Chapter 8

Summation Techniques, Padé Approximants, and Continued Fractions

8.1 Accelerated Convergence

Conditionally convergent series, such as

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}\dots=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}=\ln 2, \tag{8.1}$$

converge *very* slowly. The same is true for absolutely convergent series, such as

$$\sum_{n=1}^{\infty}\frac{1}{n^2}=\zeta(2)=\frac{\pi^2}{6}. \tag{8.2}$$

If we call the partial sum for the latter

$$\sum_{n=1}^{N}\frac{1}{n^2}=S_N, \tag{8.3}$$

the difference between the limit S and the Nth partial sum is

$$S-S_N=\sum_{n=N+1}^{\infty}\frac{1}{n^2}\approx\int_N^{\infty}\frac{dn}{n^2}=\frac{1}{N}, \tag{8.4}$$

which means that it takes 10^6 terms to get 6-figure accuracy.

Thus, to evaluate a convergent series, the *last* thing you want to do is actually literally carry out the sum. We need a method to accelerate the convergence, and get good accuracy from a few terms in the series. There are several standard methods.

8.1.1 Shanks' Transformation

The Shanks transformation is good for alternating series, or oscillating partial sums, such as Eq. (8.1). For the series

$$S = \sum_{n=1}^{\infty} a_n, \tag{8.5}$$

consider the Nth partial sum

$$S_N = \sum_{n=1}^{N} a_n. \tag{8.6}$$

Let us suppose that, for sufficiently large N,

$$S_N = S + Ab^N, \tag{8.7}$$

where $-1 < b < 0$, so that as $N \to \infty$, $S_N \to S$. We will take this as an ansatz for all N, to obtain an estimate for the limit S. Then, successive partial sums satisfy

$$S_{N-1} = S + Ab^{N-1}, \tag{8.8a}$$
$$S_N = S + Ab^N, \tag{8.8b}$$
$$S_{N+1} = S + Ab^{N+1}, \tag{8.8c}$$

so that

$$b = \frac{S_{N+1} - S}{S_N - S} = \frac{S_N - S}{S_{N-1} - S}, \tag{8.9}$$

which may be immediately solved for S,

$$S_{(N)} = \frac{S_{N+1}S_{N-1} - S_N^2}{S_{N+1} + S_{N-1} - 2S_N}, \tag{8.10}$$

where now we've inserted the (N) subscript on the left to indicate this is an estimate for the limit, based on the N, $N+1$, and $N-1$ partial sums.

For the series (8.1) the first 5 partial sums are

$$S_1 = 1, \quad S_2 = \frac{1}{2} = 0.5, \quad S_3 = \frac{5}{6} = 0.833, \quad S_4 = \frac{7}{12} = 0.5833,$$
$$S_5 = \frac{47}{60} = 0.7833, \tag{8.11}$$

which oscillate around the correct limit $\ln 2 = 0.693147$, but are not good approximations. Using the Shanks transformation (8.10) we obtain much better approximants:

$$S_{(1)} = \frac{7}{10} = 0.700, \quad S_{(2)} = \frac{29}{42} = 0.690, \quad S_{(3)} = \frac{25}{36} = 0.6944, \tag{8.12}$$

which use only the first 3, 4, and 5 terms in the original series. We can do even better by iterating the Shanks transformation,

$$S^{[2]}_{(N)} = \frac{S_{(N+1)}S_{(N-1)} - S^2_{(N)}}{S_{(N+1)} + S_{(N-1)} - 2S_{(N)}}, \tag{8.13}$$

and then we find using the same data (only 5 terms in the series)

$$S^{[2]}_{(2)} = \frac{165}{238} = 0.693277, \tag{8.14}$$

an error of only 0.02%! For more detailed comparison of Shanks estimates for this series, see Table 8.2 on page 373 of Bender and Orzag.

8.1.2 Richardson Extrapolation

For monotone series, Richardson extrapolation is often very useful. In this case we are considering partial sums S_N which approach their limit S monotonically. In this case we assume an asymptotic form for large N

$$S_N \sim S + \frac{a}{N} + \frac{b}{N^2} + \frac{c}{N^3} + \dots. \tag{8.15}$$

The first Richardson extrapolation consists of keeping only the first correction term,

$$S_N = S + \frac{a}{N}, \quad S_{N+1} = S + \frac{a}{N+1}, \tag{8.16}$$

which may be solved for the limit

$$S^{[1]}_{(N)} = (N+1)S_{N+1} - NS_N, \tag{8.17}$$

where again we've inserted on the left a superscript [1] indicating the first Richardson extrapolation, and a subscript (N) to indicate the approximant comes from the Nth and $N+1$st partial sums.

We consider as an example Eq. (8.2). Here, the first 4 partial sums are

$$S_1 = 1, \quad S_2 = \frac{5}{4} = 1.25, \quad S_3 = \frac{49}{16} = 1.361, \quad S_4 = \frac{205}{144} = 1.424, \tag{8.18}$$

to be compared with $\pi^2/6 = 1.644934$. The first three Richardson extrapolants are much better:

$$S^{[1]}_{(1)} = \frac{3}{2} = 1.5, \quad S^{[1]}_{(2)} = \frac{19}{12} = 1.58, \quad S^{[1]}_{(3)} = \frac{29}{18} = 1.611. \tag{8.19}$$

Iteration of these results by inserting $S^{[1]}_{(N)}$ in (8.17) yields further improvement: $5/3 = 1.667$, but this iteration improves only slowly with N.

To do better we keep the first two terms in (8.15). This gives the second Richardson extrapolant,

$$S^{[2]}_{(N)} = \frac{1}{2}\left[(N+2)^2 S_{N+2} - 2(N+1)^2 S_{N+1} + N^2 S_N\right]. \tag{8.20}$$

When applied to the series (8.2) the first three terms in the series yields nearly 1% accuracy:

$$S^{[2]}_{(1)} = \frac{13}{8} = 1.625. \tag{8.21}$$

For further numerical details, see Table 8.4 on page 377 of Bender and Orzag.

8.2 Summing Divergent Series

The series encountered in physics, typically perturbation expansions, are usually divergent. How can one extract a meaningful number from such series, which represent physical processes and so reflect real processes?

On the surface, it would seem impossible to attach any meaning to such obviously divergent series as

$$1 + 1 + 1 + 1 + 1 + \dots, \tag{8.22a}$$

$$1 - 1 + 1 - 1 + 1 - \dots. \tag{8.22b}$$

However, as we will now see, perfectly finite numbers can be associated with these series. Again there are various procedures, of which we give a sampling. Throughout, we are considering a divergent series of the form

$$\sum_{n=0}^{\infty} a_n. \tag{8.23}$$

8.2.1 Euler Summation

Suppose

$$\sum_{n=0}^{\infty} a_n x^n = f(x) \tag{8.24}$$

converges if $|x| < 1$. Then we *define* the limit of the series (8.23) by

$$S = \lim_{x\to 1} f(x). \tag{8.25}$$

Thus, for the series (8.22b),

$$S = \sum_{n=0}^{\infty} (-1)^n, \tag{8.26}$$

$f(x)$ is

$$f(x) = \sum_{n=0}^{\infty} (-1)^n x^n = \frac{1}{1+x}, \tag{8.27}$$

so $S = 1/2$. To supply more credence to this result, we note that it is reproduced by the Shanks transformation. The partial sums of the series are

$$S_0 = 1, \quad S_1 = 0, \quad S_2 = 1, \quad S_3 = 0, \quad \dots, \tag{8.28}$$

so

$$S=\frac{S_{N+1}S_{N-1}-S_n^2}{S_{N+1}+S_{N-1}-2S_n}=\frac{1}{2} \tag{8.29}$$

for all N.

What if we apply Euler summation to the series

$$1+0-1+1+0-1+1+0-1+1+0-1+\ldots? \tag{8.30}$$

Now

$$\begin{aligned} f(x) &= 1-x^2+x^3-x^5+x^6-x^8+x^9-\ldots \\ &= \sum_{n=0}^{\infty} x^{3n}-x^2\sum_{n=0}^{\infty} x^{3n} \\ &= \frac{1-x^2}{1-x^3}=\frac{1+x}{1+x+x^2}, \end{aligned} \tag{8.31}$$

so the sum of (8.30) is

$$S=f(1)=\frac{2}{3}. \tag{8.32}$$

Thus the process of summation is not (infinitely) associative. In this case the Shanks transformation does not work.

8.2.2 Borel Summation

Now we use the Euler representation of the Gamma function, or the factorial,

$$n!=\int_0^\infty dt\, t^n e^{-t}. \tag{8.33}$$

Then we formally interchange summation and integration:

$$S=\sum_{n=0}^\infty a_n\frac{1}{n!}\int_0^\infty dt\, t^n e^{-t}=\int_0^\infty dt\, e^{-t}\sum_{n=0}^\infty \frac{1}{n!}a_n t^n, \tag{8.34}$$

which defines the sum if

$$g(t)=\sum_{n=0}^\infty \frac{1}{n!}a_n t^n \tag{8.35}$$

exists.

Thus for (8.22b),

$$g(t)=\sum_{n=0}^\infty (-1)^n\frac{t^n}{n!}=e^{-t}, \tag{8.36}$$

and so

$$S=\int_0^\infty dt\, e^{-2t}=\frac{1}{2}, \tag{8.37}$$

which coincides with the result found by Euler summation. In general, Borel summation is more powerful than Euler summation, but if both Euler and Borel sums exist, they are equal.

In fact, we can prove that any summation that is both

1. linear, meaning that if
$$\sum_{n=0}^{\infty} a_n = A, \quad \sum_{n=0}^{\infty} b_n = B, \tag{8.38a}$$
then
$$\sum_{n=0}^{\infty}(\alpha a_n + \beta b_n) = \alpha A + \beta B, \tag{8.38b}$$
and
2. satisfies
$$\sum_{n=0}^{\infty} a_n = a_0 + \sum_{n=1}^{\infty} a_n, \tag{8.39}$$

is unique. In fact, from these two properties alone (which are satisfied by both Euler and Borel summation) we can find the value of the sum. Thus for example,

$$1-1+1-1+1-1+\ldots = S = 1-(1-1+1-1+1-1+\ldots) = 1-S, \tag{8.40}$$

implies $S = 1/2$. Slightly more complicated is

$$\begin{aligned} S &= (1+0-1+1+0-1+1+0-1+\ldots) \\ &= 1+(0-1+1+0-1+1+0-1+\ldots) \\ &= 1+0+(-1+1+0-1+1+0-1+1+0-\ldots), \end{aligned} \tag{8.41}$$

where adding the three lines gives

$$3S = 2 + (0+0+0+0+0+\ldots) = 2, \tag{8.42}$$

or $S = 2/3$ as before.

But there are sums resistant to such schemes. An example is (8.22a), because the above process leads to

$$S = 1 + (1+1+1+\ldots) = 1 + S, \tag{8.43}$$

which is only satisfied by $S = \infty$. Yet such a series can be summed.

8.2.3 Zeta-function Summation

Recall that the zeta function is defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \mathrm{Re}\, s > 1. \tag{8.44}$$

In fact, $\zeta(s)$ exists for all $s \neq 1$, so we can use that function to define the sum almost everywhere in the complex s plane. In particular, for $s = 0$:

$$1 + 1 + 1 + 1 + \ldots = \zeta(0) = -\frac{1}{2}. \tag{8.45}$$

Even a more divergent sum can be evaluated this way:

$$\sum_{n=1}^{\infty} n = \zeta(-1) = -\frac{1}{12}. \tag{8.46}$$

Note the remarkable fact that these sums are not only finite, but negative, even though each term in the sum is positive!

8.2.4 Casimir Effect

Here we give a physical example of the utility of this last mode of summation. The physics is that of a pair of parallel metallic plates, separated by a distance a in the vacuum. Because the plates modify the properties of the vacuum, there is a change in the zero-point energy of the electromagnetic field, which feels the plates because they are conductors. The result is an attraction between the plates, the famous Casimir effect, predicted by Casimir in 1948 (the same year that Schwinger discovered how to renormalize quantum electrodynamics), and now verified by many experiments at the percent level. The zero-point energy (per unit area) of modes confined by the plane boundaries at $z = 0$ and $z = a$ is

$$E = \frac{1}{2}\sum \hbar\omega = \frac{\hbar c}{2}\sum_{n=1}^{\infty}\int \frac{d^2k}{(2\pi)^2}\sqrt{k^2 + \left(\frac{n\pi}{a}\right)}, \tag{8.47}$$

where in the mode sum we have integrated over the two transverse wavenumbers k_x and k_y, and summed over the discrete modes, which, say, must vanish at $z = 0$ and a, that is, be given by an (unnormalized) mode function

$$\phi(z) = \sin\frac{n\pi}{a}z. \tag{8.48}$$

Now we write the square root as integral, putting its argument in the exponential:

$$\sqrt{k^2 + \left(\frac{n\pi}{a}\right)^2} = \frac{1}{\Gamma\left(-\frac{1}{2}\right)}\int_0^{\infty}\frac{ds}{s}s^{-1/2}e^{-(k^2+(n\pi/a)^2)s}, \tag{8.49}$$

and then interchange the two integrals:

$$E = \frac{\hbar c}{2}\sum_{n=1}^{\infty}\int_0^{\infty}\frac{ds}{s^{3/2}}e^{-(n\pi/a)^2 s}\left(\int_{-\infty}^{\infty}\frac{dk}{2\pi}e^{-k^2 s}\right)^2\frac{1}{-2\sqrt{\pi}}. \tag{8.50}$$

Here we have recognized that the two-dimensional integral over $\mathbf{k} = (k_x, k_y)$ can be broken into the product of two one-dimensional integrals because

$$e^{-(k_x^2+k_y^2)s} = e^{-k_x^2 s}e^{-k_y^2 s}. \tag{8.51}$$

These one-dimensional integrals are simply Gaussians, so the squared factor in (8.50) is simply $1/(4\pi s)$. The remaining s-integral is again a gamma function:

$$\begin{aligned} E &= -\frac{\hbar c}{16\pi^{3/2}} \sum_{n=1}^{\infty} \int_0^\infty \frac{ds}{s^{5/2}} e^{-(n\pi/a)^2 s} \\ &= -\frac{\hbar c}{16\pi^{3/2}} \Gamma\left(-\frac{3}{2}\right) \sum_{n=0}^{\infty} \left(\frac{n\pi}{a}\right)^3 \\ &= -\frac{\hbar c \pi^2}{1440 a^3}, \end{aligned} \tag{8.52}$$

where we have used the facts that

$$\Gamma\left(-\frac{3}{2}\right) = \frac{4}{3}\sqrt{\pi}, \quad \zeta(-3) = \frac{1}{120}, \tag{8.53}$$

together with the zeta-function continuation embodied in Eq. (8.44) When multiplied by 2, for the two polarization states of the photon, this is exactly Casimir's result, which implies an attractive force per unit area between the plates,

$$P = -\frac{\partial}{\partial a} E = -\frac{\hbar c \pi^2}{240 a^4} = -1.30 \times 10^{-27} \mathrm{N\,m}^2/a^4. \tag{8.54}$$

8.3 Padé Approximants

Consider a partial Taylor sum,

$$T_{N+M}(z) = \sum_{n=0}^{N+M} a_n z^n, \tag{8.55}$$

which is an $N+M$th degree polynomial. Write this in a rational form,

$$P_M^N(z) = \frac{\sum_{n=0}^N A_n z^n}{\sum_{m=0}^M B_m z^m}, \tag{8.56}$$

which is called the $[N, M]$th Padé approximant. Here the coefficients are determined from the Taylor series coefficients as follows: We set $B_0 = 1$, and determine the $(N+M+1)$ coefficients $A_0, A_1, \ldots, A_N$ and $B_1, B_2, \ldots, B_M$ by requiring that when the rational function (8.56) be expanded in a Taylor series about $z = 0$ the first $N+M+1$ coefficients match those of the original Taylor expansion (8.55).

Example

Consider the exponential function

$$e^z = 1 + z + \frac{1}{2} z^2 + \ldots. \tag{8.57}$$

The $[1,1]$ Padé of this is of the form

$$P_1^1(z) = \frac{A_0 + A_1 z}{1 + B_1 z}, \tag{8.58}$$

which, when expanded in a series about $z = 0$ reads

$$P_1^1(z) \approx A_0 + (A_1 - B_1 A_0)z + (B_1^2 A_0 - A_1 B_1)z^2. \tag{8.59}$$

Matching this with Eq. (8.57), we obtain the equations

$$\begin{aligned} A_0 &= 1, && (8.60a) \\ A_1 - B_1 A_0 &= 1, && (8.60b) \\ B_1(B_1 A_0 - A_1) &= \frac{1}{2}, && (8.60c) \end{aligned}$$

so we learn immediately that

$$A_0 = 1, \quad B_1 = -\frac{1}{2}, \quad A_1 = \frac{1}{2}, \tag{8.61}$$

so the $[1,1]$ Padé is

$$P_1^1(z) = \frac{1 + \frac{1}{2}z}{1 - \frac{1}{2}z}. \tag{8.62}$$

How good is this? For example, at $z = 1$,

$$P_1^1(1) = 3, \tag{8.63}$$

which is 10% larger than the exact answer $e = 2.718281828\ldots$, and is not quite as good as the result obtained from the first three terms in the Taylor series,

$$1 + z + \frac{1}{2}z^2\bigg|_{z=1} = 2.5, \tag{8.64}$$

about 8% low. However, in higher orders, Padé approximants rapidly outstrip Taylor approximants. Table 8.1 compares the numerical accuracy of P_N^M with T_{N+M}.

Note that typically the Padé approximant, obtained from a partial Taylor sum, is more accurate than the latter. This comes at a price, however; the Padé, being a rational expression, has poles, which are not present in the original function. Thus, e^z is an entire function, while the $[1,1]$ Padé approximant of this function has a pole at $z = 2$.

Example

Here's another example:

$$\frac{1}{z}\log(1+z) = 1 - \frac{z}{2} + \frac{z^2}{3} - \frac{z^3}{4} + \frac{z^4}{5} - \frac{z^5}{6} + \ldots. \tag{8.65}$$

$T_{N+M}(1)$	$P_M^N(1)$	Relative error of Padé
$T_3(1) = 2.667$	$P_2^1(1) = 2.667$	-1.9%
$T_4(1) = 2.708$	$P_2^2(1) = 2.71429$	-0.15%
$T_5(1) = 2.717$	$P_3^2(1) = 2.71875$	$+0.017\%$
$T_6(1) = 2.71806$	$P_3^3(1) = 2.71831$	$+0.00103\%$
$T_7(1) = 2.71825$	$P_4^3(1) = 2.71827957$	-0.000083%

Table 8.1: Comparison of partial Taylor series with successive Padé approximants for the exponential function, evaluated at $z = 1$. Note that precisely the same data is incorporated in T_{N+M} and in P_M^N.

Approximant	$z = 0.5$	$z = 1$	$z = 2$
Exact	0.810930216	0.69314718	0.549306
P_3^3	0.810930365	0.69315245	0.549403
P_4^3	0.810930203	0.69314642	0.549285

Table 8.2: Padé approximations for the function $(1/z)\log(1+z)$ compared with the exact values. Note that the Taylor series for this function has a radius of convergence of unity, yet the Padé approximations converge rapidly even beyond the circle of convergence.

It is a simple algebraic task to expand the form of an $[N, M]$ Padé in a Taylor series and compute the Padé coefficients by matching with the above. This can, of course, be easily implemented in a symbolic program. For example, in Mathematica,

$$P_M^N(z) = PadeApproximant[f[z], \{z, 0, \{N, M\}\}]. \tag{8.66}$$

Doing so here yields

$$P_3^3(z) = \frac{1 + \frac{17}{14}z + \frac{1}{3}z^2 + \frac{1}{140}z^3}{1 + \frac{12}{7}z + \frac{6}{7}z^2 + \frac{4}{35}z^3}. \tag{8.67}$$

Table 8.2 shows representative numerical values for P_3^3 and P_4^3. The Padé approximants rapidly converge to the correct value even well beyond the circle of convergence of the original series. Note further in this example that

- P_N^N is larger than the function, and decreases monotonically toward it, and

- P^N_{N+1} is smaller than the function, and increases monotonically toward it.

This bounding behavior is typical of a class of functions. For more detail see C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978), pp. 383ff.

Field Theory Examples

The following function occurs in the field theory of a massless particle in zero dimensions,

$$\begin{aligned} Z(\delta) &= \int_{-\infty}^{\infty} \frac{dx}{\sqrt{\pi}} e^{-(x^2)^{1+\delta}} \\ &= \frac{2}{\sqrt{\pi}} \frac{1}{(2+2\delta)} \int_0^\infty \frac{dt}{t} t^{1/(2+2\delta)} e^{-t} = \frac{2}{\sqrt{\pi}} \frac{1}{(2+2\delta)} \Gamma\left(\frac{1}{2+2\delta}\right) \\ &= \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{3+2\delta}{2+2\delta}\right), \end{aligned} \tag{8.68}$$

where the gamma function was defined by Euler as

$$\Gamma(z) = \int_0^\infty \frac{dt}{t} t^z \, e^{-t}, \tag{8.69}$$

and satisfies the identity

$$\Gamma(z+1) = z\Gamma(z). \tag{8.70}$$

The gamma function generalizes the factorial to complex values:

$$\Gamma(n+1) = n!, \quad n = 0, 1, 2, \ldots. \tag{8.71}$$

Because the gamma function $\Gamma(z)$ has poles when $z = -N$, $N = 0, 1, 2, \ldots$, this function has an infinite number of singularities between $\delta = -3/2$ and $\delta = -1$. Thus the radius of convergence of the Taylor series about $\delta = 0$ is 1. Yet low order Padé's for $E(\delta) = -\log Z(\delta)$ give an excellent approximation well outside of this radius, as Table 8.3 shows.

The "partition function" for a zero-dimensional field theory with a mass μ is given by the function

$$Z(\delta) = \mu \sqrt{\frac{2}{\pi}} \int_0^\infty dx \, e^{-\frac{\mu^2}{2} x^2 - \lambda (x^2)^{1+\delta}}. \tag{8.72}$$

We consider two cases. If $\mu^2 > 0$, the power series in δ again has radius of convergence 1, but the Padé approximants are accurate far beyond this radius, as shown in Table 8.4.

If, on the other hand $\mu^2 < 0$ (which corresponds to the "Higgs mechanism" in particle physics), the Taylor series converges nowhere, yet the Padé approximant is still quite good, as seen in Table 8.5.

δ	$T_{10}(\delta)$	$T_{20}(\delta)$	$P_2^3(\delta)$	$P_4^5(\delta)$	$E(\delta)$
-2.0	-1266.97	-2.0×10^6	-0.651267	-0.692962	-0.693147
-0.5	-0.120055	-0.120781	-0.120831	-0.12078223848	-0.12078223764
0.5	-0.00781712	-0.00759091	-0.00759097	-0.0075905958951	-0.0075905958949
1.0	-0.367098	-0.516940	-0.0225167	-0.022510401233	-0.022510401213
2.0	-465.821	-688611	-0.0458145	-0.04575620415	-0.04575620349
5.0	-5.5×10^6	-7.8×10^{13}	-0.0786672	-0.078172915	-0.078172899

Table 8.3: Approximations to the function (8.68). What is approximated is $E(\delta) = -\log Z(\delta)$. The Padé approximants based on 6 and 10 terms in the Taylor series of this function are far more accurate that the 10 and 20 term truncated Taylor series, and even are remarkably accurate far outside the circle of convergence, where the Taylor series is meaningless.

δ	$T_8(\delta)$	$P_4^4(\delta)$	$Z(\delta)$
0.5	1.04631	1.04630	1.04630
1.0	1.07719	1.07436	1.07436
2.0	1.81047	1.10647	1.10649
5.0	745.176	1.14253	1.14285

Table 8.4: Comparison of $Z(\delta)$, Eq. (8.72), $\mu^2 > 0$, with the 8-term truncated power series, and the corresponding $[4, 4]$ Padé. Here we have taken $\mu^2 = 1$, $\lambda = 1$.

δ	$T_8(\delta)$	$P_4^4(\delta)$	$Z(\delta)$
0.1	0.94808	0.94790	0.94790
0.5	137.697	0.88388	0.88381
1.0	40109.3	0.87323	0.87253
2.0	1.1×10^7	0.88334	0.87974
5.0	1.8×10^{10}	0.91830	0.90517

Table 8.5: Comparison of $Z(\delta)$, Eq. (8.72), $\mu^2 < 0$, with the 8-term truncated power series, and the corresponding $[4, 4]$ Padé. Here we have taken $\mu^2 = -1$, $\lambda = 1$.

8.4 Continued Fractions

8.4.1 Number Theory

The most familiar way of representing real numbers is in terms of a decimal fraction, which is nonterminating and nonrepeating if the number is irrational. However, there are other representations which, if less familiar, can be very useful. For example, the base of the natural logarithms e can be written in the form of a continued fraction,

$$e = 2 + \cfrac{1}{1 + \cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{4+\dots}}}}}. \tag{8.73a}$$

Because this built-up form is cumbersome to write, we could write this as

$$e = 2+1/(1+1/(2+1/(1+1/(1+1/(4+1/(1+1/(1+1/(6+1/(1+1/(1+\dots, \tag{8.73b}$$

or even more compactly as

$$e = 2 + \frac{1}{1+}\frac{1}{2+}\frac{1}{1+}\frac{1}{1+}\frac{1}{4+}\frac{1}{1+}\frac{1}{1+}\frac{1}{6+\dots}. \tag{8.73c}$$

The form seen here is the representation of a real number x in the form

$$x = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\frac{1}{a_3+}\frac{1}{a_4+\dots}, \tag{8.74}$$

where the numbers a_n are integers called *partial quotients*. The rational number formed by including only the first $n+1$ partial quotients a_0, a_1, $\dots$, a_n is called the n *convergent* of x. So the continued fraction is given by the set of a_ns:

$$e = \{2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, 1, 1, 12, \dots\}, \tag{8.75}$$

and the successive convergents, which rapidly approach $e = 2.718281828\dots$, are

$$\begin{aligned} &\left\{2, 3, \frac{8}{3}, \frac{11}{4}, \frac{19}{7}, \frac{87}{32}, \frac{106}{39}, \frac{193}{71}, \frac{1264}{465}, \frac{1457}{536}, \frac{2721}{1001}, \frac{23225}{8544}, \frac{25946}{9545}, \frac{49171}{18089}, \dots\right\} \\ = &\ \{2, 3, 2.666666667, 2.750000000, 2.714285714, 2.718750000, 2.717948718, \\ &\quad 2.718309859.2.718279570, 2.718283582, 2.718281718, 2.718281835, \\ &\quad 2.718281823, 2.718281829, 2.718281828, \dots\}. \end{aligned} \tag{8.76}$$

The partial quotients of x are determined by successively determining the unique integer that provides a bound for x for a given truncation of the partial fraction. Thus in the above example, where in each case $0 < r < 1$,

$$2 < e, \tag{8.77a}$$

$$\frac{5}{2} < 2 + \frac{1}{1+r} < 3, \tag{8.77b}$$

$$\frac{8}{3} < 2 + \frac{1}{1+}\frac{1}{2+r} < \frac{11}{4}, \tag{8.77c}$$
$$\frac{19}{7} < 2 + \frac{1}{1+}\frac{1}{2+}\frac{1}{1+r} < \frac{11}{4}, \tag{8.77d}$$
$$\frac{19}{7} < 2 + \frac{1}{1+}\frac{1}{2+}\frac{1}{1+}\frac{1}{1+r} < \frac{30}{11}, \tag{8.77e}$$

and so on. The successive convergents are the upper and lower bounds corresponding to $r = 0$.

The partial fraction representation of real numbers can be generated using your favorite symbolic program. For example, in Mathematica the first n partial quotients of x are given by

$$ContinuedFraction[x, n], \tag{8.78}$$

and the first n convergents are given by

$$Convergents[x, n]. \tag{8.79}$$

Let us conclude this subsection with the following comments.

- Evidently, a rational number is represented by a terminating continued fraction. For example,

  $$\frac{12357}{1234567890} = \{0, 99908, 2, 1, 1, 1, 1, 3, 3, 2, 1, 2\} \tag{8.80}$$

  exactly.

- An algebraic number, that is one which is a solution of an algebraic equation, which is not rational, is represented by a repeating pattern of partial quotients. For example,

  $$\sqrt{137} = \{11, 1, 2, 2, 1, 1, 2, 2, 1, 22, 1, 2, 2, 1, 1, 2, 2, 1, 22, \ldots\}. \tag{8.81}$$

- A trancendental number is represented by a nonrepeating pattern. That pattern is simple in the case of e, but not so for the case of π:

  $$\begin{aligned}\pi = \{&3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, 2, 1, 1, 2, 2, 2, 2, 1, 84, 2, 1, 1, 15,\\ &3, 13, 1, 4, 2, 6, 6, 99, 1, 2, 2, 6, 3, 5, 1, 1, 6, 8, 1, 7, 1, 2, 3, 7, \ldots\}\end{aligned} \tag{8.82}$$

  The first few convergents are

  $$\pi = \{3, 3.14285714, 3.141509434, 3.141592920, 3.141592653, 3.141592654, \ldots\}, \tag{8.83}$$

  so ten figure accuracy requires 6 terms. However, there are other continued fraction representations for π that have simple patterns:

  $$\frac{4}{1+}\frac{1^2}{2+}\frac{3^2}{2+}\frac{5^2}{2+}\frac{7^2}{2+\ldots}, \tag{8.84}$$

which has rather poor partial sums:

$$\pi = \{4, 2.6667, 3.4667, 2.89524, 3.33968, \ldots\}; \tag{8.85}$$

$$\pi = 3 + \frac{1}{6+}\frac{3^2}{6+}\frac{5^2}{6+}\frac{7^2}{6+\ldots}, \tag{8.86}$$

where the convergents are somewhat better

$$\pi = \{3, 3.16667, 3.13333, 3.14524, 3.13968, 3.14271, \ldots\}; \tag{8.87}$$

$$\pi = \frac{4}{1+}\frac{1^2}{3+}\frac{2^2}{5+}\frac{3^2}{7+\ldots}, \tag{8.88}$$

which is comparable,

$$\pi = \{4, 3, 3.16667, 3.13725, 3.14234, \ldots\}. \tag{8.89}$$

All of these are much worse than the rapid convergence of the standard convergents. But it is the existence of simple patterns that is perhaps remarkable.

8.4.2 Continued Fraction Representation of Functions

If a function is represented by a power series about the origin,

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \tag{8.90}$$

we can also write it in a continued-fraction form. The standard approach here is to write

$$\begin{aligned} f(x) &= \cfrac{b_0}{1 + \cfrac{b_1 x}{1 + \cfrac{b_2 x}{1 + \cfrac{b_3 x}{1 + \cfrac{b_4 x}{1+\ldots}}}}} \\ &= \frac{b_0}{1+}\frac{b_1 x}{1+}\frac{b_2 x}{1+}\frac{b_3 x}{1+}\frac{b_4 x}{1+\ldots}. \end{aligned} \tag{8.91}$$

Evidently, there is a one-to-one correspondance between the Taylor-series coefficients $\{a_n\}$ and the continued-fraction coefficients $\{b_n\}$, which may be determined by expanding the continued fraction in a power series for small x. The theory of such a representation is discussed also in the book by Bender and Orzag.

Let us consider a function with the property $f(0) = 1$; this is merely a convenient choice of normalization. Then the relation between the continued fraction coefficients and the series coefficients is easily found to be

$$a_0 = b_0 = 1, \tag{8.92a}$$

$$a_1 = -b_1, \tag{8.92b}$$

$$a_2 = b_1(b_1 + b_2), \tag{8.92c}$$

$$a_3 = -b_1[b_2 b_3 + (b_1 + b_2)^2], \tag{8.92d}$$

$$a_4 = b_1[b_2 b_3(b_3 + b_4) + 2(b_1 + b_2)b_2 b_3 + (b_1 + b_2)^3], \tag{8.92e}$$

and so on. This constitutes a nonlinear mapping from the set of numbers $\{b_n\}$ to the set $\{a_n\}$ or vice versa.

This mapping seems to be quite remarkable in that the sequence of b_ns is typically much simpler than the sequence of a_ns. Here are some examples:

Example 1

Let $b_n = n$, that is, $b_1 = 1$, $b_2 = 2$, etc. Then by computing the first few a_ns from the above formulæ we find

$$b_n = n \quad \Rightarrow |a_n| = (2n-1)!!. \tag{8.93}$$

Example 2

Let the continued-fraction sequence be $\{b_n\} = \{1, 1, 2, 2, 3, 3, 4, 4, \ldots\}$. Then the power series coefficients are given by the factorial,

$$|a_n| = n!. \tag{8.94}$$

Example 3

What if $b_n = n^2$? The first few a_n are

$$a_1 = -1, \tag{8.95a}$$
$$a_2 = 5, \tag{8.95b}$$
$$a_3 = -61, \tag{8.95c}$$
$$a_4 = 1385. \tag{8.95d}$$

These are recognized as the first few Euler numbers, defined by the generating function

$$\frac{1}{\cosh t} = \sum_{n=0}^{\infty} E_n \frac{t^n}{n!}, \tag{8.96}$$

namely

$$E_0 = 1, \tag{8.97a}$$
$$E_2 = -1, \tag{8.97b}$$
$$E_4 = 5, \tag{8.97c}$$
$$E_6 = -61, \tag{8.97d}$$
$$E_4 = 1385, \tag{8.97e}$$

and we conclude

$$a_n = E_{2n}. \tag{8.98}$$

Example 4

This suggests that we ask what sequence of b_ns corresponds to the Bernoulli numbers. It takes a bit of playing around to find the correct normalization, which matters since the transformation is nonlinear. If we take

$$a_n = 6B_{2n+2}, \tag{8.99}$$

we find that the corresponding continued-fraction coefficients are given by

$$b_n = \frac{n(n+1)^2(n+2)}{4(2n+1)(2n+3)}. \tag{8.100}$$

Although the latter seems a bit complicated, it is a closed algebraic expression. It further grows with n as a low power. Neither of these features hold for the Bernoulli numbers, which grow more rapidly than exponentially, and have no closed-form representation.

These ideas are provocative, yet the general significance of these results remains elusive. There appears also to be some deep connection to field theory. See C. M. Bender and K. A. Milton, *J. Math. Phys.* **35**, 364 (1994) for more details.