Generalized Multivariate Padé Approximants

Philippe Guillaume and Alain Huard

Mathématiques pour l’Industrie et la Physique, Institut National des Sciences Appliquées,
Complexe Scientifique de Rangueil, UMR 5640, CNRS, 31077 Toulouse Cedex, France

and

Vincent Robin

CERFACS, 42, Avenue G. Coriolis, 31057 Toulouse Cedex, France

Communicated by Hans Wallin

Received February 17, 1997; accepted October 29, 1997

A new definition of multivariate Padé approximation is introduced, which is a natural generalization of the univariate Padé approximation and consists in replacing the exact interpolation problem by a least squares interpolation. This new definition allows a straightforward extension of the Montessus de Ballore theorem to the multivariate case. Except for the particular case of the so-called homogeneous Padé approximants, this extension has up to now been impossible to obtain in the classical formulation of the multivariate Padé approximation. Besides, the least squares formulation can also be applied to the univariate case, and provides an alternative to the classical Padé interpolation.

AMS Subject Classification: 65D15.
Key Words: Padé approximants; multivariate Padé approximants; convergence of multivariate Padé approximants.

1. INTRODUCTION

Many applications in mechanics or electromagnetism aim to compute a physical quantity \( f \) like strain, stress, reflection, or transmission coefficients. The computation of such a function \( f \) requires solving a partial differential equation which usually depends upon several parameters like shape, material properties, or frequency. One needs the knowledge of the quantity not only for a particular value of each of these parameters, but for a whole range of their values. Recently, new methods have been introduced which allow the function \( f \) to be expressed as a Taylor series with respect to the parameters [16–18]. Due to the nature of the problem, the function \( f \) is often not holomorphic but only meromorphic with respect to these
parameters, and a multivariate Padé representation of this function is more appropriate than a simple Taylor expansion.

The problem is that many difficulties appear when one tries to apply the techniques developed in the univariate Padé approximation theory to two or more variables. The natural ordering of $\mathbb{N}$ is lost on $\mathbb{N}^2$, the concept of degree is no longer clear, and “no formal equation analogous to (the univariate case) gives the correct number of linear equations to determine the ratios of the coefficients in the approximation” [6]. Thus several choices have been made in order to define multivariate Padé approximants. We mention only classical Padé approximants, i.e., rational fractions in several variables. For other types of approximants, we refer to [5, 15, 23].

J. Chisholm proposed a particular choice in [6], and A. Cuyt proposed another one in [8], the so-called homogeneous approximants. These two approximants are diagonal in the sense that they need a symmetric knowledge of the power series coefficients of the function to be approximated. This is not convenient in applications where the behavior of the function is different with respect to each variable, and/or where the derivatives with respect to each variable do not have the same computational cost.

A general definition including the previous ones was given by D. Levin in [20]. Unfortunately, only one uncontested proof of uniform convergence has been established, which concerns the particular case of the homogeneous approximants [9] (see [15] for some comments). The reason why no convergence has been obtained in the general case is that there is a lack of consistency if one chooses to write as many equations as unknowns for more than one variable, where consistency means that if $f = h/g$ is a rational fraction, then its Padé approximant $P/Q$ must be equal to $f$ if the “degrees” of $P$ and $Q$ are correctly chosen. Thus it seems necessary to introduce a new definition of multivariate Padé approximation.

If we abandon the arbitrary choice of having as many equations as unknowns, we will see that it is possible to obtain consistency and that the difficulties encountered in proving the multivariate Montessus de Ballore theorem vanish. It will appear that the equations which define the Padé coefficients have to be solved in a weighted least squares sense. In the univariate case and for a particular choice of the interpolation set of indices, this new definition coincides with the usual definition of the Padé approximation; thus it is a generalization of the univariate Padé approximation to the multivariate case.

The outline of this paper is as follows. In Section 2, the consistency of a rational approximation is discussed and a new definition of multivariate Padé approximation is proposed. In Section 3, the convergence is established and illustrated on a simple example.
2. GENERALIZED MULTIVARIATE PADÉ APPROXIMANTS

First, we recall some standard notation.

2.1. Notation

Let \( d \) be a positive integer. For all finite subset \( M \subseteq \mathbb{N}^d \), we denote by \( P_M \) the set of the polynomials \( P \in \mathbb{C}[z] \) having the form \( P(z) = \sum_{\alpha \in M} P_\alpha z^{\alpha} \). We use the standard notation \( z^{\alpha} = \prod_{i=1}^d z_i^{\alpha_i} \) for \( \alpha \in \mathbb{N}^d \) and \( z \in \mathbb{C}^d \).

As usual, a finite subset \( M \subset \mathbb{N}^d \) has the rectangular inclusion property if the conditions \( \alpha \in \mathbb{N}^d, \beta \in M, \) and \( \alpha \leq \beta \) imply \( \alpha \in M \). The standard partial order of \( \mathbb{N}^d \) is used; that is, \( \alpha \leq \beta \) means \( \alpha_i \leq \beta_i \) for \( 1 \leq i \leq d \). All the finite subsets of \( \mathbb{N}^d \) which are considered in this paper are supposed to have the rectangular inclusion property, and we denote by \( |M| \) the number of elements of \( M \).

We say that a polynomial \( P \in \mathbb{P}_M \) is \( M \)-maximal if for all polynomial \( Q \in \mathbb{P} \), the condition \( PQ \mid P_M \) implies \( Q \in \mathbb{C} \).

If \( M \) and \( N \) are two finite subsets of \( \mathbb{N}^d \), we denote by \( M \ast N \) the set \( \{ \alpha + \beta : \alpha \in M, \beta \in N \} \). If \( P \in \mathbb{P}_M \) and \( Q \in \mathbb{P}_N \), then \( PQ \in \mathbb{P}_{M \ast N} \).

For a function \( f \) which is holomorphic around the origin, we denote by \( f_\alpha \) the coefficient of \( z^{\alpha} \) in the power series expansion \( f(z) = \sum_{\alpha \in \mathbb{N}^d} f_\alpha z^{\alpha} \).

2.2. Consistency of a Rational Approximation

Let us recall the definition of a Padé approximant \( P/Q \) of a function \( f \) in the univariate case. The coefficients of the polynomials \( P \in \mathbb{P}_M \) and \( Q \in \mathbb{P}_N \) are defined by the three sets \( M = \{0, 1, ..., m\}, N = \{0, 1, ..., n\}, E = \{0, 1, ..., e\} \), the linear and homogeneous system

\[
(Qf - P)_\alpha = 0, \quad \forall \alpha \in E, 
\]

and the conditions

\[
E = M \ast N, 
\]

\[
|E| = |M| + |N| - 1. 
\]

An extra condition such as \( Q(0) = 1 \) is usually added in order to avoid the zero solution, and a square system is obtained.

The two conditions (2) and (3) are equivalent in the univariate case, and it seems to be a rather complicated way of writing \( e = m + n \). Nevertheless it is impossible to preserve both of them for more than one variable, because the equality \( |M \ast N| = |M| + |N| - 1 \) holds only for \( d = 1 \).

Up to now, almost all the attempts to define multivariate Padé approximants have been made in order to preserve condition (3), which leads to a square system. The drawback of this choice is the lack of
consistency. This can be illustrated by the following example. Let \( f(x, y) = \frac{1}{(1-x)(1-y)} \), \( M = \{(i, j) \in \mathbb{N}^2; i + j \leq 2\} \), \( N = \{0, 1\}^2 \), \( E = \{0, 1, 2\} \), \( P(x, y) = 1 + x + y + x^2 + y^2 \in \mathbb{P}_M \), and \( Q(x, y) = 1 - xy \in \mathbb{P}_N \). Then \( P \) and \( Q \) are solutions to Eq. (1), condition (3) is fulfilled, but \( PQ \neq f \). Such an indetermination of the denominator coefficients appears also for higher degrees of the numerator, making it impossible to obtain uniform convergence of the Padé approximants to the function \( f \) on compact subsets of \( \{(x, y) \in \mathbb{C}^2; (1-x)(1-y) \neq 0\} \) when \( M \) increases.

We choose to preserve condition (2), which yields the following elementary result.

**Proposition 2.1.** Let \( f = \frac{h}{g} \) be an irreducible fraction, where \( h \in \mathbb{P}_M \), \( g \in \mathbb{P}_N \), \( g \) is \( N \)-maximal, and \( g(0) \neq 0 \). Let \( (P, Q) \in \mathbb{P}_M \times \mathbb{P}_N \) be a solution to the system

\[
(Qf - P)_{\alpha} = 0, \quad \forall \alpha \in M \ast N.
\]

Then there exists a constant \( c \in \mathbb{C} \) such that

\[
P = ch, \quad Q = cg.
\]

**Proof.** Due to \( g(0) \neq 0 \) and the rectangular inclusion property, system (4) is equivalent to \( (Qh - Pg)_{\alpha} = 0 \), for all \( \alpha \in M \ast N \). Since \( Qh - Pg \in \mathbb{P}_{M \ast N} \), we have \( Qh = Pg \). It follows from the Gauss lemma that there exists a polynomial \( c \) such that

\[
P = hc, \quad Q = gc.
\]

Now \( Q \in \mathbb{P}_N \) and \( g \) is \( N \)-maximal, thus \( c \in \mathbb{C} \).

Observe that system (4) has a non-trivial solution, e.g., \( (h, g) \), although it is usually strongly over-determined. However, in the case where \( h \) is not a polynomial, this system generally has no solution other than zero. For this reason, we propose a least squares formulation of the problem.

**2.3. Definition**

Let a function \( f: \mathbb{C}^d \rightarrow \mathbb{C} \) be holomorphic around the origin. For two polynomials \( P \in \mathbb{P}_M \), \( Q \in \mathbb{P}_N \), for \( E \subseteq \mathbb{N}^d \) and \( \rho \in \mathbb{R}_+^E \), we define

\[
f(P, Q) = \left( \sum_{\alpha \in E} \rho^{2\alpha} \| (Qf - P)_{\alpha} \|^2 \right)^{1/2}.
\]

The choice of \( \rho \) will be stated precisely in the next section.
**Definition 2.1.** Let $M, N \in \mathbb{N}^d$, and $E \supseteq M \ast N$. A generalized multivariate Padé approximant of $f$ is a fraction $P/Q$ with $(P, Q) \in \mathbb{P}_M \times \mathbb{P}_N$, $\sum_{x \in N} |Q_x|^2 = 1$, and
\[
|j(P, Q)| \leq j(R, S), \quad \forall (R, S) \in \mathbb{P}_M \times \mathbb{P}_N, \quad \sum_{x \in N} |S_x|^2 = 1.
\] (6)

A solution to this problem will be denoted by $[M, N]_f$.

**Remark 2.1.** For $d = 1$ and $E = M \ast N$, this definition coincides with the standard definition of univariate Padé approximants (except that the usual condition $Q(0) = 1$ is replaced by $\sum_{x \in N} |Q_x|^2 = 1$). For $d = 1$ and $E \supseteq M \ast N$, we obtain a least squares formulation of the univariate Padé approximation which provides an alternative to the exact Padé interpolation.

**3. CONVERGENCE OF THE GENERALIZED MULTIVARIATE PADÉ APPROXIMANTS**

In this section we give the main result of this paper, which is the proof of the convergence. It is based on the technique used by J. Karlsson and H. Wallin for proving the convergence in the univariate case [19].

Let a function $f$ be meromorphic on a neighborhood of the polydisc $\bar{B}(0, \rho) = \{z \in \mathbb{C}^d; |z_i| \leq \rho, i = 1, \ldots, d\}$. We suppose that the function $f$ is of the form
\[
f(z) = \frac{h(z)}{g(z)},
\]
where $g$ is a polynomial such that $g(0) \neq 0$, and $h$ is holomorphic on a neighborhood of $\bar{B}(0, \rho)$. A finite subset $N \subset \mathbb{N}^d$ is chosen in such a way that $g$ is $N$-maximal.

We need the following definition: for all sequences $(M_n)_{n \geq 0}$, $M_n \subset \mathbb{N}^d$, we say that $\lim_{n \to \infty} M_n = \infty$ if for all bounded subsets $B$ of $\mathbb{N}^d$, there exists an integer $k$ such that $B \subset M_n$ for all $n \geq k$. For the sake of simplicity, we will omit the subscript $n$ and write $M \to \infty$.

Consider the decomposition of the polynomial $g$ into irreducible factors
\[
g = \prod_{i=1}^{l} g_i^{\gamma_i},
\]
and the associated algebraic set $G$ with its decomposition into irreducible components $G_i$.
\[ G = \{ z \in \mathbb{C}^d; g(z) = 0 \}, \]
\[ G_i = \{ z \in \mathbb{C}^d; g_i(z) = 0 \}. \]

**Theorem 3.1.** Assume that \( G_i \cap B(0, \rho) \neq \emptyset \) for \( 1 \leq i \leq l \), and that \( h(z) \neq 0 \) for all \( z \in G \cap B(0, \rho) \). Let \( ([M, N])_M \) be a sequence of generalized multivariate Padé approximants with \( M \to \infty \). Then

\[
\lim_{M \to \infty} [M, N]_f(z) = f(z),
\]
uniformly on all compact subsets of \( \{ z \in B(0, \rho); g(z) \neq 0 \} \).

**Proof.** Let \( (P^M, Q^M) \in \mathbb{P}_M \times \mathbb{P}_N \) be a solution to problem (6). Consider the function

\[
H^M = g(Q^M f - P^M),
\]
which is holomorphic on a neighborhood of \( B(0, \rho) \). We need the following lemma, whose proof is given at the end of the section.

**Lemma 3.2.** We have

\[
\lim_{M \to \infty} H^M(z) = 0,
\]
uniformly on all compact subsets of \( B(0, \rho) \).

According to Definition 2.1, the sequence \( (Q^M)_M \) is bounded in the finite dimensional space \( \mathbb{P}_N \) equipped with the norm \( \| P \| = (\sum_{x \in N} |P_x|^2)^{1/2} \). Consider an arbitrary subsequence, still denoted by \( (Q^M)_M \) for simplicity, which converges to a polynomial \( Q \in \mathbb{P}_N \) with \( \| Q \| = 1 \). The subsequence \( (Q^M)_M \) converges also to \( Q \) uniformly on all compact subsets of \( \mathbb{C}^d \) when \( M \to \infty \).

Due to the assumptions of the theorem, the set \( G_i \cap B(0, \rho) \) is non-empty. For \( z \in G_i \cap B(0, \rho) \), we have \( H^M(z) = h(z) Q^M(z) \) and \( (H^M)_M \) converges to 0 on \( B(0, \rho) \), thus \( h(z) Q(z) = 0 \), and \( Q(z) = 0 \) because \( h(z) \neq 0 \). The set of regular points of \( G_i \) is open, connected, and dense in \( G_i \), thus \( Q = 0 \) on \( G_i \), and it follows from the Nullstellensatz that \( g \) divides \( Q \) [2, 24]. Similarly we can prove that \( g^j \) divides \( Q \) by considering the partial derivatives of \( H^M \), which implies that \( g \) divides \( Q \). But \( Q \in \mathbb{P}_N \) and \( g \) is \( N \)-maximal, thus there exists \( c \in \mathbb{C}, |c| = 1/\| g \| \), such that

\[ Q = cg. \]
Hence \( \lim_{M \to \infty} Q^M = cg \) uniformly on all compact subsets of \( \mathbb{C}^d \), and if we divide (7) by \( gQ^M \), we obtain with (8)

\[
\lim_{M \to \infty} \left( f \frac{P^M}{Q^M} \right)(z) = 0,
\]

uniformly on all compact subsets of \( \{ z \in B(0, \rho); g(z) \neq 0 \} \). As this holds for all convergent subsequences of the bounded sequence \( (Q^M)_M \), the whole sequence \( f - P^M/Q^M \) converges to zero in the same way.

**Proof of the lemma.** Denote by \( \Gamma_+ \) the torus \( \{ z \in \mathbb{C}^d, |z_i| = \rho, 1 \leq i \leq d \} \) with its usual orientation. We have \( gP^M \in \mathbb{P}_{M \ast N} \) and \( E \supseteq M \ast N \), thus

\[
H^M = \begin{cases} 
(hQ^M)_x & \text{if } x \notin E, \\
(g(Q^M f - P^M)_x) & \text{if } x \in E,
\end{cases}
\]

where \( (Q^M f - P^M)_E(z) = \sum_{x \in E} (Q^M f - P^M)_x z^x \). Using the Cauchy integral we can write

\[
H^M = \frac{1}{(2\pi)^d} \int_{\Gamma_+} \frac{hQ^M}{z^{x+1}} \, dz, \quad \text{if } x \notin E, \quad (9)
\]

\[
H^M = \frac{1}{(2\pi)^d} \int_{\Gamma_+} \frac{g(Q^M f - P^M)_E}{z^{x+1}} \, dz, \quad \text{if } x \in E, \quad (10)
\]

where \( x + 1 = (x_1 + 1, \ldots, x_d + 1) \). We study the two cases (9) and (10) separately.

- **The case** \( x \notin E \). The sequence \( (Q^M)_M \) is bounded in \( \mathbb{P}_N \); thus there exists a constant \( c_1 \) such that \( |Q^M(z)| \leq c_1 \) for all \( z \in \Gamma_+ \). The function \( h \) is continuous on \( \Gamma_+ \), and, with \( c_2 = \max_{z \in \Gamma_+} |h(z)| \) and \( c = c_1 c_2 \), we obtain

\[
|H^M_x| \leq \frac{c}{\rho^x} \quad \text{if } x \notin E. \quad (11)
\]

- **The case** \( x \in E \). Using the change of variable \( z = (\rho_1 \exp(2\pi \theta_1), \ldots, \rho_d \exp(2\pi \theta_d)) \) in Eq. (10), we have

\[
H^M_x = \left[ \int_{[0, 1]^d} \frac{g(Q^M f - P^M)_E}{z^x} \, d\theta \right] \, d\theta.
\]

The Cauchy–Schwarz inequality yields

\[
|H^M_x| \leq \left( \int_{[0, 1]^d} \frac{|g|^2}{\rho^x} \, d\theta \right)^{1/2} \left( \int_{[0, 1]^d} |(Q^M f - P^M)_E|^2 \, d\theta \right)^{1/2}, \quad (12)
\]
and the second integral reads with the Parseval formula
\[
\left( \int_{[0, 1]} |(Q^M f - P^M)_e|^2 d\theta \right)^{1/2} = \left( \int_{[0, 1]} \left| \sum_{a \in E} (Q^M f - P^M)_a \rho_a^2 e^{2i\pi a} \right|^2 d\theta \right)^{1/2} = \left( \sum_{a \in E} \rho_a^{2a} \left| \left( Q^M f - P^M \right)_a \right|^2 \right)^{1/2},
\]
where \( \alpha \cdot \theta = \alpha_1 \theta_1 + \cdots + \alpha_d \theta_d \). Thus, using the definition (5) of the function \( j \) and \( c' = \max_{\gamma \in \Gamma} |g(\gamma)| \), we have
\[
|H^E_z| \leq C' \frac{1}{\rho} j(P^M, Q^M) \quad \text{if} \quad \alpha \in E. \quad (13)
\]
Due to the definition of \( P_M, Q_M \) and to \( (h_M, g) \in P_M \times P_N \), we have \( j(P^M, Q^M) \leq j(h_M, g) \), and gathering Eqs. (11) and (13), we obtain
\[
|H^M(z)| \leq C' j(h_M, g) \sum_{a \in E} \left| \frac{z}{\rho} \right|^a + c \sum_{a \neq E} \left| \frac{z}{\rho} \right|^a, \quad (14)
\]
where \(|z/\rho|^a = |z_1/\rho_1|^a \cdots |z_d/\rho_d|^a|\).

Once more using the Parseval formula and the change of variable \( z = \rho \exp(2i\pi \theta) \) we have
\[
j(h_M, g) = \left( \sum_{a \in E} \rho^{2a} \left| (h - h_M)_a \right|^2 \right)^{1/2} = \left( \frac{1}{2\pi} \int_{\Gamma_e} \frac{|h_{E \setminus M}|^2}{z} \, dz \right)^{1/2}.
\]
The function \( h \) is holomorphic on a neighborhood of \( \overline{B}(0, \rho) \); thus \( h_{E \setminus M} \) converges to 0 uniformly on \( \Gamma_e \), when \( M \to \infty \), and
\[
\lim_{M \to \infty} j(h_M, g) = 0. \quad (15)
\]

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f(x, y) )</td>
</tr>
<tr>
<td>( m = 5 )</td>
</tr>
<tr>
<td>( x = y = 1.5 )</td>
</tr>
<tr>
<td>( x = y = 2 )</td>
</tr>
<tr>
<td>( x = y = 3 )</td>
</tr>
<tr>
<td>Cond. number</td>
</tr>
</tbody>
</table>
Equations (14) and (15) prove that \( \lim_{M \to \infty} |H^M(z)| = 0 \), uniformly on all compact subsets of \( B(0, \rho) \).

**Remark 3.1.** Although the choice of \( \rho \) is important in the proof of convergence, numerical experiments have shown that the result is not very sensitive with respect to this choice.

**Example 3.1.** We illustrate the convergence of the approximants on the function

\[
 f(x, y) = \exp \left( \frac{xy}{2} \right) \left( \frac{1}{1 - \frac{x}{4} - \frac{y}{2}^2} + \frac{1}{1 - \frac{2x}{3} - \frac{y}{3}^2} + \frac{1}{1 + x - y} + \frac{1}{1 - xy} \right).
\]

Here we take \( N = \{0, \ldots, 6\}^2 \), \( M = \{0, \ldots, m\}^2 \), and we let \( m \) increase.

Table I compares the approximations \([M, N]f(x, y)\) for different values of \( m \), evaluated at the three points \((x, y)\) represented.

![Real poles of the approximants \([M, N]f\).](image-url)
FIG. 2. Normalized error $|[(M, N), (x, y) - f(x, y)]/\exp(xy/2)|$.

by a * in the last picture of Fig. 1. A good convergence can be observed from $m = 5$ to 13. Numerical instability appears for $m \geq 15$, due to the increasing condition number of the system, which is given in the last line of Table I. The computations have been performed in double precision (IEEE standard).

Figure 1 shows the singular set of the approximants for some values of $m$. The non-real roots of the denominators have been ignored.

Figure 2 shows the convergence of the approximants. Each graph gives the "normalized" error $|[(M, N), (x, y) - f(x, y)]/\exp(xy/2)|$ computed on the diagonal segment $\{(x, y) \in \mathbb{R}^2; y = x, -4 \leq x \leq 4\}$.

4. FINAL COMMENTS

We have shown that if the degree condition $|E| = |M| + |N| - 1$ is used in order to define multivariate Padé approximants, where the sets $E$, $M$, and $N$ satisfy the rectangular inclusion property, then the consistency and
the Montessus de Ballore theorem may be lost, whereas the condition $E \geq M \star N$ preserves both consistency and convergence. The proofs are natural generalizations of the univariate proofs, and the corresponding numerical method is not more difficult to implement than in the univariate case.

The application of generalized Padé approximants to a scattering problem in electromagnetism is actually under way, and will be presented in a forthcoming paper.

ACKNOWLEDGMENT

The authors are grateful to the referees for their thorough reading and their suggestions concerning the presentation of this paper.

REFERENCES


