functions satisfy a first order ordinary differential equation (ODE). We give detailed estimates of the rate of convergence of the polynomial coefficients of y, y', etc., in (1.6) (i.e., in the defining equations of the integral approximants) in the limit as $\partial\{T(z)\} \to \infty$. We find that they approach the limit as various powers of that degree.

In Section 3, we consider the integral approximants to functions expressible as

$$f(z) = A(z)(1-z)^{-\gamma} + B(z)(1-z)^{-\theta} + C(z),$$

where $A(z), B(z)$ and $C(z)$ are analytic in $|z| < \rho$ for some $\rho > 1$. Such functions satisfy a second order ODE. This case is referred to as a confluent singularity because it has two independent singularities at the same point. Again, we give detailed estimates of the rate of convergence of the polynomial coefficients in the equations defining the integral approximants. These rates also turn out as powers of $\partial\{T(z)\}$, but the exponents are not normally integers, in contrast with the integer powers which occur in Section 2.

In the final section, we give two theorems concerning integral approximants to the class of functions discussed in the previous two sections. In the first theorem, we prove uniform convergence of the approximants on compact subsets of $|z| < 1$. For the functions of Section 2, we prove in our second theorem that their integral approximants converge in the portion of all Riemann sheets accessible in $|z| < \rho$ which can be lifted onto the lune $|z| < 1$, $|1-z| < \rho - 1$.

2. Hermite-Padé approximation using first order ODEs. Our aims are the reconstruction of a function $f(z)$ and its properties as accurately as possible from a knowledge of a finite number of its power series coefficients. We make certain other hypotheses about $f(z)$, of which the main one is that $f(z)$ satisfies an ODE of the form

$$(1-z)f'(z) + G(z)f(z) = H(z), \tag{2.1}$$

where $H(z)$ and $G(z) := \sum_{i=0}^{\infty} G_i(1-z)^i$ are analytic in the disk $|z| < \rho$ for some $\rho > 1$. We assume that $\gamma := -G_0$ is not an integer, and then the solution of (2.1) may be expressed [**16, 17**] as

$$f(z) = A(z)(1-z)^{-\gamma} + B(z), \tag{2.2}$$

where $A(z)$ and $B(z)$ are analytic in $|z| < \rho$. Our third assumption about $f(z)$ is that

$$A(1) \neq 0, \tag{2.3}$$

so that $f(z)$ necessarily has the singularity structure explicitly exhibited by (2.2). To introduce our first method of Hermite-Padé approximation, which is suitable for a function $f(z)$ of the form given by (2.2), we suppose that the first $L+M+2$ coefficients of its Maclaurin series (and, thereby, the first $L+M+1$ coefficients of $f'(z)$) are given. Using these coefficients, we form a set of $L+M+1$ homogeneous linear equations which are then solved and values of $\lambda^{(L)}$, $\alpha^{(L)}$, $\{g_i^{(L)}\}_{i=0}^{M-1}$, $\{h_i^{(L)}\}_{i=0}^{L-1}$ are found so that

$$\left[\lambda^{(L)}(1-z)+\alpha^{(L)}\right] f'(z)+g^{(L)}(z)f(z)=h^{(L)}(z)+O(z^{L+M+1}), \tag{2.4}$$

where

$$g^{(L)}(z) := \sum_{i=0}^{M-1} g_i^{(L)}(1-z)^i, \tag{2.5}$$

$$h^{(L)}(z) := \sum_{i=0}^{L-1} h_i^{(L)} z^i \tag{2.6}$$

and the $O(\cdot)$ notation used in (2.4) only implies an accuracy-through-order condition. The solution $y(z)$ of

$$[\lambda^{(L)}(1-z)+\alpha^{(L)}]y'(z)+g^{(L)}(z)y(z)=h^{(L)}(z) \tag{2.7}$$

satisfying $y(0)=f(0)$ is called the Hermite-Padé, integral approximant to $f(z)$, defined by the method stated.

Our main result in this section is

Theorem 2.1. *With the hypotheses and the construction of* (2.1)–(2.7), *we may take* $\lambda^{(L)}=1$ *for* L *large enough. With this normalization,*

$$\alpha^{(L)} = O(L^{-M-1}), \tag{2.8}$$

$$g^{(L)}(z) \to \sum_{i=0}^{M-1} G_i(1-z)^i \quad \text{as } L \to \infty, \tag{2.9}$$

and

$$h^{(L)}(z) \to H(z) - \left[G(z) - \sum_{i=0}^{M-1} G_i(1-z)^i\right] f(z), \quad |z| < 1, \tag{2.10}$$

as $L \to \infty$.

Remarks. Before beginning the proof of this theorem, first we mention its predecessor, next we establish some notation and then a lemma. If $G(z)$ is a polynomial of degree $M-1$ or less, much stronger results giving convergence in the cut-plane and on other sheets, for $|z| < \rho$, were established by Baker [**2**] and then were amplified by Baker, Oitmaa and Velgakis [**7**]. For such a $G(z)$, it happens that $h^{(L)}(z) \to H(z)$ with a geometric rate of convergence in $|z| < \rho$. This situation is a natural analogue of de Montessus' [**18**] theorem.

Notation. Let $[f(z)]_j$ denote the coefficient of z^j in the Maclaurin expansion of $f(z)$. We also define

$$\begin{aligned}
&(\mu)_j := \mu(\mu+1)\cdots(\mu+j-1), \qquad j \ge 1,\\
&(\mu)_j := 1, \quad j = 0,\\
&(\mu)_j := 0, \quad j < 0,
\end{aligned}$$

and the $O(\cdot)$ notation is used in its strong sense [**13**] unless otherwise stated.

Lemma 2.1. *Let* $\{a_i^{(L)}\}_{i=0}^{2M}$ *denote a set of* $2m+1$ *real or complex numbers, each depending on* L. *We assume that* $\sup_L \max_i |a_i^{(L)}| < \infty$, *and that these numbers satisfy the constraint equations*

$$\sum_{j=0}^{2M} a_j^{(L)}(\gamma+1-j)_i = \eta_i^{(L)}, \qquad i = L, L+1, \dots, L+M, \tag{2.11}$$

where, for some constant κ, *not depending on* L *or* i,

$$|\eta_i^{(L)}| < \kappa|(\gamma-2M)_i|, \qquad i = L, L+1, \dots, L+M. \tag{2.12}$$

Then

$$a_j^{(L)} = O(L^{-M+j-1}), \qquad j = 0, 1, \dots, M. \tag{2.13}$$

Proof. Define matrices X, Y elementwise by

$$X_{ij} = (\gamma + 1 - j)_{L+i}, \quad i, j = 0, 1, \dots, M,$$
$$Y_{ij} = (\gamma - M - j)_{L+i}, \quad i = 0, 1, \dots, M, \quad j = 0, 1, \dots, M-1.$$

Define vectors $\mathbf{a}$, $\tilde{\mathbf{a}}$ and $\boldsymbol{\eta}$ by

$$\mathbf{a} := \left(a_0^{(L)}, a_1^{(L)}, \dots, a_M^{(L)}\right)^T, \qquad \tilde{\mathbf{a}} := \left(a_{M+1}^{(L)}, a_{M+2}^{(L)}, \dots, a_{2M}^{(L)}\right)^T,$$
$$\boldsymbol{\eta} := \left(\eta_L^{(L)}, \eta_{L+1}^{(L)}, \dots, \eta_{L+M}^{(L)}\right)^T.$$

The constraint equations (2.11) can now be expressed in the compact form

$$X\mathbf{a} = -Y\tilde{\mathbf{a}} + \boldsymbol{\eta}. \tag{2.14}$$

For $i = 0, 1, \dots, M$, divide row i of (2.14) by $(\gamma+1)_{L-M+i}$. Then subtract rows repeatedly until the coefficient matrix of $\mathbf{a}$ has a triangular structure, so that (2.14) has been reduced to

$$\tilde{X}\mathbf{a} = -\tilde{Y}\tilde{\mathbf{a}} + \tilde{\boldsymbol{\eta}} \tag{2.15}$$

where

$$\tilde{X}_{ij} = (\gamma + L - M + i + 1)_{M-i-j}(\gamma - j + 1)_j(M + 1 - i - j)_i, \tag{2.16}$$

$$\tilde{Y}_{ij} = \frac{(j+1)_i(\gamma - M - j)_{j+M+1}}{(\gamma + L - M - j - 1)_{i+j+1}}. \tag{2.17}$$

An example of this reduction is given as Example 2.1, where the triangular structure of $\tilde{X}$ is explicitly displayed. From (2.12), we find that

$$\left|\tilde{\eta}_i^{(L)}\right| < \kappa|(\gamma - 2M)_{L+M}|/|(\gamma + 1)_L|$$

for all i and for L large enough. Hence,

$$\left|\tilde{\eta}_i^{(L)}\right| = O(L^{-M-1}). \tag{2.18}$$

Using the equations (2.15) in reverse order, and by back-substitution, we obtain

$$a_j^{(L)} = O(L^{-M+j-1}), \quad j = 0, 1, \ldots, M. \quad \square$$

Example 2.1. We exhibit the previous equations for the case of $M = 2$. Equations (2.14) are

$$\begin{bmatrix} (\gamma+1)_L & (\gamma)_L & (\gamma-1)_L \\ (\gamma+1)_{L+1} & (\gamma)_{L+1} & (\gamma-1)_{L+1} \\ (\gamma+1)_{L+2} & (\gamma)_{L+2} & (\gamma-1)_{L+2} \end{bmatrix} \begin{bmatrix} a_0^{(L)} \\ a_1^{(L)} \\ a_2^{(L)} \end{bmatrix}$$
$$= -\begin{bmatrix} (\gamma-2)_L & (\gamma-3)_L \\ (\gamma-2)_{L+1} & (\gamma-3)_{L+1} \\ (\gamma-2)_{L+2} & (\gamma-3)_{L+2} \end{bmatrix} \begin{bmatrix} a_3^{(L)} \\ a_4^{(L)} \end{bmatrix} + \begin{bmatrix} \eta_0^{(L)} \\ \eta_1^{(L)} \\ \eta_2^{(L)} \end{bmatrix}.$$

After division of rows 0, 1, 2 by $(\gamma+1)_{L-2}$, $(\gamma+1)_{L-1}$, $(\gamma+1)_L$, respectively, we obtain

$$\begin{bmatrix} (\gamma+L-1)(\gamma+L) & \gamma(\gamma+L-1) & \gamma(\gamma-1) \\ (\gamma+L)(\gamma+L+1) & \gamma(\gamma+L) & \gamma(\gamma-1) \\ (\gamma+L+1)(\gamma+L+2) & \gamma(\gamma+L+1) & \gamma(\gamma-1) \end{bmatrix} \begin{bmatrix} a_0^{(L)} \\ a_1^{(L)} \\ a_2^{(L)} \end{bmatrix}$$
$$= -\begin{bmatrix} \frac{\gamma(\gamma-1)(\gamma-2)}{\gamma+L-2} & \frac{\gamma(\gamma-1)(\gamma-2)(\gamma-3)}{(\gamma+L-2)(\gamma+L-3)} \\ \frac{\gamma(\gamma-1)(\gamma-2)}{\gamma+L-1} & \frac{\gamma(\gamma-1)(\gamma-2)(\gamma-3)}{(\gamma+L-1)(\gamma+L-2)} \\ \frac{\gamma(\gamma-1)(\gamma-2)}{\gamma+L} & \frac{\gamma(\gamma-1)(\gamma-2)(\gamma-3)}{(\gamma+L)(\gamma+L-1)} \end{bmatrix} \begin{bmatrix} a_3^{(L)} \\ a_4^{(L)} \end{bmatrix} + \begin{bmatrix} \tilde{\eta}_0^{(L)} \\ \tilde{\eta}_1^{(L)} \\ \tilde{\eta}_2^{(L)} \end{bmatrix},$$

where $\tilde{\eta}_i^{(L)} = O(L^{-3})$. In this example, it is easy to see how the rows are subtracted sequentially:

$$\begin{bmatrix} (\gamma+L-1)(\gamma+L) & (\gamma+L-1)\gamma & \gamma(\gamma-1) \\ (\gamma+L)2 & \gamma & 0 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0^{(L)} \\ a_1^{(L)} \\ a_2^{(L)} \end{bmatrix}$$
$$= -\begin{bmatrix} \frac{\gamma(\gamma-1)(\gamma-2)}{\gamma+L-2} & \frac{\gamma(\gamma-1)(\gamma-2)(\gamma-3)}{(\gamma+L-2)(\gamma+L-3)} \\ \frac{\gamma(\gamma-1)(\gamma-2)}{(\gamma+L-1)(\gamma+L-2)} & \frac{\gamma(\gamma-1)(\gamma-2)(\gamma-3)2}{(\gamma+L-1)(\gamma+L-2)(\gamma+L-3)} \\ \frac{\gamma(\gamma-1)(\gamma-2)2}{(\gamma+L)(\gamma+L-1)(\gamma+L-2)} & \frac{\gamma(\gamma-1)(\gamma-2)(\gamma-3)2\cdot 3}{(\gamma+L)\cdots(\gamma+L-3)} \end{bmatrix} \begin{bmatrix} a_3^{(L)} \\ a_4^{(L)} \end{bmatrix} + \hat{\boldsymbol{\eta}}^{(L)},$$

with $\hat{\boldsymbol{\eta}}^{(L)} = O(L^{-3})$. From the last row of these equations, we find that $a_0^{(L)} = O(L^{-3})$. By backsubstitution, it follows that $a_1^{(L)} = O(L^{-2})$ and $a_2^{(L)} = O(L^{-1})$.

Having established Lemma 2.1, we turn our attention back to finding $A(z)$ in (2.2). By direct substitution of (2.2) into (2.1), we find that

$$[\gamma A + (1-z)A' + GA](1-z)^{-\gamma}$$

must be analytic in $|z| < \rho$. Because γ is not an integer, a single-valuedness argument shows that

$$\gamma A(z) + (1-z)A'(z) + GA(z) = 0, \qquad |z| < \rho. \tag{2.19}$$

The solution of (2.19) is

$$A(z) = A(1)\exp\int_1^z \frac{G(t)+\gamma}{t-1}\,dt, \qquad |z| < \rho; \tag{2.20}$$

because $\gamma := -G(1)$, we see that $A(z)$ is analytic in $|z| < \rho$.

Proof of Theorem 2.1. We use the variable $\zeta := 1 - z$ and retain the prime $'$ to denote differentiation with respect to z. We use (2.2) to define coefficients c_i and d_i via

$$\sum_{i=0}^{\infty} c_i z^i := f(z) = A\zeta^{-\gamma} + B, \tag{2.20}$$

$$\sum_{i=0}^{\infty} d_i z^i := f'(z) = A'\zeta^{-\gamma} + A\gamma\zeta^{-\gamma-1} + B'. \tag{2.21}$$

Following (2.4), we derive values of the $M+2$ unknowns $\lambda^{(L)}$, $\alpha^{(L)}$, $\{g_i^{(L)}\}_{i=0}^{M-1}$ from the set of $M+1$ homogeneous linear equations

(2.22)
$$\left[\left\{\lambda^{(L)}(1-z) + \alpha^{(L)}\right\}\left[\sum_{j=0}^{\infty} d_j z^j\right] + g^{(L)}(z)\left[\sum_{j=0}^{\infty} c_j z^j\right]\right]_i = 0,$$
$$i = L, L+1, \dots, L+M.$$

Because (2.22) has a nontrivial solution for $\{\lambda^{(L)}, \alpha^{(L)}, \{g_i^{(L)}\}\}$ for each L, we can arrange that $\lambda^{(L)} \geq 0$ and that

(2.23)
$$\max\left[|\lambda^{(L)}|, |\alpha^{(L)}|, |g_0^{(L)}|, \cdots, |g_{M-1}^{(L)}|\right] = \max\{1, |G_0|, \dots, |G_{m-1}|\}.$$

By substituting (2.20) and (2.21) into (2.22), we obtain

$$\begin{aligned}&\left[\left\{\lambda^{(L)}(1-z) + \alpha^{(L)}\right\}\left\{A'\zeta^{-\gamma} + \gamma A\zeta^{-\gamma-1}\right\} + g^{(L)}(z)A\zeta^{-\gamma}\right]_i \\ &\quad = -\left[\left\{\lambda^{(L)}(1-z) + \alpha^{(L)}\right\}B' + g^{(L)}(z)B\right]_i, \\ &\qquad\qquad i = L, L+1, \dots, L+M,\end{aligned}$$

and, hence,

(2.24)
$$\left[\left\{\lambda^{(L)}\zeta + \alpha^{(L)}\right\}\left\{A'\zeta^{-\gamma} + \gamma A\zeta^{-\gamma-1}\right\} + g^{(L)}(z)A\zeta^{-\gamma}\right]_i = O\left[\hat{\rho}^{-L}\right], \quad i = L, L+1, \dots, L+M$$

for any $\hat{\rho}$ such that $1 < \hat{\rho} < \rho$. By substituting for A' from (2.19) into (2.24), we obtain

$$\text{(2.25)}\quad -\alpha^{(L)}\left[AG\zeta^{-\gamma-1}\right]_i + \left[\left[g^{(L)} - \lambda^{(L)}G\right]A\zeta^{-\gamma}\right]_i = O\left[\hat{\rho}^{-L}\right], \quad i = L, L+1, \dots, L+M.$$

Now we expand $A(z) = \sum_{i=0}^{\infty} A_i\zeta^i$, so that the coefficients of $\zeta^{-\gamma-1}$ and $\zeta^{-\gamma}$ in (2.25) can be analyzed. We define

$$\text{(2.26)}\qquad a_0^{(L)} = \alpha^{(L)}A_0\gamma,$$

$$\text{(2.27)}\quad a_i^{(L)} = -\alpha^{(L)}\sum_{j=0}^{i} A_jG_{i-j} + \sum_{j=0}^{\text{Min}} A_{i-j-1}\left[g_j^{(L)} - \lambda^{(L)}G_j\right], \quad i = 1, 2, \dots, 2M,$$

where Min $:= \min(M-1, i-1)$. Then (2.25) takes the form

(2.28)
$$\begin{aligned}&\left[a_0^{(L)}(1-z)^{-\gamma-1} + a_1^{(L)}(1-z)^{-\gamma} + \cdots + a_{2M}^{(L)}(1-z)^{-\gamma+2M-1}\right]_i \\ &\quad = \left[E_L(z)(1-z)^{-\gamma+2M}\right]_i + O\left[\hat{\rho}^{-L}\right], \\ &\qquad i = L, L+1, \dots, L+M,\end{aligned}$$

where $E_L(z)$ is analytic in $|z| < \rho$. In fact, the functions $\{E_L(z)\}_{L+M}^{\infty}$ are uniformly bounded in $|z| < \hat{\rho}$ for the following reasons. From (2.26), (2.27), we find that $a_0^{(L)}, a_1^{(L)}, \dots, a_{2M}^{(L)}$ are linear combinations of $\lambda^{(L)}, a^{(L)}, \{g_i^{(L)}\}_{i=0}^{M-1}$, and so they are also uniformly bounded (independently of L). We give Example 2.2 later to clarify this relationship. Because each $E_L(z)$ is a finite linear combination of functions analytic in $|z| < \hat{\rho}$, and, because the coefficients involved are bounded, $\{E_L(z)\}$ are analytic and uniformly bounded in $|z| < \hat{\rho}$. Therefore, we can use Darboux's theorem [**9, 10**] to show that a constant κ_1 exists, which is independent of L and j, for which

$$\left|\left[E_L(z)(1-z)^{-\gamma-2M}\right]_i\right| < \kappa_1 |(\gamma - 2M)_i|/i!, \qquad i = L, L+1, \dots, L+M. \tag{2.29}$$

From (2.28) and (2.29), it follows that

$$\sum_{j=0}^{2M} a_j^{(L)} (\gamma + 1 - j)_i = \eta_i^{(L)}, \qquad i = L, L+1, \dots, L+M$$

where

$$|\eta_i^{(L)}| < \kappa |(\gamma - 2M)_i|, \qquad i = L, L+1, \dots, L+M,$$

for L large enough and some constant κ. Therefore, $\{a_j^{(L)}\}$ satisfy the conditions of Lemma 2.1, and so

$$a_i^{(L)} = O(L^{-M-i-1}), \qquad i = 0, \dots, M. \tag{2.30}$$

From (2.26), we obtain

$$\alpha^{(L)} = O(L^{-M-1}). \tag{2.31}$$

From (2.27) and (2.30), we obtain

$$g_i^{(L)} - \lambda^{(L)} G_i = O(L^{-M+i}), \qquad i = 0, 1, \dots, M-1. \tag{2.32}$$

Equations (2.31) and (2.32) constitute the central estimates of this section. Since $\{\lambda^{(L)}\}$ is bounded by (2.23), the Bolzano-Weierstrass theorem implies that a finite limit Λ and its corresponding subsequence $\mathcal{L}$ exist for which

$$\lim_{\substack{L\to\infty \\ L\in\mathcal{L}}} \lambda^{(L)} = \Lambda. \tag{2.33}$$

Suppose that $|\Lambda| \neq 1$. From (2.32), we have

$$\lim_{\substack{L\to\infty \\ L\in\mathcal{L}}} g_i(L) = \Lambda G_i. \tag{2.34}$$

From (2.31), (2.33) and (2.34), we obtain

$$\begin{aligned} \lim_{\substack{L\to\infty \\ L\in\mathcal{L}}} \max &\left\{|\lambda^{(L)}|, |\alpha^{(L)}|, |g_0^{(L)}|, \cdots, |g_{M-1}^{(L)}|\right\} \\ &= |\Lambda| \max\left\{1, |G_0|, \ldots, |G_{M-1}|\right\}. \end{aligned} \tag{2.35}$$

This contradicts (2.33) and, therefore, $|\Lambda| = 1$. Because each $\lambda^{(L)} \geq 0$, it must be that $\Lambda = 1$. Using the usual argument about deletion of subsequences, it follows from (2.33) that

$$\lim_{L\to\infty} \lambda^{(L)} = 1.$$

From (2.31) and (2.32), we obtain

$$\alpha^{(L)} = O(L^{-M-1}) \tag{2.36}$$

$$g_i^{(L)} = G_i + O(L^{-M+i}), \qquad i = 0, 1, \ldots, M-1. \tag{2.37}$$

and

$$g^{(L)}(z) \to \sum_{i=0}^{M-1} g_i (1-z)^i \quad \text{as } L \to \infty.$$

From (2.4), we find

$$h^{(L)}(z) = \sum_{j=0}^{L} z^j \left\{ h_j - \left[\left[G(z) - g^{(L)}(z) \right] \sum_{i=0}^{L} c_i z^i \right]_j \right\},$$

and, hence,

$$h^{(L)}(z) \to H(z) - \left\{ G(z) - \sum_{i=0}^{M-1} G_i (1-z)^i \right\} f(z) \quad \text{in } |z| < 1. \ \square$$

We conclude this section with

Example 2.2. For the case of $M = 2$, equations (2.26), (2.27) take the matrix form

$$\begin{bmatrix} a_0^{(L)} \\ a_1^{(L)} \\ a_2^{(L)} \\ a_3^{(L)} \\ a_4^{(L)} \end{bmatrix} = \begin{bmatrix} A_0 & & \\ A_1 & A_0 & \\ A_2 & A_1 & A_0 \\ A_3 & A_2 & A_1 \\ A_4 & A_3 & A_2 \end{bmatrix} \begin{bmatrix} \gamma\alpha^{(L)} \\ g_0^{(L)} - \lambda^{(L)} G_0 - \alpha^{(L)} G_1 \\ g_1^{(L)} - \lambda^{(L)} G_1 - \alpha^{(L)} G_2 \end{bmatrix},$$

and we see that $a_0^{(L)}, a_1^{(L)}, \dots, a_4^{(L)}$ are linear combinations of $\lambda^{(L)}$, $\alpha^{(L)}$, $g_0^{(L)}$ and $g_1^{(L)}$.

3. Hermite-Padé approximation using ODEs. In this section, we suppose that the function $f(z)$ (which we wish to reconstruct from its power series) satisfies an ODE of the form

$$(1-z)^2 f''(z) + (1-z) G_0 f'(z) + K(z) f(z) = H(z), \tag{3.1}$$

where G_0 is a constant and $K(z)$, $H(z)$ are analytic in the disk $|z| < \rho$ for some $\rho > 1$. We assume that the quadratic equation

$$\nu^2 + (G_0 + 1)_\nu + K(1) = 0 \tag{3.2}$$

has two roots γ, θ, neither of which is an integer, and that $\operatorname{Re}\gamma > \operatorname{Re}\theta > \operatorname{Re}\gamma - 1$. We discover that the Hermite-Padé approximants specified here reproduce the main features of $f(z)$. In particular, the approximants have a singular point near $z = 1$ (at sufficiently high order) and we find estimates of the exponents γ, θ from these approximants which converge to γ, θ, respectively. The rate of convergence is slower than the rate of convergence of the estimates of γ found in Section 2. The solution of (3.1) may be expressed as

$$f(z) = A(z)(1-z)^{-\gamma} + B(z)(1-z)^{-\theta} + C(z), \tag{3.3}$$

where $A(z), B(z)$ and $C(z)$ are analytic in $|z| < \rho$. Thus, the expansions $A(z) = \sum_{i=0}^{\infty} A_i \zeta^i$, $B(z) = \sum_{i=0}^{\infty} B_i \zeta^i$ in $\zeta := 1 - z$ converge in $|\zeta| < \rho - 1$. We also assume that $A_0 \neq 0 \neq B_0$. For reasons similar to those given in Section 2, we find that $A(z)$ and $B(z)$ satisfy

(3.4)
$$(1-z)^2 A'' + 2\gamma(1-z)A' + \gamma(\gamma+1)A + (1-z)G_0 A' + G_0\gamma A + KA = 0,$$

(3.5)
$$(1-z)^2 B'' + 2\theta(1-z)B' + \theta(\theta+1)B + (1-z)G_0 B' + G_0\theta B + KB = 0,$$

respectively. To obtain Hermite-Padé approximants for the solution of (3.1), we first need to find the values of $\lambda^{(L)}, \alpha^{(L)}, \beta^{(L)}, \sigma^{(L)}, \tau^{(L)}$, $\{k_i^{(L)}\}_{i=0}^{M-1}$ and $\{h_i^{(L)}\}_{i=0}^{L-1}$ for which

(3.6)
$$\{\lambda^{(L)}(1-z)^2 + \alpha^{(L)}(1-z) + \beta^{(L)}\}f''(z) + \{\sigma^{(L)} + \tau^{(L)}(1-z)\}f'(z) + k^{(L)}(z)f(z) = h^{(L)}(z) + O(z^{L+M+4}),$$

where $k^{(L)}(z) = \sum_{i=0}^{M-1} k_i^{(L)} \zeta^i$ and $h^{(L)}(z) = \sum_{i=0}^{L-1} h_i^{(L)} z^i$. Obviously, $L + M + 4$ terms of each of the power series $\sum_{i=0}^{\infty} c_i z^i = f(z)$, $\sum_{i=0}^{\infty} d_i z^i = f'(z)$ and $\sum_{i=0}^{\infty} e_i z^i = f''(z)$ are required to set up the equations. The values of $\lambda^{(L)}, \alpha^{(L)}, \beta^{(L)}, \sigma^{(L)}, \tau^{(L)}$ and $\{k_i^{(L)}\}_{i=0}^{M-1}$ are found by solving the set of homogeneous equations

(3.7)
$$\left[\{\lambda^{(L)}(1-z)^2 + \alpha^{(L)}(1-z) + \beta^{(L)}\} \sum_{j=0}^{\infty} e_j z^j + \{\sigma^{(L)} + \tau^{(L)}(1-z)\} \sum_{j=0}^{\infty} d_j z^j + k^{(L)}(z) \sum_{j=0}^{\infty} c_j z^j \right]_i = 0,$$
$$i = L, L+1, \ldots, L+M+3.$$

Using much the same analysis as that of (2.23), (2.33)–(2.35), we find that we may introduce the normalization $\lambda^{(L)} = 1$ for L large enough. At this rather premature stage, we introduce the normalization $\lambda^{(L)} = 1$ for $L \geq L_0$, so as to avoid repetition and further complication of the analysis. The values of the other parameters now follow from (3.7) (and (3.6) for $\{h_j^{(L)}\}_{j=0}^{L-1}$). Using these values, we obtain the Hermite-Padé integral approximants for $f(z)$ as the solutions $y(z)$ of

(3.8)
$$\begin{aligned}\{(1-z)^2 + \alpha^{(L)}(1-z) + \beta^{(L)}\}y''(z) + \{\sigma^{(L)} + \tau^{(L)}(1-z)\}y'(z)\\ + k^{(L)}(z)y(z) = h^{(L)}(z)\end{aligned}$$

for $L = L_0, L_0+1, L_0+2, \dots$. The initial conditions associated with the ODE (3.8) are $y(0) = c_0$, $y'(0) = c_1$ for each L. Our general method of Hermite-Padé, integral approximation is exemplified by the case of $M = 1$.

Example 3.1. The case $M = 1$. We substitute the known form of the equation (3.3) into (3.6) and obtain

(3.9)
$$\begin{aligned}\Big[&\{\zeta^2 + \alpha^{(L)}\zeta + \beta^{(L)}\}\{\gamma(\gamma+1)A\zeta^{-\gamma-2} + 2\gamma A'\zeta^{-\gamma-1} + A''\zeta^{-\gamma}\\ &+ \theta(\theta+1)B\zeta^{-\theta-2} + 2\theta B'\zeta^{-\theta-1} + B''\zeta^{-\theta} + C''\}\\ &+ \{\sigma^{(L)} + \tau^{(L)}\zeta\}\{\gamma A\zeta^{-\gamma-1} + A'\zeta^{-\gamma} + \theta B\zeta^{-\theta-1} + B'\zeta^{-\theta} + C'\}\\ &+ k_0^{(L)}\{A\zeta^{-\gamma} + B\zeta^{-\theta} + C\}\Big]_i = 0, \quad i = L, L+1, \dots, L+4.\end{aligned}$$

Using (3.4), (3.5) to eliminate A'', B'' from (3.9), we obtain

(3.10)
$$\begin{aligned}\Big[&(\zeta^2 + \alpha^{(L)}\zeta + \beta^{(L)})\{C'' - (\zeta G_0 A' + G_0\gamma A + KA)\zeta^{-\gamma-2}\\ &\qquad\qquad - (\zeta G_0 B' + G_0\theta B + KB)\zeta^{-\theta-2}\}\\ &+ \{\sigma^{(L)} + \tau^{(L)}\zeta\}\{\gamma A\zeta^{-\gamma-1} + A'\zeta^{-\gamma} + \theta B\zeta^{-\theta-1} + B'\zeta^{-\theta} + C'\}\\ &+ k_0^{(L)}\{A\zeta^{-\gamma} + B\zeta^{-\theta} + C\}\Big]_i = 0, \quad i = L, L+1, \dots, L+4.\end{aligned}$$

Following the method of Section 2, we pick out the coefficients of $\zeta^{-\gamma-2}, \zeta^{-\theta-2}, \zeta^{-\gamma-1}, \ldots$ in (3.10) and define

(3.11a) $$a_0^{(L)} = -\beta^{(L)}\{G_0\gamma A_0 + K_0A_0\},$$

(3.11b) $$b_0^{(L)} = -\beta^{(L)}\{G_0\theta B_0 + K_0B_0\},$$

(3.11)
$$\begin{aligned}
a_1^{(L)} &= \sigma^{(L)}\gamma A_0 - \alpha^{(L)}\{G_0\gamma A_0 + K_0A_0\} + \beta^{(L)}t_1,\\
b_1^{(L)} &= \sigma^{(L)}\theta A_0 - \alpha^{(L)}\{G_0\theta B_0 + K_0B_0\} + \beta^{(L)}t_2,\\
a_2^{(L)} &= \tau^{(L)}\gamma A_0 + k_0^{(L)}A_0 - G_0\gamma A_0 - K_0A_0 + \beta^{(L)}t_3 + \alpha^{(L)}t_4 + \sigma^{(L)}t_5,\\
b_2^{(L)} &= \tau^{(L)}\theta B_0 + k_0^{(L)}B_0 - G_0\theta B_0 - K_0B_0 + \beta^{(L)}t_6 + \alpha^{(L)}t_7 + \sigma^{(L)}t_8,
\end{aligned}$$

where $t_1, t_2, \ldots$ are constants not depending on L. Notice that $a_0^{(L)} \propto b_0^{(L)} \propto \beta^{(L)}$ for all L. In this sense, (3.11b) is linearly dependent on its predecessor (3.11a), whereas the other equations of (3.11) are linearly independent. For this reason, (3.11b) is treated differently from the other equations.

In our analysis, we need to use the order notation in a slightly modified sense, and we define

(3.12) $$x^{(L)} = O(\theta)_L$$

to mean that $|x^{(L)}/(\theta)_L|$ is bounded as $L \to \infty$. From (3.10), (3.11), we obtain

(3.13)
$$\begin{bmatrix} (\gamma+2)_L & (\gamma+1)_L & (\theta+1)_L & (\gamma)_L & (\theta)_L \\ (\gamma+2)_{L+1} & (\gamma+1)_{L+1} & (\theta+1)_{L+1} & (\gamma)_{L+1} & (\theta)_{L+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ (\gamma+2)_{L+4} & (\gamma+1)_{L+4} & (\theta+1)_{L+4} & (\gamma)_{L+4} & (\theta)_{L+4} \end{bmatrix} \begin{bmatrix} a_0^{(L)} \\ a_1^{(L)} \\ b_1^{(L)} \\ a_2^{(L)} \\ b_2^{(L)} \end{bmatrix} =$$

$$- \begin{bmatrix} (\theta+2)_L & (\gamma-1)_L & (\theta-1)_L & \cdots & (\theta-4)_L \\ (\theta+2)_{L+1} & (\gamma-1)_{L+1} & (\theta-1)_{L+1} & \cdots & (\theta-4)_{L+1} \\ \vdots & \vdots & \vdots & & \vdots \\ (\theta+2)_{L+4} & (\gamma-1)_{L+4} & (\theta-1)_{L+4} & \cdots & (\theta-4)_{L+4} \end{bmatrix} \begin{bmatrix} b_0^{(L)} \\ a_3^{(L)} \\ b_3^{(L)} \\ \vdots \\ b_6^{(L)} \end{bmatrix}$$

$$+ O(\gamma)_{L-1}$$

These equations (3.13) are simplified by introducing the definitions

$$\tilde{a}_j^{(L)} = a_j^{(L)}(\gamma+2-j)_L, \quad \tilde{b}_j^{(L)} = b_j^{(L)}(\theta+2-j)_L, \tag{3.14}$$

$j = 0, 1, \ldots, 6$, which induce a scaling of the columns. We also find it convenient to introduce the notation

$$(a)_{\bar{j}} := a(a-1)(a-2)\cdots(a-j+1). \tag{3.15}$$

We perform elementary row operations on the matrix coefficient on the left-hand side of (3.12) and obtain
(3.16)

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ \gamma^{-\theta+2} & \gamma^{-\theta+1} & 1 & \gamma^{-\theta} & 0 \\ (\gamma^{-\theta+2})\cdot 2 & (\gamma^{-\theta+1})\cdot 1 & 1\cdot(\theta-\gamma+1) & 0 & 0 \\ (\gamma^{-\theta+2})_{\bar{2}}\cdot 2 & (\gamma^{-\theta+1})_{\bar{2}}\cdot 1 & 0 & 0 & 0 \\ (\gamma^{-\theta+2})_{\bar{2}}\cdot 2! & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{a}_0^{(L)} \\ \tilde{a}_1^{(L)} \\ \tilde{b}_1^{(L)} \\ \tilde{a}_2^{(L)} \\ \tilde{b}_2^{(L)} \end{bmatrix}$$

$$= -\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 2 & \gamma-\theta-1 & -1 & \cdots & -4 \\ 2(\theta-\gamma+2) & (\gamma-\theta-1)(-1) & (-1)(\theta-\gamma-1) & \cdots & (-4)(\theta-\gamma-4) \\ 2(\theta-\gamma+2)_{\bar{2}} & (\gamma-\theta-1)_{\bar{2}}(-1) & (-1)_{\bar{2}}(\theta-\gamma-1) & \cdots & (-4)_{\bar{2}}(\theta-\gamma-4) \\ 2!(\theta-\gamma+2)_{\bar{2}} & (\gamma-\theta-1)_{\bar{2}}(-1)_{\bar{2}} & (-1)_{\bar{2}}(\theta-\gamma-1)_{\bar{2}} & \cdots & (-4)_{\bar{2}}(\theta-\gamma-4)_{\bar{2}} \end{bmatrix} \begin{bmatrix} \tilde{b}_0^{(L)} \\ \tilde{a}_3^{(L)} \\ \tilde{b}_3^{(L)} \\ \vdots \\ b_6^{(L)} \end{bmatrix}$$

$$+ O(\gamma)_{L-1}.$$

From (3.11), (3.12), (3.14) and (3.16), we see that

$$|\tilde{a}_0^{(L)}| \le \chi \max\{|\tilde{b}_0^{(L)}|, |\tilde{a}_3^{(L)}|, |\tilde{b}_3^{(L)}|, \cdots, |\tilde{b}_6^{(L)}|\}, \quad L \ge L_0 \tag{3.17}$$

for some constant χ. If $|\tilde{b}_0^{(L)}|$ were the largest in the list in (3.17), then

$$|a_0^{(L)}| \le |b_0^{(L)}| L^{-\gamma+\theta-\varepsilon}, \quad L \ge L_1 \tag{3.18}$$

for some $L_1 \ge L_0$ and some $\varepsilon > 0$, which contradicts (3.11a,b). Recalling the previous analysis, (2.33)–(2.35), of the normalization, we obtain the estimates

$$\begin{aligned} a_0^{(L)} &= O(L^{-3}), & a_1^{(L)} &= O(L^{-2}), \\ b_1^{(L)} &= O(L^{\gamma-\theta-2}), & a_2^{(L)} &= O(L^{-1}), \\ b_2^{(L)} &= O(L^{\gamma-\theta-1}), \end{aligned}$$

and, from (3.11),

$$\begin{aligned} &\beta^{(L)} = O(L^{-3}), \quad \sigma^{(L)}, \alpha^{(L)} = O(L^{\gamma-\theta-2}), \\ &\tau^{(L)} = G_0 + O(L^{\gamma-\theta-1}), \\ &k_0^{(L)} = K_0 + O(L^{\gamma-\theta-1}). \ \square \end{aligned} \tag{3.19}$$

These estimates suffice to establish convergence of the exponents of the Hermite-Padé, integral approximants. For the case of $M \geq 2$, we use the method of Example 3.1 and find that

$$\begin{aligned} &\beta^{(L)} = O(L^{\theta-\gamma-M-1}), \qquad \sigma^{(L)}, \alpha^{(L)} = O(L^{-M}), \\ &\tau^{(L)} = G_0 + O(L^{1-M}), \\ &k_0^{(L)} = K_0 + O(L^{1-M}), \\ &k_j^{(L)} = K_j + O(L^{1-M+\theta-\gamma+j}), \quad j = 1, 2, \ldots, M-1. \end{aligned} \tag{3.20}$$

The estimates of (3.20) can be used to give estimates of the location of the leading singular point and of its associated singular indices. We first remark that each integral approximant which results from solving (3.6) has, in general, two distinct algebraic singular points near $z = 1$, rather than one confluent singularity. These singularities are located at

$$\begin{aligned} Z_\pm^{(L)} &= 1 + \frac{1}{2}\alpha^{(L)} \pm \left[\left[\frac{1}{2}\alpha^{(L)}\right]^2 - \beta^{(L)}\right]^{\frac{1}{2}} \\ &= 1 + O\left[L^{-\frac{1}{2}(M+1+\gamma-\theta)}\right], \quad M \geq 2, \\ &= 1 + O(L^{-3/2}), \quad M = 1. \end{aligned} \tag{3.21}$$

To show leading order, the exponents of $y^{(L)}(z)$ at $z_\pm^{(L)}$ are

$$\psi_\pm = -\frac{1}{2}G_0 - 1 = \frac{1}{2}(\theta + \gamma - 1), \tag{3.22}$$

whereas the exponents of the originating function are γ and θ. In order to estimate the latter exponents, we first note that

$$\begin{aligned} z_s^{(L)} := \frac{1}{2}\left[z_+^{(L)} + z_-^{(L)}\right] &= 1 + O(L^{-M}), \quad M \geq 2, \\ &= 1 + O(L^{\gamma-\theta-2}), \quad M = 1, \end{aligned} \tag{3.23}$$

and (3.23) is, in general, a more accurate estimate than (3.21) of the singular point. To estimate the values of γ, θ, one may insert the values of $\tau^{(L)}$ and $k_0^{(L)}$ for G_0 and K_0 in (3.2). The estimates of the error in both γ, θ are then $O(L^{\gamma-\theta-1})$ for $M = 1$, from (3.19), and $O(L^{1-M})$ for $M \geq 2$, from (3.20).

4. Properties of the approximants. The general theme of this paper is the approximation of functions which satisfy an ODE of the general form (4.1), using integral approximants which themselves are integral of specially constructed ODEs. We pay particular attention to convergence of "horizontal sequences" of integral approximants. In Sections 2 and 3, the examples show how the rate of convergence of the coefficients of (4.2) may be estimated. In this section, we shall take results having the character of those of the previous sections as hypotheses and use these hypotheses to study the behavior of the integrals of the corresponding ODEs.

Our assumptions in Sections 2 and 3 are compatible with the more general hypothesis that the function to be approximated satisfies an m-th order ODE of the general form,

$$\sum_{j=0}^{m} Q_j(z) f^{(j)}(z) + \Phi(z) = 0 \tag{4.1}$$

where $Q_m(z)$ is a polynomial of degree M_m, and $Q_0, Q_1, \dots, Q_{m-1}(z)$ and Φ are functions which are analytic in $\mathcal{D}_\rho := \{z : |z| \leq \rho\}$ for some $\rho > 1$. We assume that $Q_m(1) = 0$, that $z = 1$ is a regular singular point of (4.1) and that $f(z)$ has a branch point at $z = 1$; we also assume that $Q_m(z) \neq 0$ for all other values of $z \in \mathcal{D}_\rho$. In fact, we assume that $f(z)$ has disk monodromy dimension m with respect to $\mathcal{D}_\rho$ [**2**]. The implication of this assumption is that when $f(z)$ is analytically continued on any path in $\mathcal{D}_\rho$, there are no singular points except $z = 1$ and that precisely m linearly independent coverings of $\mathcal{D}_\rho$ are generated, no matter how often the point $z = 1$ is encircled.

Our sequence of integral approximants for $f(z)$ are solutions of ODEs having the general form

$$\sum_{j=0}^{m} P_j^{(L)}(z) y_L^{(j)}(z) + \pi_L(z) = 0, \tag{4.2}$$

where the polynomials $P_j^{(L)}(z)$ and $\pi_L(z)$ are determined from a knowledge of the power series coefficients of $f(z)$, as described previously. Likewise, the initial coefficients of this power series supply the initial conditions which uniquely determine the required solution of (4.2). To emphasize the dependence of the solution on the choice of the degrees of the polynomials of (4.2), we adopt the notation

$$y^{(L)}(z) = [L/p_0; \dots ; p_{m-1}; M_m]. \tag{4.3}$$

We refer to the sequence with $p_0, p_1, \dots , p_{m-1}, M_m$ fixed and $L \to \infty$ as a *horizontal sequence.*

Our *hypotheses* needed in this section are for L large enough (say $\geq L_0$):

$$P_m^{(L)}(z) \to Q_m(z), \tag{4.4a}$$

$$P_j^{(L)}(z) \to \sum_{k=0}^{p_j} Q_j^{(k)}(1)(z-1)^k/k!, \tag{4.4b}$$

and that the rate of convergence of (4.4a,b) is $O(L^{-\omega})$ uniformly for $z \in \mathcal{D}_\rho$ and some $\omega > 0$. Equation (4.4b) is primarily a hypothesis about convergence of $P_j^{(L)}(z)$ to the p_j-th Taylor section of $Q_j(z)$ about $z = 1$ and not to $Q_j(z)$ itself. The preceding hypotheses suffice for the following general convergence theorem pertaining to the disk $|z| < 1$.

Theorem 4.1. *Under the hypotheses of this section, there exists a finite L_0, such that the horizontal sequence of integral approximants,*

$$\mathcal{H} := \{[L/p_0; p_1; \dots ; p_{m-1}; M_m] : p_0, p_1, \dots , p_{m-1}, M_m \textit{ fixed}, \\ L = L_0, L_0 + 1, L_0 + 2, \dots \},$$

converges to $f(z)$ uniformly in any compact subset of $|z| < 1$, where M_m is defined by (4.1).

Proof. From (4.2) and (4.4a,b), we are led to define

$$\hat{\Phi}(z) := \Phi(z) + \sum_{j=0}^{m-1} \left\{ Q_j(z) - \sum_{k=0}^{p_j} Q_j^{(k)}(1)(z-1)^k/k! \right\} f^{(j)}(z). \tag{4.5}$$

Since the P_m converge to Q_m and $Q_m(z) \neq 0$ except at $z = 1$, and since the P_j also converge, as explained in (4.4) for L large enough ($L \geq L_0$), substitution of these results in (4.2) and comparison with (4.1) and (4.5) leads to the result

$$\lim_{L\to\infty} \pi_L(z) = \hat{\Phi}(z), \qquad |z| < 1. \tag{4.6}$$

Therefore, we can select an L_0 such that, for all $L > L_0$, the approximants (p_j fixed) of (4.3) have the property that $|P_j/P_m|$ and $|\pi_L/P_m|$ are uniformly bounded on any compact subset of $|z| < 1$. If, for a given $z_0, |z_0| < 1$, we choose the path $\mathcal{P} := \{\zeta = rz_0/|z_0|, 0 \leq r \leq 1\}$, then these conditions are sufficient to apply Baker's theorem [**2**] on the uniqueness of convergence and, so, conclude the theorem. □

Our eventual aim is to extend the result of Theorem 4.1 onto Riemann sheets accessible within $|z| \leq \rho$. For the case of integral approximants specified in Section 2, we have progressed toward the objective with the following result.

Theorem 4.2. *Let $f(z)$ satisfy the ODE*

$$(1-z)f'(z) + G(z)f(z) = H(z), \tag{4.7}$$

where $G(1)$ is not an integer, and $G(z)$ and $H(z)$ are analytic in $|z| \leq \rho$, $\rho > 1$. Then

$$\lim_{M\to\infty} \lim_{L\to\infty} [L/M; 1] = f(z), \;\; |z| < 1, |1-z| < \rho - 1 \tag{4.8}$$

on compact subsets of any (finitely numbered) Riemann sheets of $f(z)$ accessible from the disk $|z| \leq \rho$.

Proof. First, we treat the case $\operatorname{Re}(G(1)) < 0$. The solution to (4.7) can be written explicitly as [**15, 17**],

$$f(z) = f_S(z) + f_A(z), \tag{4.9a}$$

(4.9b)

$$f_S(z) = (1-z)^{-\gamma} e^{-g(z)} \left[\frac{f(a)}{(1-a)^{-\gamma}} + \int_a^1 H(\eta)(1-\eta)^{\gamma-1} e^{g(\eta)}\, d\eta \right],$$

$$f_A(z) = (1-z)^{-\gamma} e^{-g(z)} \int_1^z H(\eta)(1-\eta)^{\gamma-1} e^{g(\eta)}\, d\eta, \tag{4.9c}$$

where a is the initial point from which the integration of (4.7) is begun. Using this separation (4.9), we find that $f_A(z)$ is analytic at $z = 1$ because, if we expand the integrand in powers of $(1-\eta)$ and then integrate, we see that the prefactor $(1-z)^{-\gamma}$ exactly cancels all the $(1-z)^{\gamma}$ dependence of the integrand. The function $g(z)$ is defined by

$$g(z) = \int_a^z \frac{G(\zeta)+\gamma}{1-\zeta}\, d\zeta, \quad G(1) = -\gamma. \tag{4.10}$$

Let us select a point r, $\max(0, 2-\rho) < r < 1$, an $M_\varepsilon > \mathrm{Re}\,(\gamma)$, and an $\varepsilon > 0$, so that

$$\left| \sum_{j=0}^{M_\varepsilon - 1} G_j (1-z)^j - G(z) \right| < \varepsilon, \tag{4.11}$$

for all $|1-z| \leq 1-r$, and where $G_j = G^{(j)}(1)/j!$ as in Section 2. We may impose (4.11) by the assumed analyticity of $G(z)$ and Taylor's theorem with remainder. Next, let us select L_ε so that, for all $L \geq L_\varepsilon$,

$$|f(r) - [L/M_\varepsilon - 1; 1](r)| < \varepsilon. \tag{4.12}$$

We may take this step by Theorem (4.1). The $[L/M_\varepsilon - 1; 1]$ has its singularity at z_L, the zero of $\lambda^{(L)}(1-z) + \alpha^{(L)}$. From (2.35), (2.36), we see that

$$z_L = 1 + O(L^{-M_\varepsilon - 1}). \tag{4.13}$$

If we now define $\big(g^{(L)}(t)$ is from (2.4)$\big)$

$$\hat{g}_L(z) = \int_r^z \frac{g^{(L)}(t) + \lambda_L \gamma_L}{\lambda_L (z_L - t)}\, dt, \qquad g^{(L)}(z_L)/\lambda_L = -\gamma_L, \tag{4.14}$$

where $\hat{g}_L(z)$ and, later, $\hat{g}(z)$ will indicate that the initial point for the integration of the differential equation is r [$a = r$ in (4.9)], then we can write the approximant as

$$y_L(z) = y_{L,S}(z) + y_{L,A}(z), \tag{4.15a}$$

(4.15b) $y_{L,S}(z) =$

$$(z_L - z)^{-\gamma L} e^{-\hat{g}_L(z)} \left[\frac{y_L(r)}{(z_L - r)^{-\gamma_L}} + \int_r^{z_L} H_L(\eta)(z_L - \eta)^{\gamma_L - 1} e^{\hat{g}_L(\eta)}\, d\eta \right],$$

(4.15c) $$y_{L,A}(z) = (z_L - z)^{-\gamma L} e^{-\hat{g}_L(z)} \int_{z_L}^{z} H_L(\eta)(z_L - \eta)^{\gamma_L - 1} e^{\hat{g}_L(\eta)}\, d\eta,$$

where, as in (4.9) for $f_A(z)$, here $y_{L,A}(z)$ can be seen to be analytic about $z = z_L$. $H_L(z)$ is defined in (4.17) below. If we consider now a path from $z = r$ to $z = r$ on $|1 - z| = 1 - r$ encircling $z = 1$ n times, we obtain, from (4.9), with $a = r$, and (4.15), the results

(4.16)
$$\begin{aligned} f(r^{[n]}) &= e^{-2\pi i \gamma n} f_S(r^{[0]}) + f_A(r^{[0]}), \\ y_L(z) &= e^{-2\pi i \gamma_L n} \gamma_{L,S}(r^{[0]}) + y_{L,A}(r^{[0]}), \end{aligned}$$

where $r^{[j]}$ is the point on the j-th Riemann sheet which lifts onto r on the 0-th or principal Riemann sheet. The integration of the terms involving H and H_L from $r^{[0]}$ to $r^{[n]}$ disappears by Cauchy's theorem because of the analyticity in the disk $|1 - z| \le 1 - r$. Likewise, $\hat{g}(r^{[0]}) = \hat{g}(r^{[n]}) = 0$ and $\hat{g}_L(r^{[0]}) = \hat{g}_L(r^{[n]}) = 0$. Thus, we obtain result (4.16). By Theorem (4.1) and our hypothesis $\gamma_L \to \gamma$ as $L \to \infty$, we can conclude that, except for the $\int_r^{z_L}$ part, the $y_{L,S}$ term in (4.16) converges to the f_S term. That integral part of (4.15b) is identical to (4.15c), and its convergence will follow immediately when we prove the convergence of (4.15c). It now remains to consider the f_A and $y_{L,A}$ terms.

Following (4.5), we may write

(4.17)
$$\begin{aligned} H_L(z) = \Bigg[H(z) &- \Bigg\{ (G(z) - \sum_{i=0}^{M_\varepsilon - 1} G_i (1-z)^i \Bigg\} f(z) \\ &+ \Bigg\{ (\sum_{i=0}^{M_\varepsilon - 1} (g_i^{(L)} - G_i)(1-z)^i \Bigg\} f(z) \Bigg]_L, \end{aligned}$$

where $[\ ]_L$ denotes the Maclaurin section through z^L, and $g_i^{(L)}$ are the Taylor series coefficients about $z = 1$ of $g^{(L)}(z)$, the approximant to $G(z)$. The first term on the right-hand side of (4.17) is independent of

L and a convergent series in the disk $|z| \le \rho$. The singular multiplying factor in (4.15c), $(z_L - \eta)^{\gamma_L - 1}$, is integrable as, again, $\mathrm{Re}\,(\gamma_L) > 0$ for large enough L by the hypothesis on γ and the convergence of γ_L, and so the contribution from the first term on the right-hand side of (4.17) to (4.15c) tends in the limit $L \to \infty$ to (4.9c).

By (4.11) and Schwarz's lemma, the absolute value of the second term on the right-hand side of (4.17) is bounded in $|1-z| \le 1-r$ by

$$\varepsilon \left[\frac{|1-z|}{1-r} \right]^{M_\varepsilon} |f(z)|. \tag{4.18}$$

This result bounds the order and coefficient of the dominant singularity of the second term on the right-hand side of (4.17). By Darboux's theorem [**10**], we need only consider it to get the asymptotic estimate of the series-truncation error. Therefore, if we expand the term $(1-z)^{M_\varepsilon - \gamma}$ to order z^L, the error is of the order $\sum_{i=L}^{\infty} \binom{M_\varepsilon - \gamma}{l} (-z)^l$. If we use the estimate [**10**] $\binom{M_\varepsilon - \gamma}{l} = O(l^{\gamma - M_\varepsilon - 1})$, then, for $|z| \le 1$, we can bound the series-truncation error by $\sum_{i=L}^{\infty} O(l^{\gamma - M_\varepsilon - 1}) = O(L^{\gamma - M_\varepsilon})$ which goes to zero if $M_\varepsilon > \mathrm{Re}\,(\gamma)$ as we have here. Thus, the integral of the second term is found to be of order ε by the use of (4.11) directly in (4.17) without series truncation, for substitution into (4.15c).

The singular part of the last term in (4.17) when substituted into (4.15c) can, by the estimates of (2.37), be re-expressed as

$$\begin{aligned} \left| \left[\sum_{i=0}^{M_\varepsilon - 1} O(L^{-M_\varepsilon + i}) (1-z)^{i-\gamma} \right]_L \right| &\le \left| \sum_{i=0}^{[\gamma]} O(L^{-M_\varepsilon + i}) \sum_{j=0}^{L} \binom{i-\gamma}{j} (-z)^j \right| \\ &\quad + \left| \left[\sum_{i=[\gamma]+1}^{M_\varepsilon - 1} O(L^{-M+i}) (1-z)^{i-\gamma} \right]_L \right|, \end{aligned} \tag{4.19}$$

where $[\gamma]$ is the greatest integer less than or equal to $\mathrm{Re}\,(\gamma)$. The last term on the right-hand side of (4.19) consists of terms which vanish at $z=1$. By the use of the arguments given above for the second term of (4.17), the truncation error in L vanishes as $L \to \infty$.

The principal part vanishes because of the $O(L^{-M+i})$ coefficient so the whole term converges to zero as $L \to \infty$. In the first term on the right-hand side of (4.19), the binomial coefficients increase as j increases when $\mathrm{Re}\,(\gamma) \geq i+1$. In these terms, we can majorize all the binomial coefficients by $\binom{i-\gamma}{L}$ and, again using the estimate for binomial coefficients, the sums of those terms are bounded by $O(L^{M_\varepsilon-\gamma})$. There remains the term,

(4.20)

$$\Big[O(L^{-M_\varepsilon+[\gamma]})(1-z)^{[\gamma]-\gamma}\Big]_L = O(L^{-M_\varepsilon+[\gamma]})\bigg(1+\sum_{j=1}^{L} O(j^{\gamma-[\gamma]-1})\bigg) = O(L^{-M_\varepsilon+\gamma}),$$

which goes to zero as $M_\varepsilon > \mathrm{Re}\,(\gamma)$. Since z_L also converges to unity by the results of Section 2, we conclude that $y_L(r^{[n]})$ converges to $f(r^{[n]}) + O(\varepsilon)$ as $L \to \infty$. This argument can be extended by taking M_ε large enough to include any point in the lune $|z| < 1$, $|1-z| < \rho - 1$ and any $\varepsilon > 0$.

Now we discuss the case $\mathrm{Re}\,(G(1)) > 0$. Instead of beginning with (4.9), we need to recast the solution of (4.7). First, it is convenient to define

$$F(\eta) = H(\eta)e^{g(\eta)}. \tag{4.21}$$

The required form for the solution of (4.7) is

$$f(z) = f_S(z) + f_A(z), \tag{4.22a}$$

(4.22b)

$$f_S(z) = (1-z)^{-\gamma}e^{-g(z)} \times \left[\frac{f(a)}{(1-a)^{-\gamma}} + \int_a^1 \left\{F(\eta) - \sum_{j=0}^{-1-[\gamma]} \frac{F^{(j)}(1)(\eta-1)^j}{j!}\right\}(1-\eta)^{\gamma-1}\,d\eta + \sum_{j=0}^{-1-[\gamma]} \frac{F^{(j)}(1)(a-1)^j}{(j!)(\gamma+j)}\right],$$

(4.22c)

$$f_A(z) =$$
$$(1-z)^{-\gamma}e^{-g(z)}\int_1^z \left\{F(\eta) - \sum_{j=0}^{-1-[\gamma]} \frac{F^{(j)}(1)(\eta-1)^j}{j!}\right\}(1-\eta)^{\gamma-1}\,d\eta$$
$$- e^{-g(z)} \sum_{j=0}^{-1-[\gamma]} \frac{F^{(j)}(1)(z-1)^j}{(j!)(\gamma+j)}.$$

Now the integrals are convergent and so the separation into the singular and analytic parts is well defined when $\mathrm{Re}\,(\gamma) < 0$. A similar recasting is also required for (4.15). The proof for this case now follows the same method as that given above for the case $\mathrm{Re}\,(\gamma) > 0$. Therefore, Theorem 4.2 follows. $\square$

Note. The preview of this work [**3**] contains at (2.37) an incorrect, miswritten version of (4.8).

Acknowledgment. This work was supported in part by the U.S. DOE (contract W-7405-ENG-36) and, in part, by the NSF (grant INT-8618690). P.R. Graves-Morris is also grateful to the Royal Society for a travel grant.

REFERENCES

1. G.A. Baker, Jr., *The essentials of Padé approximants*, Academic Press, New York, 1975.

2. ———, *Approximate analytic continuation beyond the first Riemann sheet*, Rational approximation and interpolation (P.R. Graves-Morris, E.B. Saff and R.S. Varga, eds.) Springer-Verlag LNIM 1105, Heidelberg, 1984, 285–294.

3. ———, *Integral approximants for functions of higher monodromic dimension*, Non-linear numerical methods and rational approximation (A. Cuyt, ed.) Reidel, Dordrecht, 1984, 3–22.

4. ———, J.L. Gammel and J.G. Willis, *An investigation of the applicability of the Padé approximant method*, J. Math. Anal. Appl. **2** (1961), 405–418.

5. ——— and P.R. Graves-Morris, *Padé approximants*, in *Encyclopedia of Mathematics*, Addison-Wesley/Cambridge University Press, London, **14** 1981, Chapter 1.

6. ———, and D.S. Lubinsky, *Convergence theorems for rows of differential and algebraic Hermite-Padé approximants*, J. Comput. Appl. Math. **18** (1987), 29–51.

7. ———, J. Oitmaa and M.J. Velgakis, *Series analysis of multivalued functions*, Phys. Rev. A **38** (1988), 5316–5331.

8. ———, G.S. Rushbrooke and H.E. Gilbert, *High-temperature series expansions for the spin*-1/2 *Heisenberg model by the method of irreducible representations of the symmetric group*, Phys. Rev. **135**A (1964), 1272–1277.

9. G. Darboux, *Sur l'approximation des fonctions de très-grands nombres et sur une classe étendue de dévelopments en série*, J. Math. **3** (1878), 377–416.

10. D.S. Gaunt and A.J. Guttmann, *Asymptotic analysis of coefficients*, Phase transitions and critical phenomena, Vol. 3 (C. Domb and M.S. Green, eds.), Academic, London, 1974, 181–245.

11. M.L. Goldberger and K.M. Watson, *Collision theory*, J. Wiley, New York, 1964, Chapter 8.

12. A.J. Guttmann, *The high-temperature susceptibility and spin-spin correlation function of the three-dimensional Ising model*, J. Phys. A. **20** (1987), 1855–1863.

13. G.H. Hardy, *Divergent series*, Oxford University Press, London, 1956, Chapter 8.

14. C. Hermite, *Sur la généralisation des fractions continues algébriques*, Ann. of Math. (2) **21** (1893), 289–308.

15. D.L. Hunter and G.A. Baker, Jr., *Methods of series analysis* III. *Integral approximant methods*, Phys. Rev. B **19** (1979), 3808–3821.

16. E.L. Ince, *Ordinary differential equations*, Dover, New York, 1944.

17. E. Kamke, *Differentialgleichungen Lösungsmethoden und Lösungen*, vol. I, Akademische Verlagsgesellschaft, Leipzig, 1951.

18. R. de Montessus de Ballore, *Sur les fractions continues algébriques*, Bull. Soc. Math. France **30** (1902), 28–36.

19. R.G. Newton, *Scattering theory of waves and particles*, McGraw-Hill, New York, 1966.

20. J. Nuttall, *The convergence of Padé approximants to functions with branch points*, Padé and rational approximation, (E.B. Saff and R.S. Varga, eds.) Academic, New York, 1977, 101–109.

21. —–, *Sets of minimum capacity, Padé approximants and the bubble problem*, Bifurcation phenomena in mathematical physics and related topics (C. Bardos and D. Bessis, eds.) Reidel, Dordrecht, 1980, 185–201.

22. ———, *Asymptotics of diagonal Hermite Padé polynomials*, J. Approx. Theory **42** (1984), 299–386.

23. H. Padé, *Sur la répresentation approchée d'une fonction par des fractions rationelles*, Ann. de l'Ecole Normale Sup., 3ième sèrie, **9**, Suppl. (1892), 3–93.

24. ———, *Sur les séries entières convergents ou divergents, et les fractions continues rationelles*, Acta Math. **18** (1894), 97–112.

25. R.E. Shafer, *On quadratic approximation*, SIAM J. Numer. Anal. **11** (1974), 447–460.

26. H. Stahl, *Three different approaches to a proof of convergence for Padé approximants*, Rational approximation and its applications in mathematics and physics (J. Gilewicz, M. Pindor and W. Siemaszko, eds.) Springer-Verlag LNIM 1237, Berlin, 1987, 79–124.

27. H.E. Stanley, *Introduction to phase transitions and critical phenomena*, Oxford University Press, London, 1971, Chapter 9.

Theoretical Division, Los Alamos National Laboratory, Los Alamos NM 87545

Mathematics Department, University of Bradford, Bradford, West Yorkshire BD7 1DP, England