Comparison between Borel-Padé summation and factorial series, as time integration methods

Ahmad Deeb, Aziz Hamdouni, Dina Razafindralandy

Comparison between Borel-Padé summation and factorial series, as time integration methods

Ahmad Deeb, Aziz Hamdouni and Dina Razafindralandy

Laboratoire des Sciences de l'Ingnieur pour l'Environnement
LaSIE, UMR - 7356 CNRS
University of La Rochelle, France

Abstract

We compare the performance of two algorithms of computing the Borel sum of a time power series. The first one uses Padé approximants in Borel space, followed by a Laplace transform. The second is based on factorial series. These algorithms are incorporated in a numerical scheme for time integration of differential equations.

1 Introduction

Often, the numerical resolution of an evolution problem is very time-consuming. It is, for example, the case of the simulation of turbulent fluid flows. The main reason is that the time step Δt required by classical schemes (Euler, Runge-Kutta, ...) is subject to a restrictive stability or accuracy condition. As a consequence, Δt is generally very small (at the order of milliseconds for fluid flow and femtoseconds for particle transport), while a long-time simulation (some minutes to many hours) is necessary for engineering applications.

In [38], Razafindralandy and Hamdouni proposed an algorithm which enables bigger time steps for the resolution of differential equations. This algorithm is based on the Asymptotic Numerical Method, where the solution is sought as a (eventually zero radius of convergence) power series of t, followed by a Borel summation to enlarge the domain of validity of the series and, consequently, to speed up the resolution. For the heat equation, the time step was up to about 1000 times as large as the one required by the usual Euler explicit method. With a simple periodic problem, it was also shown that the Borel-summation-based algorithm is much faster, in terms of CPU time, than the Euler explicit and the 2-nd and 4-th order Runge-Kutta methods.

Beyond speed considerations, it was shown in [10] that, coupled with the Borel summation method, the asymptotic numerical method provides a well-behaved numerical scheme regarding to invariance properties. Indeed, it competes with symplectic schemes in solving hamiltonian systems since it respects,

1

at a high order, Liouville's theorem of volume conservation. Contrarily to Runge-Kutta or Euler explicit method, it is, moreover, able to reproduce the iso-spectrality of the solution of a Lax pair problem.

Borel summation of a divergent but Gevrey series $\hat{u}$ consists in (a) computing the Borel transform $\mathcal{B}\hat{u}$ of $\hat{u}$ which is a convergent series, (b) prolonging $\mathcal{B}\hat{u}$ along in a domain containing a semi-line from the origin, and (c) taking the Laplace transform to go back to the physical space [36, 41]. At the end, one gets the so-called Borel-sum of $\hat{u}$ which is an holomorphic function (in some domain) and is an actual solution of the equation. Note that formal time series solutions of many equations arising in fluid mechanics (including the heat equation, Burgers'equation and the Navier-Stokes equations) are generally divergent but Gevrey series [26, 28, 27, 8].

This summation technique was transformed in [38, 10] into a numerical algorithm where the prolongation of the Borel-transformed series was realised with Padé approximants and the Laplace transformation with a Gauss-Laguerre quadrature. The strength but also the weakness of this three-stage algorithm are the Padé approximation. Indeed, on one hand, the latter are known to be an efficient tool in prolonging a convergent series outside its validity disk. They may also be usefull in finding poles and then in choosing the direction of the Laplace transform. On the other hand, the standard algorithm of Padé approximation requires the resolution of a linear problem which may not be inversible when the Padé table is not normal; a reduction is then needed. Moreover this linear problem is generally ill-conditionned and may lead to the apparition of spurious zero-pole pairs (Froissart doublet) [17, 40].

To avoid the problem induced by the Padé approximants, the use of Padé-type approximants where the denominator (and the poles) are prescribed is proposed in [42]. Another method is to use a conformal mapping to extend the Borel-transformed series [21], but this method needs the knowledge of the closest pole of $\mathcal{B}\hat{u}$. All these methods provide an (Laplace-) integral representation of the Borel sum at the end of the algorithm.

The Borel sum can also be represented by a generalised factorial series (GFS) which is computed directly from the terms of the series, without passing through the Borel space. The advantage is that it avoids the three-stage computation, reducing then the source of errors. The introduction of GFS as an asymptotics to a power series dates from the beginning of the last century [35, 31, 32, 47]. More recently, the effective summation by GFS has been analysed numerically in [42, 43, 11, 48] and applied to some generic (theoretical) problems where the formal series solution $\hat{u}$ is divergent.

In this article, we propose to compare the Borel-Padé algorithm to one based on generalised factorial series for the numerical resolution of ODE's. The goal is to examine robustness of these algorithm for a "blind" use.

In section 2, the Gevrey asymptotics theory and the Borel summation method are briefly recalled. The Borel-Padé algorithm is then illustrated. Its weak points are analysed in section 3 and an algorithm based on generalised factorial series is presented in section 4. The two last sections are devoted to numerical

experiments.

2 Borel summation theory

Consider a meromorphic ordinary or partial differential equation

$$\frac{\mathrm{d}u}{\mathrm{d}t} = F(u,t), \qquad u(t=0) = u_0, \tag{1}$$

with a formal time power series solution

$$\hat{u}(t) = \sum_{n=0}^{\infty} u_n t^n. \tag{2}$$

The coefficients u_n can be calculated by injecting (2) in (1).

Series (2) may be divergent (in the sense that the radius of convergence is zero). However, we assume that $\hat{u}(t)$ is a Gevrey series of index one, meaning that its general term does not grow faster than factorials. More generally,

Definition 2.1 *A formal power series $\hat{u}(t) = \sum_{n=0} u_n t^n$ is called a Gevrey series of index $I > 0$ if there exists constants $C, A \in \mathbb{R}$ such that*

$$|u_n| < CA^n (n!)^I \qquad \forall n \geq 1. \tag{3}$$

Most of series coming from engineering applications are divergent [19] but Gevrey series of some order [37]. Knowing the terms of $\hat{u}(t)$, our goal is to find an approximate solution of (1). This approximate solution will be an analytic function in a sector and owns series (2) as Gevrey asymptotics, as defined hereafter.

Definition 2.2 *A function $u(t)$, analytic in an open sector S, is Gevrey-asymptotic of order I to a power series $\hat{u}(t) = \sum_{n \geq 0} u_n t^n$ at the origin in S if, for any compact subsector $T \subset S$, there exists two constants C and A such that*

$$\forall n \geq 1, \ \forall t \in T, \qquad \left| u(t) - \sum_{k=0}^{n-1} u_k t^k \right| < CA^n |t|^n (n!)^I . \tag{4}$$

As said, we limit ourselves to index-one Gevrey series ($I = 1$) in this article. We denote $u(t) \sim^G \hat{u}(t)$ when $\hat{u}(t)$ is Gevrey-asymptotic to $\hat{u}(t)$.

The general theory of Gevrey asymptotics [36] ensures that, given a Gevrey series, one can find a function wich has that series as Gevrey expansion. The Borel summation technique is a method of effectively computing this function.

2.1 Borel summation

As introduced, the Borel summation uses the Borel transformation and Laplace transformation. Let us define these operations.

Definition 2.3 *The Borel transform of* $\hat{u}(t) = \sum\limits_{n=0} u_n t^n$ *is the series*

$$\mathcal{B}\hat{u}(\xi) = \sum_{n=0}^{\infty} \frac{u_{n+1}}{n!} \xi^n \tag{5}$$

If $\hat{u}(t)$ is a Gevrey series, its factorial growth of u_n is then canceled out through the Borel transformation, such that $\mathcal{B}\hat{u}(\xi)$ has a non-zero convergence radius.

Let d be a semi-line d starting from the origin. We say that a function $P(\xi)$, analytic in a domain containing d, has an exponential decay at the infinity in the direction d if there exists two constants A and C such that

$$|P(\xi)| A\,e^{C|\xi|} \tag{6}$$

as $|\xi|$ tends to infinity along d. We can now set the definition of the Laplace transformation, which is, when applied formally to a power series, the inverse of the Borel transformation.

Definition 2.4 *Consider a function $P(\xi)$, analytic in a domain containing d, and having an exponential decay at the infinity in the direction d. The Laplace transform[1] of $P(\xi)$ in the direction d is*

$$\mathcal{L}P(t) = \int_d P(\xi)\,e^{-\xi/t}\,\mathrm{d}\,\xi. \tag{7}$$

$\mathcal{L}P(t)$ is analytic in a sector bissected by d.

We are now ready to set the definition of the Borel sum of a series.

Definition 2.5 *A power series $\hat{u}(t)$ is said* Borel-summable *in a direction d if*

- *$\hat{u}(t)$ is a Gevrey series,*

- *$\mathcal{B}\hat{u}(\xi)$ can be analytically prolonged into a function $P(\xi)$ in a domain containing d, and*

- *$P(\xi)$ has an exponential decay at the infinity in the direction d.*

In that case,

$$\mathcal{S}\hat{u}(t) = u_0 + \mathcal{L}P(t) \tag{8}$$

is called the Borel sum *of $\hat{u}(t)$ in the direction d.*

Theorem 2.6 *If $\hat{u}$ is Borel-summable, then it is the Gevrey expansion of its Borel sum $\mathcal{S}\hat{u}(t)$, i.e. $\hat{u}(t) \sim^G \mathcal{S}\hat{u}(t)$.*

$$\hat{u}(t) = \sum_{n=0}^{\infty} u_n t^n \qquad \sim^G \qquad \mathcal{S}\hat{u}(t) = u_0 + \int_d P(\xi) e^{-\xi/t}\, \mathrm{d}\,\xi$$

Borel $\downarrow$ $\qquad\qquad$ $\uparrow$ Laplace

$$\mathcal{B}\hat{u}(\xi) = \sum_{n=0}^{\infty} \frac{u_{n+1}}{n!}\,\xi^n \qquad \xrightarrow{\text{Prolongation}} \qquad P(\xi)$$

Table 1: Borel-Laplace summation

The three operations involved in the Borel-Laplace summation method are summarised in Table 1.

The dependence of the Borel sum on the direction d is related to Stokes phenomenon [22]. Since for nonlinear problems, there lacks theory which would determine which direction would be the most suitable, we consider only the case where d is the positive real axis.

Numerically, the series $\hat{u}(t)$ is determined up to an arbitrary order N. More precisely, $\hat{u}(t)$ is numerically represented by the truncated series

$$\hat{u}^N(t) = \sum_{n=0}^{N} u_n t^n. \tag{9}$$

To approximate the Borel sum $\mathcal{S}\hat{u}(t)$, one of the most natural algorithms is the Borel-Padé algorithm.

3 Analysis of the Borel-Padé algorithm

3.1 Borel-Padé algorithm (BP)

An algorithm which consists in transforming step-by-step the Borel-Laplace summation method into an effective numerical program is the Borel-Padé algorithm, described as follows:

(a) Compute the approximate Borel transform $\mathcal{B}\hat{u}^N(\xi)$,

(b) Extend $\mathcal{B}\hat{u}^N(\xi)$ on the positive real axis into a rational function $P^N(\xi)$ using Padé approximants, and

(c) Compute the Laplace transform using the Gauss-Laguerre quadrature method.

---
[1]Properly speaking, $\mathcal{L}P(t)$ is the Laplace transform of $P(\xi)$ at $1/t$.

The algorithm is schematised in Table 2. The function P'^N in Table 2 is linked to P_N by the Gauss quadrature formula, ξ_i are the roots of the N_G-th Laguerre polynomial and ω_i are the corresponding weights.

$$\hat{u}^N(t) = \sum_{n=0}^{N} u_n t^n \qquad\qquad u_0 + \sum_{i=1}^{N_G} P'^N(\xi_i, t)\omega_i$$

Borel $\Big\downarrow$ $\qquad\qquad$ $\Big\uparrow$ Gauss-Laguerre

$$\mathcal{B}\hat{u}^N(\xi) = \sum_{n=0}^{N-1} \frac{u_{n+1}}{n!}\,\xi^n \quad \xrightarrow{\text{Padé}} \quad P^N(\xi) = \frac{a_0 + a_1\xi + \cdots + a_{N_1}\xi^{N_1}}{1 + b_1\xi + \cdots + b_{N_2}\xi^{N_2}}$$

Table 2: Borel-Padé summation algorithm (BP)

The Padé approximant P^N is determined by the relation [15, 5, 6]:

$$1 + b_1\xi + \cdots + b_{N_2}\xi^{N_2}\mathcal{B}\hat{u}^N(\xi) - a_0 + \cdots + a_{N_1}\xi^{N_1} = O(\xi^m) \qquad (10)$$

where the degrees of the polynomials verify $N_1 + N_2 = N - 1$ and m is the highest order that can be obtained. Any choice of N_1 from 0 to N may give a suitable approximation of $P(\xi)$ but, when nothing is known on this function, the most popular choice is the diagonal Padé approximant. This choice is supported by some convergence theorems and conjectures from experiences (see [2, 39, 24]). So, in the numerical tests, we take $N_1 = N/2$ or $N_1 = (N+1)/2$ according to the parity of N.

The Gauss-Laguerre method has been chosen for the evaluation of the Laplace transform since it is one of the most popular quadrature methods of semi-infinite integrals. It integrates exactly a polynomial of degree up to $2N_G - 1$. Other methods may be considered (but this is beyond the scope of this article). One can cite for example the Fourier-Chebyshev scheme [4] wich extends the Clenshaw-Curtis method and Gauss-type quadratures which integrate exactly rational functions [23, 14, 12, 13, 46].

The principal weaknesses and source of errors in the Borel-Padé algorithm are the integration error, the non-unicity of the Padé approximants (even when the degrees N_1 and N_2 are fixed) and the apparition of poles of the Padé approximants. We will analyse these points one after the other.

3.2   Error due to the numerical integration

Consider the fundamental example of Euler's equation:

$$t^2\frac{\mathrm{d}u}{\mathrm{d}t} + u = t. \qquad (11)$$

The formal solution and its Borel transform are, respectivelly,

$$\hat{u}(t) = \sum_{n \geq 0} (-1)^n n! t^{n+1} \tag{12}$$

and

$$\mathcal{B}\hat{u}(\xi) = \sum_{n \geq 0} (-1)^n \xi^n. \tag{13}$$

$\mathcal{B}\hat{u}(\xi)$ can be naturally extended into the analytic function

$$P(\xi) = \frac{1}{1 + \xi}. \tag{14}$$

Hence, the Borel sum of $\hat{u}(t)$ is

$$\mathcal{S}\hat{u}(t) = \int_0^{+\infty} \frac{1}{1 + \xi}\, e^{-\xi/t} \mathrm{d}\xi. \tag{15}$$

In summary, from the divergent series (12), the summation method provides the solution (15) of equation (11) which is analytic in the complex half-plane where the real part of t is positive.

Borel-Padé algorithm is able to represent correctly $\mathcal{B}\hat{u}(\xi)$ and $P(\xi)$. Indeed, as soon as $N \geq 2$,

$$P^N(\xi) = P(\xi).$$

The only source of error is the numerical integration. Figure 1 compares the approximate solution to the analyic one (15) for $N = 10$. It shows the superiority of the BP solution over the truncated series approximation. It also shows the effect of the numerical integration when time grows.
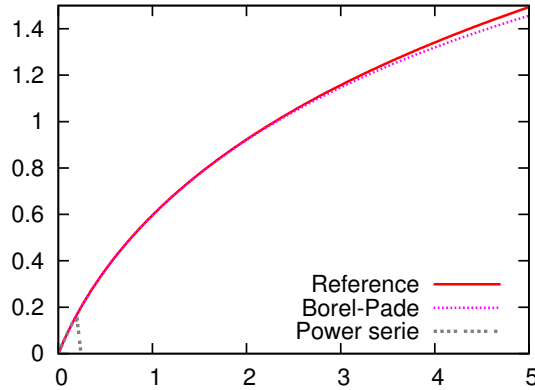

Figure 1: Effect of discretisation on Borel-Padé algorithm

The integration error is inherent to the method and cannot be avoided. It however should not be considered as a real weakness of the method since it is simply a discretisation error.

3.3 Pole of the Padé approximants on $\mathbb{R}^+$

A more problematic situation is when $\hat{u}(t)$ is Borel-summable in the positive real line direction but the Padé representation of $P(\xi)$ has a pole on this line. For example, consider the equation

$$\begin{cases} t\dfrac{\mathrm{d}u}{\mathrm{d}t} = u^2 + u, \\[3mm] \dfrac{\mathrm{d}u}{\mathrm{d}t}(0) = 1. \end{cases} \tag{16}$$

The formal solution and its Borel transform are respectively

$$\hat{u}(t) = {}_{n1}t^n \qquad \text{and} \qquad \mathcal{B}\hat{u}(\xi) = {}_{n0}\frac{\xi^n}{n!}. \tag{17}$$

The natural prolongation

$$P(\xi) = \mathrm{e}^\xi \tag{18}$$

of $\mathcal{B}\hat{u}(\xi)$ owns a Laplace transform along $d = \mathbb{R}^+$ but, for instance, when $N_1 = N_2 = 3$,

$$P^N(\xi) = \frac{1 + \dfrac{1}{2}\xi + \dfrac{1}{10}\xi^2 + \dfrac{1}{120}\xi^3}{1 - \dfrac{1}{2}\xi + \dfrac{1}{10}\xi^2 - \dfrac{1}{120}\xi^3} \tag{19}$$

has a pole on d.

Poles on the positive real line may be ignored during the computation of the Laplace transform (provided, of course, that they do not coincide with any Gauss node). But, as said in [9, 25], this may lead to a loss of precision or to a slow down.

Spurious poles may also appear due to numerical round-off [17, 16, 40], even when robust algorithms are used [29, 3]. Fortunately, they are paired with a close-by zero such that the doublet zero-pole simplifies outside of a small region, and can be ignored.

3.4 Non-unicity of Padé approximants

A classical problem arising with Padé approximation is the non-uniqueness of the representation. To remind this situation, consider the equation

$$\begin{cases} t^2\dfrac{\mathrm{d}u}{\mathrm{d}t} + u = t - t^2, \\[3mm] u(0) = 0. \end{cases} \tag{20}$$

It is straight forward to check that the formal solution

$$\hat{u}(t) = t + 2\sum_{n\geq 1}(-1)^n n! t^{n+1} \tag{21}$$

is Borel summable along $d = \mathbb{R}^+$. The Borel transform and its natural prolongation are

$$\mathcal{B}\hat{u}(\xi) = 1 + 2\sum_{n\geq 1}(-1)^n \xi^n,$$

and

$$P(\xi) = \frac{1-\xi}{1+\xi}.$$

However, definition (10) leads to a non invertible system when, for instance, $N_1 = N_2 = 2$, and the Padé approximant cannot be computed. This is due to the existence of a common divisor in the P^N. Technics exist for circumventing this problem and choosing a unique representation but at some extra computational cost [1, 18, 7].

In the next section, the generalised factorial series algorithm, which avoid the computation of a Padé approximant, is presented.

4   Generalised factorial series algorithm (GFS)

To show the link between factorial series and the Borel sum, let us formally rewrite the Laplace transform in the following way:

$$\int_0^{+\infty} P(\xi)\,\mathrm{e}^{-\xi/t}\,\mathrm{d}\xi \quad = \quad \int_0^1 P(-\ln)^{\frac{1}{t}-1}\,\mathrm{d} \tag{22}$$

$$\tag{23}$$

$$= \quad \sum_{n\geq 0}^{\infty} b_n \int_0^1 (1-)^n{}^{\frac{1}{t}-1}\,\mathrm{d} \tag{24}$$

where the b_n come from the Taylor development of $P(-\ln s)$ at $s = 1$. After integrations by parts, we get the inverse factorial representation

$$\mathcal{S}\hat{u}(t) = u_0 + \sum_{n\geq 0}^{\infty} \frac{b_n n!}{\frac{1}{t}(\frac{1}{t}+1)\cdots(\frac{1}{t}+n)}, \tag{25}$$

or

$$\mathcal{S}\hat{u}(t) = u_0 + \sum_{n\geq 0}^{\infty} \frac{b_n n! t^{n+1}}{(1+t)\cdots(1+nt)} =: \mathcal{I}(t). \tag{26}$$

9

The inverse factorial series $\mathcal{I}(t)$ is unique, is absolutely convergent for $Re(1/t)$ big enough [30, 31, 33, 34, 45] and converges effectively to $\mathcal{S}\hat{u}(t)$.

Denote $\mathcal{I}^N(t)$ be the truncation of the factorial series (26) at order N. As shown in [11, 26], $\mathcal{I}^N(t)$ approximates asymptotically the Borel sum as follows.

Theorem 4.1 *Consider a Borel-summable series $\hat{u}(t)$. Let A and C be the constants involved in the exponential decay of the prolongation $P(\xi)$ of its Borel transform (see equation 6) and $T = 1/C$. Then the inverse factorial series representation $\mathcal{I}(t)$ of $\mathcal{S}\hat{u}(t)$ verifies*

$$|\mathcal{S}\hat{u}(t) - \mathcal{I}^N(t)| \leq C_N \left| \frac{t^{N+1}N!}{(1+t)\cdots(1+tN)(Re\frac{1}{t} - \frac{1}{T})} \right| \tag{27}$$

where

$$C_N = AT^{\frac{1}{T}} \frac{(N+1+\frac{1}{T})^{N+1+\frac{1}{T}}}{(N+1)^N}. \tag{28}$$

With a suitable rotation, the sum in any other direction d can be recovered from (26).

In fact, $\mathcal{S}\hat{u}(t)$ can be represented by a more general convergent series, namely by a generalised factorial series (GFS) [44, 41]:

$$\mathcal{S}\hat{u}(t) = u_0 + \sum_{n \geq 0}^{\infty} \frac{b_n s_1 \ldots s_n t^{n+1}}{(1+ts_1)\cdots(1+ts_n)}. \tag{29}$$

for any sequence of complex numbers (s_n) located on the semi-line d' which is the mirror symmetry of d about the real axis. The coefficients b_n are linked to $u_k = P^{(k-1)}(0)$ by the relation

$$b_n = \frac{1}{s_1 \ldots s_n} \sum_{k=1}^{n+1} |S(n, k-1)| u_k. \tag{30}$$

$S(n, k)$ are the generalised Stirling numbers of first kind.

The GFS converges for $Re(1/t) > 1/T$ and $s_n = n\omega e^{-i\theta}$ for any $\omega > \omega_0$ and some $\omega_0 > 0$. ω_0 can be linked to the location of the closest singularity of P in the complex Borel plane. It is said that, numerically, the convergence of the GFS is more and more rapid as ω get close to ω_0 [41]. For linear equations, the closest singularity can be "read" from the coefficients of the equations and can be evaluated algorithmically. But for non-linear equations, the determination of ω_0 is harder.

Due to the explosive growth of the Stirling numbers, relation (30) should not be used as is, if the truncature order N is very big. Instead, the reccurence relation

$$|S(j+1, k)| = s_j |S(j, k)| + |S(j, k-1)|, \quad 1 \leq k \leq j+1, \tag{31}$$

on the Stirling numbers permits to compute the first terms of (29) as follows (see [41]).

Assume that $\theta = 0$. Set $\tau = 1/t$ and v_{n+1} the n-th term of the GFS, that is

$$
\begin{aligned}
v_{n+1} &= \frac{b_n n!}{\tau(\tau + s_1) \cdots (\tau + s_k)} \\[2ex]
&= \frac{\displaystyle\sum_{k=0}^{n} |S(n,k)| v_{k+1}^{(1)} \tau^{k+1}}{\tau(\tau + s_1) \cdots (\tau + s_k)}
\end{aligned}
$$

where $v_{k+1}^{(1)} = u_k t^{k+1}$. Without explicitly using the Stirling numbers, v_{n+1} can be computed recursively by the relations

$$
\begin{cases}
v_{k+1}^{(1)} &= u_k t^{k+1}, & k \geq 0 \\[2ex]
v_{k+1}^{(j+1)} &= \dfrac{s_{j-1} v_{k+1}^{(j)} + \tau v_k^{(j)}}{+ s_j}, & k \geq j \geq 1 \\[2ex]
v_{n+1} &= v_{n+1}^{(n+1)}, & n \geq 0.
\end{cases}
\tag{32}
$$

5 Numerical comparisons

In this section, we numerically compare the GFS algorithm with the Borel-Padé procedure. We take $s_i = i$.

5.1 Discretisation errors

In Section 3.2, we took Euler's equation (11) to analyse the effect of numerical discretisation on the Borel-Padé algorithm. We now apply the GFS algorithm to the same equation. As seen on Figure 2, the Borel-Padé solution is closer to the reference solution than the GFS one. We conclude that, if the quadrature of the Laplace integral is the only source of error, then BP is better than GFS.

However, in realistic applications, there is only slim chance that the Padé representation P is exact. With the next example, the effect of the Padé approximation error combined to the quadrature is illustrated.

Consider for example the following equation:

$$
\begin{cases}
\dfrac{\mathrm{d}u}{\mathrm{d}t} + u^2 = 0 \\[2ex]
u(0) = 1.
\end{cases}
\tag{33}
$$

The formal series solution is:
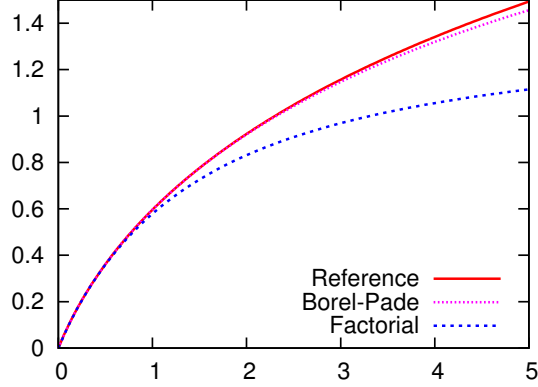
Figure 2: Approximate solutions of Euler's equation

$$\hat{u}(t) = \sum_{k \geq 0}(-1)^k t^k \tag{34}$$

which is convergent inside the unit disc. The Borel transform of (34) is

$$\sum_{k \geq 0}\frac{(-1)^{k+1}}{k!}\,\xi^k. \tag{35}$$

It prolonges into the exponential function

$$P(\xi) = -e^{-\xi}, \tag{36}$$

After the Laplace transformation, we retrieve the exact solution

$$1 - \int_0^{+\infty} e^{-\xi}e^{-\xi/t}\,\mathrm{d}\xi = \frac{1}{1+t}. \tag{37}$$

In this case, the Borel summation is, theoretically, exact.

Numerically, with BP, say when $N_1 = N_2 = 3$, the function (36) is approximated by the Padé approximant

$$P^N(\xi) = \frac{1 - \dfrac{1}{2}\xi + \dfrac{1}{10}\xi^2 - \dfrac{1}{120}\xi^3}{1 + \dfrac{1}{2}\xi + \dfrac{1}{10}\xi^2 + \dfrac{1}{120}\xi^3} \tag{38}$$

which is analytic on the real positive axis. This representation leads to a good approximation of the solution for small values of t. But when t grows, Figure 3 shows that BP suffers the effect of the numerical integration and of the truncation. As for it, GFS gives the exact Borel sum.
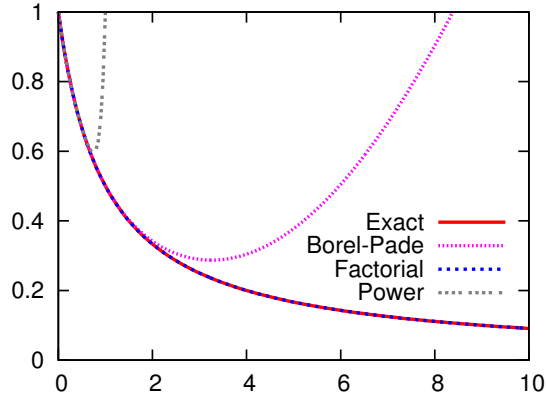
12

Figure 3: Approximate solution to equation (33)

This example clearly shows the advantage of avoiding the Padé approximation and Laplace transform in the numerical computation of the Borel sum of the solution, for some problems.

In the next section, we examine the case where a pole appears in the Padé approximant.

5.2 Pole in Padé approximant

Consider the system

$$\begin{cases} t\dfrac{\mathrm{d}u}{\mathrm{d}t} = u^2 + u, \\[2mm] \dfrac{\mathrm{d}u}{\mathrm{d}t}(0) = \lambda. \end{cases} \tag{39}$$

Since Padé approximants are invariant under dilatations [15, 6], it can easily be seen from (19) that, contrarily to $P(\xi)$, the rational approximation P^7 has a pole on $\mathbb{R}^+$, for any $\lambda > 0$.

With $\lambda = 8$, Figure 4 shows that, for small values of t, the pole can be simply ignored. But for some values of t, it produces significant errors in BP. With GFS, the appproximate solution does not present a weird behaviour.

5.3 Non-unique Padé representation

We now examine the case where the Padé approximant is not unique. Let us consider problem (20) again with $N_1 = N_2 = 2$. Theoretically, the Borel summation method gives the exact solution. However, as discussed previously, BP does not give any approximate solution without an additional optimisation. As for it, GFS encounters no problem. The approximate solution is presented on Figure 5.
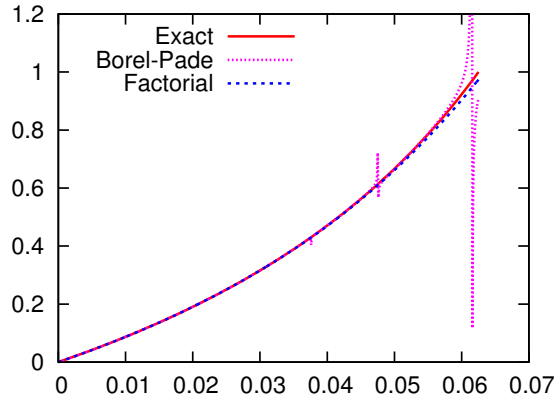
13

Figure 4: Pole in the Padé approximant

The above tests show that many paramaters have to be considered with BP for it to be successful. On the other hand, GFS will always give an approximate solution, and is then more robust. How good this solution is depends on the approximation order N (and, of course, on the problem).

In the last section, some other numerical experiments are presented.

6   Other interesting tests

Consider the initial value problem

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t} + u^3 = 0 \\[2mm] u(0) = 1 \end{cases} \tag{40}$$

Figure 6 compares the approximate solutions with BP and GFS. It shows that both algorithms reproduce the exact solution with a relatively good precision for small values of t. Next, around $t = 1$, differences with the exact solution start to be visible, first with BP, and then with GFS.

If we take a closer look at the error, we can see (Figure 7) that BP is more precise for small t. But around $t = 0.4$ the error curves intersect and GFS becomes better.

Similar conclusions can be drawn with the following problem

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t} - \dfrac{1 + u^2}{1 + t^2} = 0, \\[2mm] u(0) = -1. \end{cases} \tag{41}$$
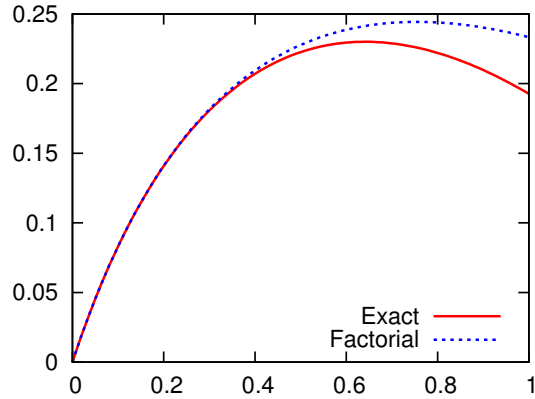
14

Figure 5: Equation (20)

The exact solution

$$u(t) = \tan(\arctan(t) - 1) \qquad (42)$$

is correctly predicted by both algorithms, as seen on Figure 8 for small t.

Figure 9 shows that at the beginnig, the error is smaller with BP than with GFS, but the tendency is reversed for bigger t.

Lastly, consider the logistic equation [20]

$$\frac{\mathrm{d}u}{\mathrm{d}t} = ru(1 - u) - h, \qquad (43)$$

which models the dynamics of a population. The growth rate r is set to 1, and the harvesting rate h to $1/5$. Figure 10 presents the approximate solutions with an initial population size $u(0) = 3$. It shows that the factorial series algorithm predicts a good approximate solution even for large values of t, contrarily to BP.

7   Conclusion

In this article, we analysed the weak points of BP algorithm for numerical computation of the Borel sum. We also saw that GFS algorithm does not suffer from the same problem of robustness, and can be considered as a serious alternative to BP.

Only one step of these algorithms have been considered here. For realistic resolutions, one has to iterate the algorithms. This means that, when the error reaches a prescribed tolerance value, a new series expansion is computed with the last acceptable solution as initial condition, and the resummation algorithm is applied again. This operation is known as a continuation. Relations such as (27) may be used to evaluate this error.
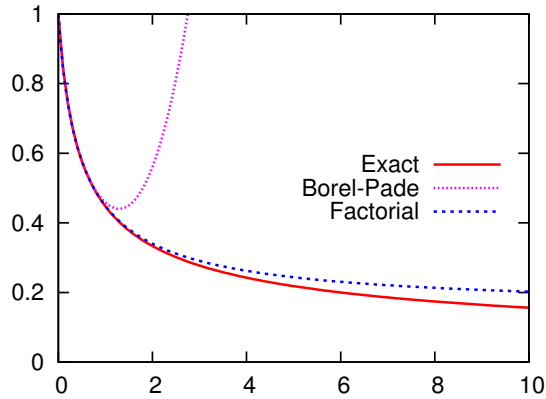
Figure 6: Approximate solutions of equation (40)

The next step would be a CPU-time comparison between BP with robustness amelioration and GFS, with continuation.

References

[1] V. Adukov and O. Ibryaeva, A new algorithm for computing padé approximants, *ArXiv*, **1112.5694v**.

[2] G. Baker, J. Gammel and J. Wills, An investigation of the applicability of the Padé approximant method, *Journal of Mathematical Analysis and Applications*, **2** (1961), 405–418.

[3] B. Beckermann and A. Ana Matos, Algebraic properties of robust Padé approximants, *arXiv:1310.2438 [math.NA]*.

[4] J. Boyd, Exponentially convergent Fourier-Chebshev quadrature schemes on bounded and infinite intervals, *Journal of Scientific Computing*, **2**.

[5] C. Brezinski, Rationnal approximation to formal power serie, *Journal of Approximation Theory*, 295–317.

[6] C. Brezinski and J. Van Iseghem, Padé approximations, in *Handbook of Numerical Analysis* (eds. P. G. Ciarlet and J. L. Lions), vol. 3, Elsevier, 1994, 47 – 222,

[7] A. Bultheel, Recursive algorithms for nonnormal Pade tables, *SIAM Journal on Applied Mathematics*, **39** (1980), 106–118.

[8] O. Costin, G. Luo and S. Tanveer, Divergent expansion, Borel summability and three-dimensional Navier-Stokes equation, *Philosophical Transactions*
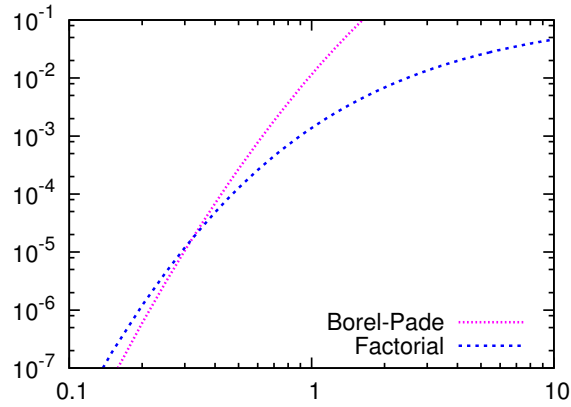
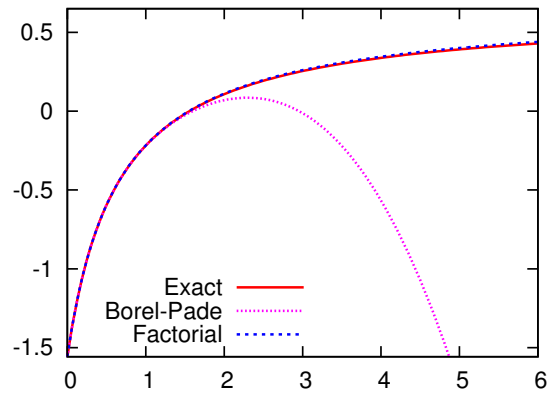Figure 7: Absolute errors on equation (40) in log-log scale

Figure 8: Exact and approximate solution of equation (41)

of the Royal Society A: Mathematical, Physical and Engineering Sciences, **36** (2008), 2775–2788.

[9] P. J. Davis and P. Rabinowitz, Ignoring the singularity in approximate integration, *Journal of the Society for Industrial and Applied Mathematics: Series B, Numerical Analysis*, **2** (1965), 367–383.

[10] A. Deeb, A. Hamdouni, E. Liberge and D. Razafindralandy, Borel-Laplace summation method used as time integration scheme, *ESAIM: Procedings and Surveys*, **45** (2014), 318–327.

[11] E. Delabaere and J.-M. Rasoamanana, Sommation effective d'une somme de Borel par séries de factorielles, *Annales de l'institut Fourier*, **57** (2007), 421–456.
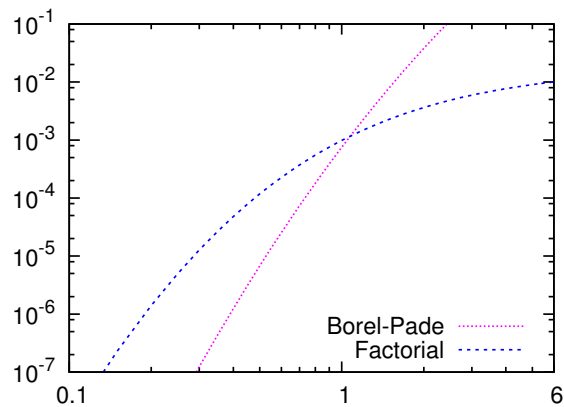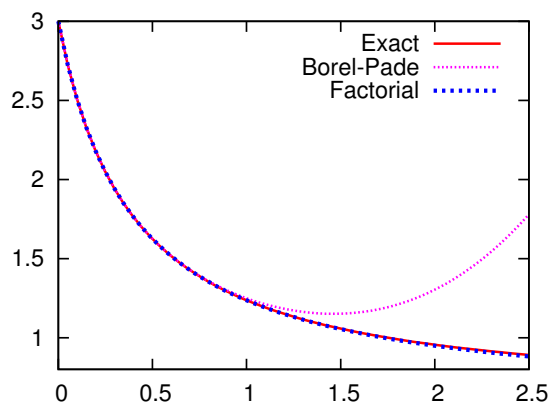
Figure 9: Absolute errors for equation (41)

Figure 10: Logistic equation with harvesting

[12] W. Gautschi, Gauss-type quadrature rules for rational functions, *arXiv:math/9307223*.

[13] W. Gautschi, The use of rational functions in numerical quadrature, *Journal of Computational and Applied Mathematics*, **133** (2001), 111–126.

[14] W. Gautschi, Quadrature formulae on half-infinite intervals, *BIT Numerical Mathematics*, **31** (1991), 437–446.

[15] J. Gilewicz, *Approximants de Padé*, vol. 667 of Lecture Notes in Mathematics, Springer-Verlag, 1978.

[16] J. Gilewicz and Y. Kryakin, Froissart doublets in Padé approximation in the case of polynomial noise, *Journal of Computational and Applied Mathematics*, **153** (2003), 235–242.

[17] J. Gilewicz and M. Pindor, Padé approximants and noise: A case of geometric series, *Journal of Computational and Applied Mathematics*, **87** (1997), 199 – 214.

[18] P. Gonnet, S. Güttel and L. Trefethen, Robust Padé approximation via SVD, *SIAM Review*, **51** (2013), 101–117.

[19] N. Hall, Interview of Sir Michael Berry by Nina Hall: Caustics, catastrophes and quantum chaos, *Nexus News*, pp. 4–5.

[20] M. Hirsch, S. Smale and R. Devaney, *Differential equations, dynamical systems, and an introduction to chaos*, Elsevier, 2013.

[21] H. Kleinert and V. Schulte-Frohlinde, *Critical properties of Φ^4-theories*, World Scientific, 2000.

[22] V. Kowalenko, *The Stokes Phenomenon, Borel Summation and Mellin-Barnes Regularisation*, Bentham, 2009.

[23] R. Kumar and J. M., Quadrature formulas for semi-infinite integrals, *Mathematics of Computation*, **28** (1974), 499–503.

[24] D. Lubinsky, Reflections on the Baker-Gammel-Wills (Padé), in *Analytic Number Theory, Approximation Theory, and Special Functions* (eds. G. V. Milovanović and M. T. Rassias), Springer New York, 2014, 561–571.

[25] D. S. Lubinsky and P. Rabinowitz, Rates of convergence of Gaussian quadrature for singular integrands, *Mathematics of Computation*, **43** (1984), pp. 219–242.

[26] D. Lutz, M. Miyake and R. Schäfke, On the Borel summability of divergent solutions of the heat equation, *Nagoya Mathematical Journal*, **154** (1999), 1–29.

[27] G. Lysik, Borel summable solutions of the Burgers equation, *Annales Polonici Mathematici*, **95** (2009), 187–197.

[28] G. Lysik and S. Michalik, Formal solutions of semilinear heat equations, *Journal of Mathematical Analysis and Applications*, **341** (2008), 372 – 385.

[29] W. Mascarenhas, Robust Padé approximants can diverge, *arXiv:1309.5753 [math.NA]*.

[30] N. Nielsen, Recherches sur les séries de factorielles, *Annales Scientifiques de l'E.N.S. 3è série*, **19** (1902), 409–453.

[31] N. Nielsen, Les séries de factorielles et les opérations fondamentales, *Mathematische Annalen*, **59** (1904), 355–376.

[32] N. Nielsen, Sur les séries de factorielles et la fonction gamma (extrait d'une lettre adressée à M.-N. de Sonin à Saint-Pétersbourg), *Annales Scientifiques de l'E.N.S. 3è série*, **23** (1906), 145–168.

[33] N. Nörlund, *Vorlesungen über differenzenrechnung*, Srpinger Verlag, 1924.

[34] N. Nörlund, *Leçons sur les séries d'interpolation*, Gauthier-Villard et Cie, 1926.

[35] S. Pincherle, Sulle serie di fattoriali. nota II, *Atti della Reale Accademia dei Lincei, Rendiconti della Classe di Scienze Fisiche, Matematiche e Naturali. Series 5*, **11** (1902), 417–426.

[36] J.-P. Ramis, Séries divergentes et théories asymptotiques, in *Journées X-UPS 1991*, 1991, 7–67.

[37] J.-P. Ramis, Les développements asymptotiques après poincaré : continuité et... divergences, *Gazettes des Mathématiciens*.

[38] D. Razafindralandy and A. Hamdouni, Time integration algorithm based on divergent series resummation, for ordinary and partial differential equations, *Journal of Computational Physics*, **236** (2013), 56–73.

[39] H. Stahl, Conjectures around the Baker-Gammel-Wills conjecture, *Constructive Approximation*, **13** (1997), 287–292,

[40] H. Stahl, Spurious poles in Padé approximation, *Journal of Computational and Applied Mathematics*, **99** (1998), 511–527.

[41] J. Thomann. Resommation des séries formelles. Solutions d'équations diffrentielles linéaires ordinaires du second ordre dans le champ complexe au voisinage de singularits irrégulières. *Numerische Mathematik*, 58:503–535, 1990.

[42] J. Thomann, Procédés formels et numériques de sommation de séries solutions d'équations différentielles, in *Journées X-UPS 1991, Séries divergentes et procédés de resommation* (ed. C. de mathématiques), 1991, 101–114.

[43] J. Thomann, *Formal and numerical summation of formal power series solutions of ODE's*, Technical report, CIRM Luminy, 2000.

[44] F. Thomlinson, Generalized factorial series, *Transactions of the American Mathematical Society*, **31**.

[45] M. Thomson, *The Calculus Of Finite Differences*, Macmillan and Company, 1933.

[46] van Deun J., A. Bultheel and P. Vera, On computing rational Gauss-Chebyshev quadrature formulas, *Mathematics of Computation*, **75** (2006), 307–326.

[47] G. N. Watson, The transformation of an asymptotic series into a convergent series of inverse factorials, *Rendiconti del Circolo Matematico di Palermo*, **34** (1912), 41–88.

[48] E. Weniger, Summation of divergent power series by means of factorial series, *arXiv:1005.0466 [math.NA]*.