7 Padé Approximants

7.1 From Power Series to Padé Sequences

It can often happen that one can calculate the coefficients c_m up to any order m in the power series

$$f(z) = \sum_{m=0}^{\infty} c_m\, z^m \tag{7.1.1}$$

for a function of interest without being able to determine an expression for the general term and hence, without being able to determine the radius of convergence of the series. If, as also happens rather frequently, one does not possess any other representation for the function, this can pose a serious problem. One has no way of locating the singularities of $f(z)$ and may not even be able to rule out the possibility of a singularity at $z = 0$ itself. Consequently, there is no way of knowing how rapidly the series converges or even whether it converges at all for the values of z one is interested in. A good illustration is provided by perturbation theory calculations of quantum mechanical transition amplitudes. These are expressed as a power series in the "coupling strength" g of the interaction responsible for the transition and one has a formal mechanism for calculating the coefficients in the series up to any desired order. The sum of the calculated terms is then evaluated by setting g equal to its physical value. In the case of the electromagnetic interaction, for example, g is just the fine structure constant $\alpha = e^2 / \hbar c = 1/137$. However, the reason why perturbation theory is used at all in such problems is the impossibility of obtaining exact solutions and the barriers to this often imply an ignorance of the analytic properties of the solutions as functions of g. In other words, one has no a priori knowledge of how or even whether the series converges at the physically relevant value of g. To make matters worse, bound and resonant states manifest themselves as singularities of transition amplitudes. Thus, a perturbation series would appear to be totally irrelevant if one's principal aim is the very important one of identifying such states. Fortunately, one can use a Taylor-like series, whether convergent or not, to define sequences of **Padé approximants** which provide much more versatile representations of the corresponding function than does the power series which is their source.

Definition: The $[L/M]$ Padé approximant to $f(z)$ is the rational function

$$f_{L/M}(z) = \frac{P_L(z)}{Q_M(z)}, \qquad Q_M(0) = 1 \tag{7.1.2}$$

where $P_L(z)$ and $Q_M(z)$ are polynomials of degree L and M respectively, such that $f_{L/M}(z)$ and $f(z)$ have the same first $L + M$ derivatives at $z = 0$; that is, such that

$$\frac{P_L(z)}{Q_M(z)} - \sum_{m=0}^{\infty} c_m\, z^m = \mathrm{O}(z^{L+M+1}) \tag{7.1.3}$$

where $O(z^{L+M+1})$ represents terms of order $L + M + 1$ and higher: $O(z^{L+M+1}) = a\,z^{L+M+1}$ $+b\,z^{L+M+2} + \dots$ for some $a, b, \dots$.

Henri Padé (1863-1953) was born in Abbeville which is a town in the Picardy region of northern France. He graduated with his Agrégation de Mathématiques from the École Normale Supérieure in Paris in 1886 and began a career teaching in secondary schools. He also began a program of mathematical research, publishing his first paper in 1888 and commencing work on a doctoral thesis in 1890. Presented to the Sorbonne in1892, Padé's thesis provided the first systematic study of the representation of functions by rational fractions or what we now call Padé approximants. In 1897 Padé received the first of a series of university appointments culminating in that of Dean of the Faculty of Science at the University of Bordeaux in 1906. In the meantime, in 1899, he published another major work on Padé approximants and by 1908, when he left Bordeaux to become a Rector of the French Academy, had written 41 papers of which 29 were on continued fractions and Padé approximants. He remained a Rector of the Academy, first at Besançon then Dijon and finally Aix-Marseille, until he retired at age 70 in 1934.

A moment's reflection suggests that the class of functions which can be usefully approximated by rational functions is bound to be larger than the class which can be approximated by polynomials. Rational functions have poles of their own and so should be able to provide a representation even in the neighbourhood of singularities. Thus, we are predisposed to expect that the domain of convergence of a sequence of approximants $f_{L/M}(z)$ is a good deal larger than that of the corresponding Taylor series. This expectation will be confirmed; we shall find that Padé approximants provide a method of analytically continuing functions whose definition is provided solely by power series.

The coefficients of the polynomials $P_L(z)$ and $Q_M(z)$ are determined by equating coefficients of like powers of z in

$$P_L(z) - Q_M(z) \sum_{m=0}^{L+M} c_m\,z^m = O(z^{L+2M+1}), \quad Q_M(0) = 1. \tag{7.1.4}$$

The full and unique solution for $f_{L/M}(z)$ is thus found to be

$$f_{L/M}(z) = \frac{\begin{vmatrix} c_{L-M+1} & c_{L-M+2} & \cdots & c_{L+1} \\ \vdots & \vdots & \vdots & \vdots \\ c_L & c_{L+1} & \cdots & c_{L+M} \\ \sum_{k=M}^{L} c_{k-M}z^k & \sum_{k=M-1}^{L} c_{k-M+1}z^k & \cdots & \sum_{k=0}^{L} c_k z^k \end{vmatrix}}{\begin{vmatrix} c_{L-M+1} & c_{L-M+2} & \cdots & c_{L+1} \\ \vdots & \vdots & \vdots & \vdots \\ c_L & c_{L+1} & \cdots & c_{L+M} \\ z^M & z^{M-1} & \cdots & 1 \end{vmatrix}} \tag{7.1.5}$$

where $c_k \equiv 0$ if $k < 0$ and the sums for which the lower limit is larger than the upper limit are also to be replaced by 0. Such a formal and formidable expression does little to inform one's intuition. Therefore, we list below the approximants that correspond to $L, M \leqslant 2$:

$$f_{0/0} = c_0, \quad f_{1/0} = c_0 + c_1 z, \quad f_{2/0} = c_0 + c_1 z + c_2 z^2$$

$$f_{0/1} = \frac{c_0}{1 - \frac{c_1}{c_0} z}, \quad f_{1/1} = \frac{c_0 + \frac{c_1^2 - c_0 c_2}{c_1} z}{1 - \frac{c_2}{c_1} z}, \quad f_{2/1} = \frac{c_0 + \frac{c_1 c_2 - c_0 c_3}{c_2} z + \frac{c_2^2 - c_1 c_3}{c_2} z^2}{1 - \frac{c_3}{c_2} z}$$

$$f_{0/2} = \frac{c_0}{1 - \frac{c_1}{c_0} z + \frac{c_1^2 - c_2 c_0}{c_0^2} z^2}, \quad f_{1/2} = \frac{c_0 + [c_1 + c_0 \frac{c_0 c_3 - c_1 c_2}{c_1^2 - c_0 c_2}] z}{1 + \frac{c_0 c_3 - c_1 c_2}{c_1^2 - c_0 c_2} z + \frac{c_2^2 - c_1 c_3}{c_1^2 - c_0 c_2} z^2}$$

$$f_{2/2} = \frac{c_0 + \left[c_1 + c_0 \frac{(c_1 c_4 - c_2 c_3)}{c_2^2 - c_1 c_3} \right] z + \left[c_2 + \frac{c_1(c_1 c_4 - c_2 c_3) + c_0(c_3^2 - c_2 c_4)}{c_2^2 - c_1 c_3} \right] z^2}{1 + \frac{c_1 c_4 - c_2 c_3}{c_2^2 - c_1 c_3} z + \frac{c_3^2 - c_2 c_4}{c_2^2 - c_1 c_3} z^2} \tag{7.1.6}$$

As this display presages, it is useful to group the approximants to a given function in an infinite array known as a **Padé table**:

$$
\begin{matrix}
f_{0/0} & f_{0/1} & f_{0/2} & f_{0/3} & \cdots \\
f_{1/0} & f_{1/1} & f_{1/2} & f_{1/3} & \cdots \\
f_{2/0} & f_{2/1} & f_{2/2} & f_{2/3} & \cdots \\
\vdots & \vdots & \vdots & \vdots &
\end{matrix} \tag{7.1.7}
$$

The approximants $f_{L/M}$ with L fixed comprise a **row** of the table, while those with M fixed form a **column**. The set $\{f_{M/M}\}$ is called the **diagonal** sequence and together with the **paradiagonal** sequences $\{f_{M+j/M}\}$ with j fixed comprise the ones of most interest. The table subsumes several more specialized types of approximations. For example, the first column consists of the approximations provided by truncating the power series. Of more interest in numerical analysis, the sequence consisting alternately of members of the diagonal and first paradiagonal,

$$f_{0/0}, \; f_{1/0}, \; f_{1/1}, \; f_{2/1}, \; f_{2/2}, \; \ldots, f_{M/M}, \; f_{M+1/M}, f_{M+1/M+1}, \; \ldots,$$

corresponds to the approximations one would obtain by truncating the **continued fraction**

$$f(z) = a_0 + \cfrac{a_1 z}{1 + \cfrac{a_2 z}{1 + \cfrac{a_3 z}{1 + \cfrac{a_4 z}{1 + \ldots}}}}.$$

Mathematicians have yet to establish an extensive theoretical knowledge of the convergence properties of Padé approximants. Numerical experimentation indicates

that the theorems that have been proven offer a limited perspective of the representational potential of Padé sequences. Consequently, following the approach used by Padé and by each successive generation of researchers, we shall examine by means of specific examples how approximants imitate the analytic structure of the functions they represent. Then, after a brief overview of the convergence theorems that have been proven, we will engage in informed speculation on just how powerful a means of analytic continuation Padé approximants really are.

7.2 Numerical Experiments

Let us commence our study of how approximants reconstruct the functions they represent by reminding ourselves that an $[L/M]$ approximant is a meromorphic function with L zeros and M poles where, for counting purposes, we are treating a zero (pole) of order n as though it were n simple zeros (poles). Thus, other meromorphic functions ought to pose a reasonably simple representational challenge. To confirm this, we look first at the trivial case provided by other rational functions.

Example: Substituting the coefficients of the geometric series

$$f(z) = \sum_{m=0}^{\infty} z^m$$

into equation (7.1.5) we see that all Padé approximants except the sequence of truncated series $\{f_{L/0}\}$ exactly reproduce the function they are supposedly approximating; that is, we have

$$f_{L/M} \equiv \frac{1}{1-z} \quad \text{for all} \quad L \geqslant 0, \; M \geqslant 1.$$

This is not a coincidental result. One can readily show that if $f(z)$ is a rational function whose numerator is of degree I and denominator degree J, then

$$f_{L/M} \equiv f(z) \quad \text{for all} \quad L \geqslant I, \; M \geqslant J.$$

Example: A more demanding task is that of reproducing the poles and zeros of a function like $\tan z$. Obviously, this is possible only in the limit that both L and $M \to \infty$. However, we should be able to get an intuitive feel for what will happen in that limit by calculating the first few approximants to $\tan z$ and identifying the location of their poles and zeros. Our starting point is the Taylor series

$$\tan z = z + \frac{1}{3} z^3 + \frac{2}{15} z^5 + \frac{17}{315} z^7 + \ldots$$

whose disc of convergence, $|z| < \pi/2$, is determined by the location of the first two poles, $z = \pm\pi/2$. For calculational convenience, we shall factor out the zero at $z = 0$ and treat the remainder as a series in z^2:

$$\tan z = z f(z^2) \quad \text{where} \quad f(z^2) = 1 + \frac{1}{3} z^2 + \frac{2}{15} (z^2)^2 + \frac{17}{315} (z^2)^3 + \ldots.$$

Using equations (7.1.6) we then find

$$f_{0/1}(z^2) = \frac{3}{3 - z^2}, \quad f_{1/1}(z^2) = \frac{1}{3}\frac{15 - z^2}{5 - 2 z^2},$$

$$f_{2/1}(z^2) = \frac{1}{15}\frac{630 - 45 z^2 - z^4}{42 - 17 z^2}, \quad f_{0/2}(z^2) = \frac{45}{45 - 15 z^2 - z^4},$$

$$f_{1/2}(z^2) = 5\frac{21 - 2 z^2}{105 - 45 z^2 + z^4},$$

which are the only approximants we can calculate in this order. Their poles and zeros are compared with those of $\tan z/z$ in Table 7.1.

Table 7.1: Zeros and Poles of Approximants of $f(z) = \tan z/z$.

Function	Zeros	Poles
$f_{0/1}$	n/a	±1.732
$f_{0/2}$	n/a	±1.601, ±4.191i
$f_{1/1}$	±3.873	±1.581
$f_{2/1}$	±3.348, ±7.497i	±1.572
$f_{1/2}$	±3.240	±1.571, ±6.522
$\frac{\tan z}{z}$	±3.142, ±6.284, ...	±1.571, ±4.712, ...

The sequence consisting of the approximants $f_{0/1}, f_{1/1}, f_{1/2}, \ldots$ acquires poles and zeros in the same alternating order as they occur for the original function. Specifically, $f_{0/1}$ has only poles and no zeros. The poles, at $z = \pm 1.732$, are within about 10% of the first two poles of $\tan z/z$, at $z = \pm\pi/2$. The next approximant in the sequence, $f_{1/1}$, has poles at $z = \pm 1.581$ and, in addition, has zeros at $z = \pm 3.873$. In other words, we have now reproduced $z = \pm\pi/2$ to within better than 1%, as well as the first two zeros of $\tan z/z$, $z = \pm\pi$, to within about 25%. Finally, $f_{1/2}$ locates these zeros to within less than 3% and provides us with a first approximation to the location of the next pair of poles at $z = \pm 3\pi/2$. Each successive approximant attempts to replicate the next pair of poles or zeros in the same order as they are encountered with increasing values of $|z|$. Moreover, each successive approximant replicates all previously located poles and zeros with ever increasing accuracy. What is perhaps most impressive is that this detailed picture of how $\tan z/z$ behaves for $|z| \geqslant \pi/2$ has been generated from a very restricted knowledge of how that function behaves for $|z| < \pi/2$, namely, a few terms of the function's Taylor expansion about $z = 0$. Thus, this Padé sequence provides a very powerful method of analytically continuing the Taylor expansion outside its circle of convergence.

The other two approximants listed in Table 7.1 are also worthy of note. While $f_{0/2}$ reproduces the poles at $z = \pm\pi/2$ with much greater accuracy than does $f_{0/1}$, it also produces extraneous poles on the imaginary axis. Similarly, $f_{2/1}$ has extraneous zeros at $z = \pm 7.5i$. These features are attributable to the structural inability of these particular approximants to mimic the alternating character of the poles and zeros of the

function they represent. If we were to go to higher order, we would find that the extraneous poles and zeros are very unstable with respect to position. This is because they do not reflect specific analytic features possessed by the original function in the finite plane but rather, its behaviour at infinity. In any case, this illustrates that there is a tangible advantage to be gained from being selective in one's choice of Padé sequence to represent a given function.

Example: To further illustrate the phenomenon of extraneous poles and zeros, let us consider the approximants of e^z which not only has no poles of its own, it also has no zeros. The [1/1] approximant is

$$f_{1/1} = \frac{2 + z}{2 - z}$$

with a zero on the negative real axis at $z = -2$ and a symmetrically placed pole on the positive real axis at $z = +2$. The [2/2] approximant is

$$f_{2/2} = \frac{12 + 6z + z^2}{12 - 6z + z^2}$$

with zeros at $z = -3 \pm \sqrt{3}i$ and poles at $z = +3 \pm \sqrt{3}i$. Thus, we perceive an interesting pattern emerging. The zeros all occur in the left half-plane with the poles in mirror-image locations in the right half-plane. Moreover, on increasing the order of the approximant, the poles and zeros have moved further away from the origin. This pattern persists as one further increases the order. The poles occur either on the real axis or on either side of it in conjugate pairs but in any case, moving ever further to the right. The zeros cluster on or about the negative real axis and move further and further to the left.

Recalling that $|e^z|$ increases without limit as $|z| \to \infty$ along the positive real axis and goes to zero as $|z| \to \infty$ along the negative real axis, this pattern becomes understandable. The poles and zeros of the approximants are simulating the exponential function's essential singularity at $z = \infty$.

This simulation evidently assists the convergence of the approximants. As one can check, an approximant gives a much better approximation of the value of e^z in the region between the poles and zeros than does the corresponding truncated Taylor series. Even at $z = -1$, a point at which the Taylor series converges fairly rapidly, the approximants improve on the accuracy of the approximation by one to two orders of magnitude. This is shown in Table 7.2 where we see that the [M/M] approximants increase their accuracy by two decimal places per unit increase in M, with the value given by $f_{4/4}$ being off by only one part in 10^8.

Example: Isolated essential singularities are not the only type of non-polar singularity that can be simulated by Padé approximants. A function with branch points is necessarily discontinuous across a cut or cuts joining the points. Simulating a discontinuity of this sort might appear to be an overly ambitious task for a sequence of functions that are continuous everywhere except for a finite set of poles. However, by clustering their poles along the cut the Padé approximants produce a representation that, very often, is convergent **everywhere** in the cut plane.

Table 7.2: $[M/M]$ Approximants of e^z Evaluated at $z = -1$.

M	$f_{M/M}$	% Deviation from e^{-1}	$\sum\limits_{m=0}^{2M} \dfrac{(-1)^m}{m!}$	% Deviation from e^{-1}
1	0.33333333	9.39	0.5	35.9
2	0.36842105	0.147	0.375	1.94
3	0.36787564	1.03×10^{-3}	0.36805555	4.79×10^{-2}
4	0.36787945	2.45×10^{-6}	0.36788194	6.79×10^{-4}
Limit	0.367879441	–	0.367879441	–

The efficacy of the simulation is perhaps best appreciated by recalling the dispersion representation for functions with branch points. For example, the function $(1+z)^{-1/2}$, with branch points at $z = -1$ and $z = \infty$, has the representation

$$(1+z)^{-1/2} = \frac{1}{\pi} \int\limits_1^\infty \frac{d\chi}{(\chi + z)(\chi - 1)^{1/2}}, \tag{7.2.1}$$

corresponding to a cut along the real axis segment $-\infty < \chi \leqslant -1$. Using this as our starting point, we can give $(1+z)^{-1/2}$ an approximate representation by simply evaluating the integrand at a finite number of carefully selected points. Such an operation is precisely that of replacing the cut by a finite number of simple poles each located at the negative image of one of the pre-selected points and therefore, distributed along the line segment where the cut used to be. Of course, a truly discontinuous function results only in the limit as the number of such poles becomes infinite.

As we know, the choice of simple curve used as a cut joining the two branch points is arbitrary. If the choice is different from the negative real axis, it would be reflected in a corresponding change to the integration contour in (7.2.1) and hence, in the location of the poles simulating the cut were we to continue to use the integral as the basis of an approximate representation. However, there is no such flexibility with the approximate representation obtained with Padé approximants. Since the location of the branch cut has no effect on the value of $(1+z)^{-1/2}$ and its derivatives at the origin, the approximants necessarily make an independent but specific choice of the curve to use in simulating the cut. The only question is whether this choice is predictable. The answer is an unqualified yes only in the case of a class of functions that are named after a mathematician called **Stieltjes**; they will be the subject of more detailed study a little further on. For more general functions one must rely on systematic behaviour observed in the course of numerical experimentation to justify any inference as to the location of the simulated cuts.

G.A. Baker Jr. and his collaborators have experimented with a variety of functions possessing two, three, or even four branch points. They found that the limit set of poles of diagonal Padé approximants forms circular arcs each of which joins a pair of branch points and, if extrapolated to form a complete circle, would pass through the origin. As usual we include straight lines in our definition of circle and so, in the case

of $(1 + z)^{-1/2}$, the poles of the $[M/M]$ approximants conform to this rule by clustering along the real axis segment $-\infty < x < -1$.

Another interesting function studied by Baker et al. is

$$f(z) = \frac{(1 + z^2)^{1/2}}{1 + z}.$$

It has branch points at $z = \pm i$. To conform with the empirical rule, its $[M/M]$ approximants should simulate a cut which runs from $z = +i$ through $z = \infty$ to $z = -i$. As Figure 7.1 shows, this expectation is born out by actual calculation: one pole in the limit set converges to $z = -1$ and the rest cluster along the imaginary axis segments $1 < y < \infty$ and $-\infty < y < -1$.
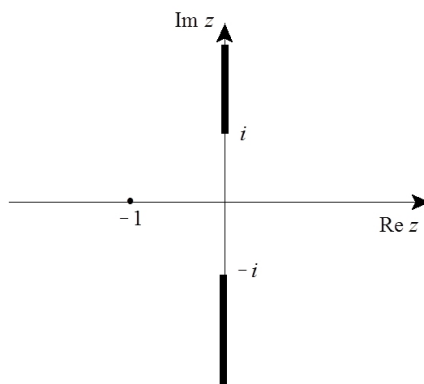

Figure 7.1: The poles of the $[M/M]$ approximants to $f(z) = (1 + z^2)^{1/2}/(1 + z)$ simulate a cut along the imaginary axis from $z = +i$ to $z = -i$, passing through $z = \infty$, as well as reproduce the polar singularity at $z = -1$.

The behaviour of these approximants at infinity is rather interesting as it reveals how cuts adversely affect convergence. From the calculated values displayed in Table 7.3 we see that the $[M/M]$ approximants converge to $\sqrt{2} \pm 1$ according as M is even or odd. Thus, this Padé sequence does not converge to anything at $z = \infty$. The sequences $[2M/2M]$ and $[2M + 1/2M + 1]$ converge separately but not to ± 1 which are the values attained by the function $(1 + z^2)^{1/2}/(1 + z)$ on alternate lips of the cut. This suggests that one should not expect convergence of Padé approximants at points located on a cut.

7.3 Stieltjes Functions

Thomas Jan Stieltjes (1856-1894) had little formal education, flunking out of the Polytechnical School of Delft in 1876, due in large part to an all-consuming passion for math-

Table 7.3: The $[M/M]$ Approximants to $(1 + z^2)^{1/2}/(1 + z)$ Evaluated at Infinity

M	$[M/M]$
1	0.333333
2	2.333333
3	0.411764
4	2.411764
5	0.414141
6	2.414141
7	0.414211
8	2.414211
Limit	$\sqrt{2} \pm 1$
Actual Value of Function	**±1**

ematics. In 1882, while employed as an assistant at the Leiden Observatory, Stieltjes began a correspondence with the distinguished mathematician Hermite. Over the remaining 12 years of Stieltjes' life, they exchanged 432 letters on his mathematical interests. Hermite became both his mentor and his sponsor, helping Stieltjes obtain academic appointments in the Netherlands and in France. Stieltjes worked on almost all branches of analysis, continued fractions and number theory but he is best remembered for the Stieltjes function defined below.

Everything that has been empirically deduced about the behaviour of Padé approximants of functions with branch points can be rigorously proven to occur for **Stieltjes functions**. These possess either a Taylor or an asymptotic series about $z = 0$ whose coefficients can be expressed in terms of a very particular type of integral.

Definition: A function $f(z)$ which admits representation by a series of the form

$$f(z) = \sum_{m=0}^{\infty} c_m(-z)^m \tag{7.3.1}$$

is a **Stieltjes function** if and only if

$$c_m = \int_0^{\infty} t^m \, dg(t) \tag{7.3.2}$$

where $g(t)$ is a bounded non-decreasing function which takes on infinitely many values in the interval $0 \leqslant t < \infty$.

The series itself is called a **series of Stieltjes** and need not be convergent. However, if it is, with radius of convergence R, then we can interchange the order of integration and summation in (7.3.1) to obtain

$$f(z) = \int_0^{\infty} \sum_{m=0}^{\infty} (-tz)^m \, dg(t), \tag{7.3.3}$$

for $|z| < R$. We know that

$$\sum_{m=0}^{\infty}(-tz)^m = \frac{1}{1+tz} \text{ for } |tz| < 1$$

and that it diverges for $|tz| \geqslant 1$. Therefore, if (7.3.3) is to be meaningful, $f(z)$ must possess the following Stieltjes integral representation

$$f(z) = \int_0^{1/R} \frac{g'(t)dt}{1+tz}. \tag{7.3.4}$$

This provides an analytic continuation of $f(z)$ outside the circle of convergence of (7.3.1) and permits identification of the function's singularities. Indeed, if we now make the substitution $t \to x = -\frac{1}{t}$, (7.3.4) assumes the more familiar features of a dispersion representation for a real function (see equation (4.1.5):

$$f(z) = \frac{1}{\pi}\int_{-\infty}^{-R} \frac{\operatorname{Im}f_+(\chi)}{\chi - z}d\chi \tag{7.3.5}$$

$$\operatorname{Im}f_+(x)\Big|_{x=-1/t} = \lim_{\varepsilon \to 0}\operatorname{Im}f[-(t+i\varepsilon)^{-1}] = -\pi t\frac{dg(t)}{dt}. \tag{7.3.6}$$

Thus, $f(z)$ is a real function, $f^*(z^*) = f(z)$, and is holomorphic everywhere in the complex plane except for a cut along the negative real axis from $z = -R$ to infinity.

One can show that these two integral representations retain their validity even in the limit of a divergent series of Stieltjes, $R \to 0$, provided that the series coefficients are such that

$$\sum_{m=1}^{\infty}(c_m)^{-1/(2m+1)}$$

also diverges. This condition is roughly equivalent to requiring that $|c_m| \leqslant (2m)!$. The cut along the negative real axis then extends from $z = 0$ to infinity and (7.3.4) becomes

$$f(z) = \int_0^{\infty} \frac{g'(t)dt}{1+tz}. \tag{7.3.7}$$

As for the series, it too retains a representational role albeit only an asymptotic one. In conformance with equation (6.1.7), we have

$$\lim_{|z| \to 0}\left\{z^{-n}\left[f(z) - \sum_{m=0}^{n}c_m z^m\right]\right\} = 0 \tag{7.3.8}$$

for any positive integer n and arg z restricted to avoid the cut joining the branch points at $z = 0$ and $z = \infty$.

Notice that the properties of $g(t)$ necessarily imply that $\text{Im}\, f_+(x) \leqslant 0$.

In summary, we can characterize all Stieltjes functions as satisfying the reality condition $f^*(z^*) = f(z)$, as possessing branch points at $z = \infty$ and a point on the negative real axis $z = -x_0$, $x_0 \geqslant 0$ and, when these points are joined by a cut along that axis, as having a negative definite, pure imaginary discontinuity across the cut. Indeed, given a function with these properties, the choice of the negative real axis as the location of its cut assures the existence of a Stieltjes integral representation since we can write

$$f(z) = \int\limits_0^\infty \frac{g'(t)dt}{1 + tz} \tag{7.3.9}$$

with the weight function given by

$$g(t) - g(t_0) = \frac{1}{\pi} \int\limits_{t_0}^t \lim_{\varepsilon \to 0} \text{Im}[f(-(\tau - i\varepsilon)^{-1})] \frac{d\tau}{\tau}. \tag{7.3.10}$$

Differentiation of (7.3.9) then generates a series of Stieltjes about $z = 0$.

Baker's empirical rule concerning the simulation of cuts by Padé approximants can actually be proven to be the case for Stieltjes functions. The relevant theorem also provides remarkably detailed information about the poles and zeros of successive approximants.

Theorem: If $\sum\limits_{m=0}^\infty c_m(-z)^m$ is a series of Stieltjes, the poles and zeros of the $[M + J/M]$, $J \geqslant -1$ Padé approximants obtained from the coefficients are on the negative real axis. Furthermore, the poles of successive approximants interlace and all the residues are positive. The roots of the numerators (the zeros) also interlace those of the corresponding denominators (the poles).

Example: Stieltjes functions are more common place than one might at first suspect. Whether by means of an adroitly chosen transformation or by more immediate measures, a very wide class of analytic functions can be cast in a Stieltjes form. Among those which are immediately recognizable as possessing properties typical of a Stieltjes function is the elementary function $f(z) = \frac{1}{z}Ln(1+z)$. It admits the following series, Stieltjes and dispersion representations:

$$\frac{1}{z}\, Ln(1 + z) = \begin{cases} 1 - \frac{z}{2} + \frac{z^2}{3} - \frac{z^3}{4} + - \ldots, & |z| < 1 \\[2mm] \int\limits_0^1 \frac{dt}{1+tz}, & -\pi < \arg(z + 1) < \pi \\[2mm] \int\limits_{-\infty}^{-1} \frac{d\chi}{\chi(\chi-z)}, & -\pi < \arg(z + 1) < \pi. \end{cases}$$

In conformance with the preceding theorem, the poles and zeros of the diagonal approximants to this function alternate along the negative real axis to the left of $z = -1$. As the order of the approximants increases, the spacing between successive poles

and zeros decreases due to the interlacing effect called for by the theorem. Thus, the picture that emerges of the limiting situation as $M \to \infty$ is of an infinite sequence of very closely spaced, alternating poles and zeros extending along the negative real axis from $z = -1$ to infinity.

As we shall see, this single-minded behaviour of the poles and zeros offers the entire cut complex plane as the domain of convergence for the diagonal approximants. However, to illustrate how an alternative and, as it turns out, inappropriate choice of Padé sequence can have a much more restricted domain of convergence, we shall follow Baker's lead and explore how the $[4M/M]$ approximants to $\frac{1}{z}Ln(1+z)$ behave.

Table 7.4: $[4M/M]$ Approximants to $\frac{1}{z}Ln(1+z)$

$z\ M$	1	2	3	4	$\frac{1}{z}Ln(1+z)$
1.0	0.69242424	0.69314873	0.69314718	0.69314718	0.69314718
6.0	−0.80000000	2.1535393	−2.6374816	5.1116289	0.32431836
9.0	−4.3735294	28.911655	−176.68475	1091.7095	0.25584279

Table 7.4 compares the values of the first four of these approximants with the corresponding values of $\frac{1}{z} Ln(1 + z)$ at a series of points on the positive real axis. We see that for $z = 1.0$ convergence to the exact value is almost immediate. For $z = 4.0$ (not shown in the table) convergence is slower but it still takes place. However, for $z = 6.0$ and 9.0 the approximants are oscillating in sign and, in the latter case particularly, diverging quite rapidly.

This startling change in behaviour is directly associated with the extraneous zeros possessed by these approximants. Each approximant requires only M zeros to alternate with the M poles along the cut. Since Stieltjes functions do not vanish anywhere in the cut plane, this leaves $3M$ zeros with no role to play in the reconstruction of the analytic properties of $\frac{1}{z} Ln(1 + z)$. Plotting these zeros for the four approximants in Table 7.4, (Figure 7.2), one finds that they form a pattern suggestive of a closed curve surrounding a roughly heart-shaped region about the origin. Evidently, this curve will only increase in definition with increasing M and so constitutes a natural boundary for the approximants' domain of convergence.

This example illustrates a general rule for the approximation of functions possessing branch points: to enjoy a maximal domain of convergence one should restrict consideration to diagonal or para-diagonal Padé sequences.

Example: There has been speculation that the perturbation series obtained in quantum electrodynamics is an asymptotic rather that a convergent series owing to a singularity at $\alpha = 0$, where α is the electromagnetic coupling strength; (the physical value of α is $e^2 / \hbar c = 1/137$). Thus, it may be useful to know something about the convergence properties of Padé sequences derived from such series.
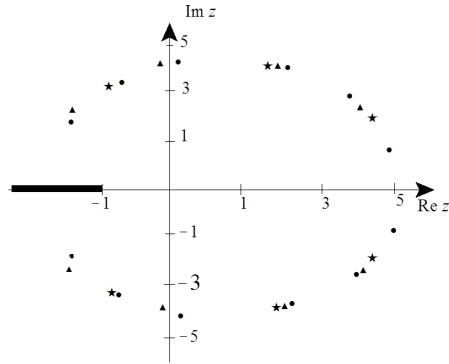
Figure 7.2: The extraneous zeros of the $[4M/M]$ Padé approximants to $[\mathrm{Ln}(1 + z)]/z$. These zeros define the boundary of the domain of convergence of this sequence of approximants; (from Baker, (1975)).

Based on the last example, we might expect favourable results, for diagonal sequences at least, if the series are series of Stieltjes. This expectation is born out by both experiment and theory. An interesting case in point is provided by the Stieltjes function

$$f(z) = \int\limits_{0}^{\infty} \frac{e^{-t}}{1 + zt} dt, \tag{7.3.11}$$

which has branch points at $z = 0$ and ∞. Its asymptotic expansion about $z = 0$ is the divergent series

$$f(z) \sim 1 - (1!)z + (2!)z^2 - (3!)z^3 + \ldots = \sum\limits_{m=0}^{\infty}(-1)^m\, m!\, z^m \tag{7.3.12}$$

which makes it a simplified analogue of the electrodynamics perturbation series since the latter's general term is expected to go like $m!\, \alpha^m$ for large values of m. Calculating the $[M/M]$ approximants to $f(z)$, one finds that they do converge towards the exact values defined by (7.3.11) and do so for **all** finite z in the cut plane. Although the rate of convergence is rather slow, it is essentially unchanged as $|z|$ increases. For example, at $z = 1.0$ the $[6/6]$ approximant has the value 0.5968 compared to an exact value of 0.5963, while at infinity $f_{6/6} = \frac{1}{7}$ compared to an exact limiting value of zero. (As $x \to \infty, f(x) \to 0$ and $f_{M/M}(x) \to \frac{1}{M+1}$). In both cases the error is proportional to $\frac{1}{M+1}$.

This leap from a divergent series to a sequence that is convergent throughout the cut plane is a little breathtaking. It is this feature, as well as the ability to locate and identify singularities of functions defined by power series alone, that best exemplifies the analytical power of Padé approximants.

7.4 Convergence Theorems

As we have already noted, the theory of convergence of Padé sequences is still far from complete. We need be aware of only a few highlights.

The earliest theorems to be proved apply to columns of the Padé Table. For example, de Montessus de Ballore established in 1902 that, if $f(z)$ is holomorphic throughout the domain $|z| < R$, except for a finite number of poles of total multiplicity m, then the sequence $\{f_{L/m}(z)\}$ converges uniformly to $f(z)$ for $|z| \leqslant \rho < R$, except at the poles of $f(z)$. This theorem was generalized around 1930 to prove that the column sequences $\{f_{L/m+\mu}(z)\}, \mu \geqslant 1$ are also uniformly convergent for $|z| \leqslant \rho < R$, provided that the singularities of $f(z)$ on $|z| = R$ are no worse than a multiple pole of order greater than μ or a branch point. Moreover, the μ additional poles, not needed to represent the poles of $f(z)$ inside $|z| < R$, simulate the singularities on $|z| = R$ with the first extra pole representing the "strongest" of these singularities, the second the next strongest, and so on.

Since one can readily prove that $[f_{L/M}]^{-1}$ is the $[M/L]$ approximant to $f^{-1}(z)$ if $f_{L/M}(z)$ is an approximant to $f(z)$, theorems concerning columns of the Padé Table and poles in the z-plane translate directly into theorems concerning rows and zeros.

From this modest beginning we move on to a series of theorems established in the 1960's by Chisholm (1966) and Beardon (1968). These concern the convergence of sequences $\{f_{L/M}(z)\}$ when **both** L and $M \to \infty$. If $f(z)$ is meromorphic in a closed disc $|z| \leqslant R$, the sequence $\{f_{L/M}(z)\}$ converges uniformly to $f(z)$ in the double limit $L, M \to \infty$ provided that z is restricted to a second disc $|z| \leqslant \rho \leqslant R$ from which small neighbourhoods of the poles of $f(z)$ and of any limit points of the poles of the approximants have been removed. The radius ρ of the disc of convergence depends on the ratio L/M and, in general, is less than R. However, equality with R can be guaranteed if we are prepared to restrict ourselves to sequences for which $L \geqslant kM$ where k is a number $\geqslant 1$ that depends on both the value of R and the identity of $f(z)$.

While helpful as an aid to our intuition, these theorems do little to meet our theoretical needs. For mathematicians and physicists alike, the diagonal and paradiagonal sequences are much the most important elements of the Padé Table. This importance stems in large part from certain properties which are uniquely associated with diagonal approximants. For example, they can be shown to be invariant under mappings of the form $w = \frac{az}{cz+d}$, a, c and d constants.

Thus, it suffices to establish convergence in a restricted domain since one can immediately extend it in size by means of one of these mappings. Further, diagonal approximants can be shown to be **unitary** in the sense that if $f^{\star}(z) = f^{-1}(z)$, then $f_{N/N}^{\star}(z) = [f_{N/N}(z)]^{-1}$. This is a critically important property in quantum mechanical applications. So, what we really need are convergence theorems that apply specifically to diagonal approximants and cover the full range of analytic functions.

The results of numerical experimentation suggest that the boundary of the domain of convergence of a diagonal sequence should be determined only by the loca-

tion of the non-polar singularities of the function it represents. However, very little progress has been made toward proving this assertion for more than a few special cases which fortunately include Stieltjes functions.

To conclude this brief survey of Padé convergence theory, we state the theorem that applies to Stieltje's functions. Given our discussion of numerical experience with approximants, it contains no surprises.

Theorem: Let $\sum_{m=0}^{\infty} c_m(-z)^m$ be a series of Stieltjes. If the series is convergent with a radius of convergence R, then any sequence of $[M+J/M], J \geqslant -1$ Padé approximants to the series converges in the cut plane $(-\infty < z \leqslant -R)$ to the function $f(z)$ defined by the series. If the series is divergent, then any sequence of $[M+J/M], J \geqslant -1$ Padé approximants converges to an analytic function in the cut plane $(-\infty < z \leqslant 0)$. If, in addition, $\sum_{m=0}^{\infty} (c_m)^{-1/(2m+1)}$ diverges, then all the sequences tend to a common limit $f(z)$ which possesses a Stieltjes representation of the form given by (7.3.9) and (7.3.10) and an asymptotic expansion about $z = 0$ given by the original series.

The proof of the various parts of this theorem is the subject of an entire chapter of Baker's 1975 monograph on Padé approximants.

7.5 Type II Padé Approximants

As an alternative to defining Padé approximants by means of the value of a function and its derivatives at a single point, one can build approximants that contain information at two or more points. An especially useful case arises when we use only the value of the function (and not of any of its derivatives) at an appropriate number of points. These are called Padé approximants of the second type or type II.

Definition: Let $z_1, z_2, \ldots z_N$ be N complex numbers and $f(z)$ be an analytic function which takes the values $f(z_i)$ at these points. We define the $[L/M]$ **Padé approximant of type II** to $f(z)$ to be the ratio of two polynomials in z, $P_L(z)$ and $Q_M(z)$ with $L+M = N-1$, which takes the values $f(z_i)$ at $z = z_i$,

$$f_{L/M}(z) \equiv \frac{P_L(z)}{Q_M(z)}$$

$$f_{L/M}(z_i) = f(z_i), \quad i = 1, \ldots N. \tag{7.5.1}$$

Type II approximants provide an extension of the Lagrange or polynomial method of interpolating between discrete points in the same way that their type I counterparts extend the Taylor or polynomial approximation of functions known only in the neighbourhood of $z = 0$. They are unique and one can prove (see, for example, J. Zinn-Justin, Phys. Reports 1C, No. 3 (1971)) that they possess essentially the same properties as do approximants of the first type, including convergence in the case of approximation of Stieltjes functions.

The interpolation provided by type II approximants is both more efficient and more effective at analytic continuation outside the region of the points z_i than that given by a Lagrange polynomial approach. Once again this is due to the capacity of Padé approximants to replicate the poles and simulate the non-polar singularities of the functions they approximate.

Example: If one builds the type II Padé approximants to $\tan z$ using values corresponding to z real and in the range $-\pi/2 < z < \pi/2$, one finds that the lowest orders give the nearby poles and zeros of $\tan z$ and that by increasing the order (number of points), more and more poles and zeros are successfully replicated. Moreover, this remains true even if one contracts the range of z to $-\pi/4 < z < \pi/4$ where $\tan z$ is well away from any of its poles and so is a relatively slowly varying function (in fact, it resembles a straight line there).

Another interesting application of type II approximants is in the summation of numerical series. If we have

$$S = \sum_{m=0}^{\infty} u_m, \tag{7.5.2}$$

we can treat the partial sums $S_N = \sum\limits_{m=0}^{N} u_m$ as functions of $\frac{1}{N}$,

$$S_N = f\left(\frac{1}{N}\right) \quad \text{say,}$$

so that the sum of the series is

$$S = f(0). \tag{7.5.3}$$

We can then compute S by extrapolating from $f(1), f\left(\frac{1}{2}\right), f\left(\frac{1}{3}\right), \ldots$ to $f(0)$; that is to say, by analytically continuing $f(z)$ to $z = 0$.

Example: Consider the series

$$S = \sum_{m=0}^{\infty} \frac{1}{m^2}$$

which is known to converge slowly. Using the first $N + M + 1$ partial sums as input, we can build the $[N/M]$ type II approximant to $f(z)$ and thus get an estimate of $f(0)$.

The first three partial sums are

$$S_1 = f(1) = 1 \quad S_2 = f\left(\frac{1}{2}\right) = 1.25 \quad S_3 = f\left(\frac{1}{3}\right) = 1.361.$$

These yield the approximant estimates

$$f_{[0/0]}(0) = 1 \quad f_{[0/1]}(0) = 1.5 \quad f_{[1/1]}(0) = 1.65.$$

The exact value is $f(0) = S = \pi^2/6 = 1.645\ldots$. Thus, convergence has been accelerated in a quite remarkable way; while the partial sums themselves give a very crude

estimate of S, using the same information as input for the [1/1] Padé approximant yields an estimate that is accurate to three significant figures.

We conclude with a reference to the use of type II approximants in the solution of algebraic equations. Suppose that we wish to solve the equation $F(z) = 0$. We begin by determining the values F_1, F_2, F_3 assumed by F at the three points z_1, z_2, z_3, respectively. Next, we build the type II approximant, $F^{(1)}_{[1/1]}(z)$, which has the same values F_i for $z = z_i$, $i = 1, 2, 3$. One can immediately read off the zero z_4 of $F^{(1)}_{[1/1]}$ which gives a first estimate of a zero of $F(z)$. Using the actual value $F_4 \equiv F(z_4)$, we repeat the process using F_2, F_3, F_4 and the points z_2, z_3, z_4 as input and determine $F^{(2)}_{[1/1]}(z)$. Its zero is a second estimate of the zero of $F(z)$. Evidently, this can be repeated until we are satisfied with the degree of convergence of our estimates yielding, if it works, a remarkably simple method of solving our original equation. The question of how well it works has been addressed by Zinn-Justin who shows that the rate of convergence from a distance is quite fast and that once we are close, it is exponential; if we are within an error ε of the zero at some point in the process, a further iteration will bring us within an error ε^3. Thus, in two steps we can go from an error of 10% to one of 1 part in 10^{-9}.