HOW WELL CAN THE CONCEPT OF PADÉ APPROXIMANT BE GENERALIZED TO THE MULTIVARIATE CASE?

Dedicated to Haakon Waadeland, on the occasion of his 70^{th} birthday

ANNIE CUYT*

KEYWORDS: Padé approximation, multivariate, rational function

RUNNING HEAD: Multivariate Padé Approximation

ABSTRACT

What we know about multivariate Padé approximation has been developed in the last 25 years. In the next sections we compare and discuss many of these results. It will become clear that simple properties or requirements, such as the uniqueness of the Padé approximant and consequently its consistency property, can play a crucial role in the development of the multivariate theory. A separate section is devoted to a discussion of the convergence properties. At the end we include an extensive reference list on the topic.

1. PADÉ APPROXIMANTS AND TAYLOR SERIES.

Given a function $f(z)$ through its Taylor series expansion at a certain point in the complex plane, the Padé approximant $[n/m]^f$ of degree n in the numerator and m in the denominator for f is defined by (for simplicity we use the Taylor series at the origin)

$$\begin{aligned} f(z) &= \sum_{i=0}^{\infty} c_i z^i \\ p(z) &= \sum_{i=0}^{n} a_i z^i \\ q(z) &= \sum_{i=0}^{m} b_i z^i \\ (fq-p)(z) &= \sum_{i \geq n+m+1} d_i z^i \end{aligned} \tag{1}$$

with $[n/m]^f$ equal to the irreducible form of p/q. Three essential remarks about this definition have to be made.

First, a count of unknowns and conditions in (1) tells us that we have to compute $n+m+2$ coefficients a_i and b_i from $n+m+1$ linear equations

$$d_i = \sum_{k=0}^{i} c_k b_{i-k} - a_i = 0 \qquad i = 0, \ldots, n+m$$

* RESEARCH DIRECTOR FWO-VLAANDEREN, DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, UNIVERSITEIT ANTWERPEN (UIA), UNIVERSITEITSPLEIN 1, B–2610 WILRIJK, BELGIUM, EMAIL: CUYT@UIA.UA.AC.BE

This is always possible and choosing at least one of the unknowns $b_0, \ldots, b_m$ in the homogeneous system $d_i = 0$ for $i = n+1, \ldots, n+m$ does not change the rational function p/q. It only influences the numerator and denominator polynomials $p(z)$ and $q(z)$ in the sense that they are in fact only determined up to multiplicative factor. The irreducible form is usually normalized in such a way that the denominator evaluates to 1 at the origin (or the point at which the Taylor series development was considered).

Second, in order for $[n/m]^f$ to exist for all natural numbers n and m, one has to obtain the Padé approximant from the polynomials p and q satisfying the linear conditions (1), rather than imposing such conditions directly on the numerator and denominator of $[n/m]^f$. The reason for this is that when solutions of (1) are reducible, the numerator and denominator of the irreducible form do not necessarily satisfy (1) anymore [2, pp. 20–21]. However, if $f(0) \neq 0$, then

$$(fq - p)(z) = \sum_{i \geq n+m+1} d_i z^i \Longrightarrow (f - [n/m]^f)(z) = \sum_{i \geq \partial p_{n,m} + \partial q_{n,m} + t + 1} e_i z^i$$

where $\partial p_{n,m}$ and $\partial q_{n,m}$ respectively indicate the exact degree of the numerator and denominator of $[n/m]^f$ and $t \geq 0$.

Third, $[n/m]^f$ as given above is well-defined because one can prove that all solutions of (1) reduce to one and the same irreducible form, for fixed f, n and m [6, pp. 68]. Although this property is simple to prove in the univariate case, it causes great problems when defining multivariate Padé approximants.

Let us now take a look at the multivariate problem. We shall not use standard multi-index notation because it may obscure some points that we are trying to make. Given a Taylor series expansion (for simplicity we describe only the bivariate case but the higher-dimensional case is only notationally more difficult)

$$f(x, y) = \sum_{(i,j) \in \mathbb{N}^2} c_{ij} x^i y^j \tag{2}$$

one can group the different definitions for multivariate Padé approximants into four main categories, depending on how one deals with the data c_{ij}.

Rewriting $f(x, y)$ as

$$f(x, y) = \sum_{k=0}^{\infty} c_{i_k j_k} x^{i_k} y^{j_k}$$

is done in what we call the **equation lattice** group of definitions. This group includes popular definitions such as [27, 43, 53, 52, 48, 38, 46, 30]. Another way to deal with the information is to rewrite $f(x, y)$ as

$$f(x, y) = \sum_{k=0}^{\infty} \left(\sum_{i+j=k} c_{ij} x^i y^j \right)$$

We will refer to this approach as the **homogeneous** approach, and some very interesting and at the same time intriguing facts have to be told about it. A third group of definitions looks at the Taylor series development as

$$\begin{aligned} f(x, y) &= \sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} c_{ij} y^j \right) x^i \\ &= \sum_{i=0}^{\infty} c_i(y) x^i \end{aligned}$$

and treats the problem partly in a **symbolic** way. Interchanging the role of x and y does not necessarily lead to the same results. Definitions in this category can be found in [85, 91]. This type of definitions has not yet been studied extensively. A fourth approach builds on the link between Padé approximants and corresponding continued fractions. Since the univariate Padé approximant $[n/m]^f$ can be obtained as the convergent of a corresponding ordinary continued fraction, multivariate definitions have been introduced that consider convergents of so-called corresponding **branched continued fractions**. Because these definitions have been reviewed in separate papers on the subject [95, 100], we only include bibliographic material on these definitions here.

A few multivariate definitions are difficult to categorize: in [1, 3] the problem is treated as a moment problem, in [4] as a model reduction problem and in [7] as a least squares problem.

2. Univariate Padé approximants.

In this section we list the properties of the univariate Padé approximant that we want to examine for each of the multivariate generalizations below. It was already pointed out that we start from (1) and not from conditions on $[n/m]^f$ itself and that $[n/m]^f$ is always and uniquely defined in that way. The unicity of the **irreducible form** $[n/m]^f$ of the rational functions p/q with p and q satisfying (1) will be a point of discussion in the sequel. In the univariate case it is based on the next theorem that states that different solutions of (1) reduce to the same rational function.

Theorem 2.1:
Let p_1 and q_1 as well as p_2 and q_2 satisfy the conditions (1). Then

$$(p_1 q_2)(z) = (p_2 q_1)(z)$$

Another point of discussion is the desirability of certain properties for the multivariate Padé approximant. For instance the univariate Padé approximant automatically satisfies a **consistency property** because of the unicity of the irreducible form. This property means that for an irreducible rational function $f(z)$, given only by its Taylor series, the Padé approximation process reconstructs the given rational function when calculating its appropriate Padé approximant. This consistency property is in fact quite logical and hence very desirable.

Theorem 2.2:
If $f(z) = g(z)/h(z)$ with $h(0) = 1$ and

$$g(z) = \sum_{i=0}^{n} g_i z^i$$

$$h(z) = \sum_{i=0}^{m} h_i z^i$$

then for $f(z)$ irreducible and $k \geq n$ and $\ell \geq m$ we find $[k/\ell]^f = [n/m]^f = f$.

The univariate Padé approximant also satisfies a number of **covariance properties**. A number of operators Φ exist that can work on the series development f and commute more or less with the Padé operator $\mathcal{P}_{n,m}$ that associates with f its Padé approximant $[n/m]^f$:

$$\Phi\left[\mathcal{P}_{n,m}(f)\right] = \mathcal{P}_{n_\Phi, m_\Phi}\left[\Phi(f)\right]$$

with n_Φ and m_Φ depending on the considered Φ. It's easy to see that the operators Φ have to be rational.

The most important covariance property is the reciprocal covariance. It allows you for instance to mirror three-term recurrence relations among Padé approximants, that are valid only for $n \leq m$, to the case $m \leq n$ by switching from f to $1/f$.

THEOREM 2.3:
Let $f(0) \neq 0$ and let $[n/m]^f = p_{n,m}/q_{n,m}$, Then

$$[m/n]^{1/f} = \frac{q_{n,m}/f(0)}{p_{n,m}/f(0)}$$

THEOREM 2.4:
Let a, b, c and d be complex numbers with $cf(0) + d \neq 0$ and let $[n/n]^f = p_{n,n}/q_{n,n}$. Then

$$[n/n]^{(af+b)/(cf+d)} = \frac{(ap_{n,n} + bq_{n,n})/(c\,f(0) + d)}{(cp_{n,n} + dq_{n,n})/(c\,f(0) + d)}$$

Is there a multivariate definition that preserves all these properties or do we have to make a choice among the multivariate generalizations depending on which theorems we want our approximant to satisfy? Moreover, do we want the multivariate Padé approximant to satisfy a **projection property**, reducing to the univariate Padé approximant when all but one variable are equated to zero in the given function and its approximant? Discussion exists about a possible **factorization property**. Some researchers desire that if $f(x,y) = g(x)h(y)$, its multivariate Padé approximant is the product of the univariate Padé approximants for g and h. We think however that this depends greatly on the functions g and h in question. Consider for instance

$$g(x) = \exp(x) \qquad h(y) = \exp(y) \qquad f(x,y) = \exp(x+y)$$

Then

$$\begin{aligned}[1/2]^f \times [1/2]^g &= \frac{(1 + x/3)(1 + y/3)}{(1 - 2x/3 + x^2/6)(1 - 2x/3 + x^2/6)} \\ &\neq \frac{1 + (x+y)/3}{1 - 2(x+y)/3 + (x+y)^2/6}\end{aligned}$$

where this last function is a much more logical candidate as a multivariate Padé approximant for $f = gh$.

In the last century many **convergence properties** for univariate Padé approximants were given, describing their approximation power for several function classes. It makes sense to approximate locally meromorphic functions having only poles in a certain region, by rational functions whose denominator degree equals at least the number of poles in the considered region. Functions with a countable number of singularities, not necessarily poles, can very well be approximated by the Padé approximants $[n/n]^f$. Most of these results have one or other multivariate counterpart. The theorems proven for the different multivariate generalizations differ slightly in the conditions they impose on the multivariate function that is being approximated. More information on this can be found in section 6.

To top off the discussion we shall comment in short on the **computational algorithms** that exist for each of the multivariate definitions that are being discussed.

3. THE EQUATION LATTICE APPROACH.

3.1. Definition.

For $f(x,y)$ given by (2), we can define a multivariate Padé approximant p/q to f by determining $p(x,y)$ and $q(x,y)$ from accuracy–through–order conditions as follows. Let the polynomials $p(x,y)$ and $q(x,y)$ be of the general form

$$p(x,y) = \sum_{(i,j)\in N} a_{ij} x^i y^j \tag{3}$$

$$q(x,y) = \sum_{(i,j)\in D} b_{ij}x^iy^j \tag{4}$$

where N (for numerator) and D (for denominator) are nonempty finite subsets of $\mathbb{N}^2$. The sets N and D indicate in a way the degree of the polynomials $p(x,y)$ and $q(x,y)$. Let us denote

$$n+1 = \#N \qquad m+1 = \#D$$

In analogy with the univariate case we also choose a set of indices E (for equations) such that

$$N \subseteq E \tag{5a}$$

$$\#(E\backslash N) = m = \#D - 1 \tag{5b}$$

$$E \text{ satisfies the inclusion property} \tag{5c}$$

Here (5c) means that when a point belongs to the index set E, then the rectangular subset of points emanating from the origin with the given point as its furthermost corner, also lies in E. In other words,

$$(i,j) \in E \Longrightarrow \{(k,\ell) \mid k \le i, \ell \le j\} \subseteq E$$

We then impose the following accuracy-through-order conditions on the polynomials $p(x,y)$ and $q(x,y)$, namely

$$(fq-p)(x,y) = \sum_{(i,j)\in \mathbb{N}^2\backslash E} d_{ij}x^iy^j \tag{6}$$

Condition (5a) enables us to split the system of equations

$$d_{ij} = 0 \qquad (i,j) \in E$$

in an inhomogeneous part defining the numerator coefficients

$$\sum_{\mu=0}^{i}\sum_{\nu=0}^{j} c_{\mu\nu}b_{i-\mu,j-\nu} = a_{ij} \qquad (i,j) \in N \tag{7a}$$

and a homogeneous part defining the denominator coefficients

$$\sum_{\mu=0}^{i}\sum_{\nu=0}^{j} c_{\mu\nu}b_{i-\mu,j-\nu} = 0 \qquad (i,j) \in E\backslash N \tag{7b}$$

ans is not as essential as (5b). In fact condition (5a) and (5b) could be replaced by $\#N + \#D = \#E + 1$. By convention $b_{k\ell} = 0$ if $(k,\ell) \notin D$. Condition (5b) guarantees the existence of a nontrivial denominator $q(x,y)$ because the homogeneous system has one equation less than the number of unknowns and so one unknown coefficient can be chosen freely. Condition (5c), together with the Leibniz product rule, finally takes care of the real Padé approximation property, namely

$$q(0,0) \neq 0 \Longrightarrow (f - \frac{p}{q})(x,y) = \sum_{(i,j)\in \mathbb{N}^2\backslash E} \tilde{d}_{ij}x^iy^j \tag{7c}$$

If E does not satisfy the inclusion property, then (6) does not imply

$$\left[\frac{1}{q}(fq-p)\right](x,y) = \left(f-\frac{p}{q}\right)(x,y) = \sum_{(i,j)\in I\!N^2\setminus E} \tilde{d}_{ij}x^iy^j$$

since in that case $f - p/q$ also contains terms resulting from the multiplication of the holes in E by $(1/q)(x,y)$. We denote the set of rational functions p/q satisfying (6) by $[N/D]_E^f$ and we call it the general multivariate Padé approximant for f.

For a univariate function $f(z)$ the above construction reduces to defining subsets N, D and E of $I\!N$ to respectively index the numerator, denominator and approximation conditions (1). These univariate index sets are then given by $N = \{0,\ldots,n\}, D = \{0,\ldots,m\}$ and $E = \{0,\ldots,n+m\}$. In going from one to many variables a variety of choices for these index sets is now introduced.

Because of the freedom in choosing the sets N, D and E, the equation lattice definition covers a variety of approximation schemes, sometimes with minor variations on the general definition above. In [50, 52, 53, 49, 55, 20, 21] rectangular schemes are studied, in [22, 21, 29, 46, 38] triangular schemes, and in [27, 43, 44, 45] a combination of both. For more information we also refer to [48, 36].

In general, uniqueness of the general multivariate Padé approximant, in the sense that all rational functions in $[N/D]_E^f$ reduce to the same irreducible form, is not guaranteed, unless the index set $E \setminus N$ supplies a homogeneous system of linearly independent equations. It is obvious that at least one non-trivial solution of (6) exists because the number of unknown coefficients b_{ij} is one more than the number of conditions in (7b). But it is not so (unlike in the univariate case) that different solutions p_1, q_1 and p_2, q_2 of (6) are necessarily equivalent, meaning that $(p_1q_2)(x,y) = (p_2q_1)(x,y)$. Hence p_1/q_1 and p_2/q_2 may be different functions. In general one can only say that

$$(p_1q_2 - q_1p_2)(x,y) = \sum_{(i,j)\in N\times D\setminus E} e_{ij}x^iy^j$$

where

$$N \times D = \{(i+k, j+\ell) \mid (i,j) \in N, (k,\ell) \in D\}$$

One way to enforce a unicity property is to choose the index set E as large as possible, by adding conditions as soon as there are linearly dependent equations in (7b), but this is not always possible. This in fact amounts to weakening (5b) to $\#N + \#D \le \#E + 1$.

This phenomenon basically also explains why the equation lattice definitions do not satisfy a consistency property, unless again the index set $E \setminus N$ supplies a homogeneous system of linearly independent equations.

3.2. Consistency property.

The consistency property would mean that for an irreducible rational function

$$f(x,y) = \frac{g(x,y)}{h(x,y)} = \frac{\sum_{(i,j)\in N} g_{ij}x^iy^j}{\sum_{(i,j)\in D} h_{ij}x^iy^j}$$

and for a solution $p(x,y)/q(x,y) \in [N/D]_E^f$ we want to find that p/q and g/h are equivalent. In other words, that

$$(ph - gq)(x,y) = 0$$

It is clear that this is the case if the general multivariate Padé approximation problem to f has a unique solution, because then both p/q and g/h satisfy the approximation conditions (6). If the solution is non-unique we can get into trouble because of the non-unicity of the irreducible form of the Padé approximant as pointed out in the previous section. A solution of the form

$$\frac{2-\alpha-(1+2\alpha)x+2\alpha y}{\alpha-2+\alpha x+(1-\alpha)y}$$

has 3 different irreducible forms:

$$\begin{aligned}
\alpha = -1: &\quad -1 \\
\alpha = 0: &\quad \frac{2-x}{-2+y} \\
\alpha = 1: &\quad \frac{1-3x+2y}{-1+x}
\end{aligned}$$

These irreducible forms cannot all together coincide with g/h. In the context of the consistency property the relations $k \geq n$ and $\ell \geq m$ in theorem 2.2 have to be interpreted as $N_k \supseteq N$ and $D_\ell \supseteq D$ where N and D respectively index the numerator and denominator polynomial of the irreducible $f(x,y)$.

3.3. Covariance properties.

Chisholm, Hughes Jones, Graves-Morris and others who studied the equation lattice definition extensively, emphasized that the index sets should be chosen so as to maximize the number of desirable properties for the multivariate Padé approximant. Covariance properties fully rely on the inclusion property (5c) of the equation lattice E and hence apply to most of the definitions in this group (some definitions drop condition (5c) in order to obtain computational advantages [20, 53, 45]).

THEOREM 3.1 [14]:
Let $p/q \in [N/D]_E^f$ which is the general multivariate Padé approximant to $f(x,y)$ as defined above and let $g(x,y) = (1/f)(x,y)$ with $f(0,0) \neq 0$. Then

$$q/p \in [D/N]_E^g$$

If also $D \subset E$ then the split into (7a) and (7b) can also be done for $[D/N]_E^g$. If we study the homographic function covariance of the general multivariate Padé approximant, we cannot consider denominator index sets D different from the numerator index set N, just like in the univariate case. Indeed, when transforming the function f into the function $\tilde{f} = (af+b)/(cf+d)$, a general Padé approximant $p/q \in [N/D]_E^f$ transforms into

$$\frac{ap+bq}{cp+dq}(x,y) = \frac{\sum_{(i,j)\in N\cup D} \tilde{a}_{ij}x^iy^j}{\sum_{(i,j)\in N\cup D} \tilde{b}_{ij}x^iy^j}$$

which can not necessarily be written in the form $\tilde{p}/\tilde{q} \in [N/D]_E^{\tilde{f}}$.

THEOREM 3.2 [52, 43]:
Let $p/q \in [N/N]_E^f$ which is the general multivariate Padé approximant to $f(x,y)$ and let $\tilde{f} = (af + b)/(cf+d)$, then

$$\tilde{p}/\tilde{q} \in [N/N]_E^{\tilde{f}}$$

with

$$\begin{aligned}
\tilde{p}(x,y) =& ap(x,y) + bq(x,y) \\
\tilde{q}(x,y) =& cp(x,y) + dq(x,y)
\end{aligned}$$

3.4. Projection property.

The general equation lattice definition usually also reduces to the univariate definition as a special case. The projection property below is valid for the multivariate Padé approximants defined in [27, 43, 52] but in general not for those defined in [48, 38, 46, 53, 20, 21].

We introduce, for a finite subset S of $I\!N^2$, the notations

$$S_x = \max\{i|(i,j) \in S\}$$
$$S_y = \max\{j|(i,j) \in S\}$$

and the two particular projection operators

$$\mathcal{P}_x(f) = f(x,0)$$
$$\mathcal{P}_y(f) = f(0,y)$$

THEOREM 3.3 [46]:
If $E_x \geq N_x + D_x$, then the univariate Padé approximant $[N_x/D_x]^{\mathcal{P}_x(f)}$ equals the irreducible form of $\mathcal{P}_x([N/D]_E^f)$. If $E_y \geq N_y + D_y$, then the univariate Padé approximant $[N_y/D_y]^{\mathcal{P}_y(f)}$ equals the irreducible form of $\mathcal{P}_y([N/D]_E^f)$.

Even when the conditions stated in Theorem 3.3 are satisfied, it should be noted that two general multivariate Padé approximants of different degree and order, $[N^1/D^1]_{E^1}^f$ and $[N^2/D^2]_{E^2}^f$, can be projected on a univariate Padé approximant of the same order. This easily follows from the fact that for different index sets S^1 and S^2, one can have $S_x^1 = S_x^2$.

3.5. Algorithms.

Concerning the algorithmic aspect, we have to make a distinction between on one hand algorithms for the very general case, where the index sets N, D and E can be chosen freely as long as (5) is satisfied (with a possible exception for (5c)), and on the other hand algorithms that apply to specific N, D and E such as the ones given in [27, 43, 20].

Let us first treat the latter. In [43] N and D are rectangular,

$$N = ([0,n_1] \times [0,n_2]) \cap I\!N^2 \qquad D = ([0,m_1] \times [0,m_2]) \cap I\!N^2$$

while the construction of E, which we do not repeat here, depends on the relation of n_1, n_2, m_1 and m_2 with respect to $\min(n_1, n_2)$ and $\min(m_1, m_2)$. The logic of the construction can be understood in terms of the so-called prong method for the computation of the approximants [44]. The i^{th} prong is defined as the vector

$$B_i = (b_{i+1,i}, b_{i+2,i}, \ldots, b_{m_1,i}, b_{i,i+1}, b_{i,i+2}, \ldots, b_{i,m_2}, b_{i,i})$$

where the b_{ij} are the denominator coefficients of the multivariate Padé approximant. Here b_{00} is already normalized to be 1 and we assume that the homogeneous system of linear equations (7b) has maximal rank. Calculating B_0 is then equivalent to calculating Padé approximants to $f(x,0)$ and $f(0,y)$, and it turns out that the computation of B_i only requires the values of B_k for $k = 0, \ldots, i-1$. In short, the prong method reduces the computation of the b_{ij} to solving a linear system with a block lower triangular coefficient matrix.

In [20] for instance the sets N, D and E do not satisfy (5c) since they are chosen as

$$N = ([0,n_1] \times [0,n_2]) \cap I\!N^2 \qquad D = ([0,m_1] \times [0,m_2]) \cap I\!N^2$$
$$E = N \cup \left[([n_1+1, n_1+m_1+1] \times [n_2+1, n_2+m_2+1]) \cap I\!N^2\right] \setminus (n_1+m_1+1, n_2+m_2+1)$$

This choice has the drawback that the order of approximation is not higher than in the case of polynomial approximation of $f(x,y)$ by $p(x,y)$ indexed by N. But it has the advantage that ϵ-like and qd-like algorithms can be developed for these approximants and hence that they are easily computable.

With respect to the former we refer the reader to [30] where the well-known ϵ-algorithm for Padé approximants is generalized to the calculation of general multivariate Padé approximants $[N/D]_E^f$. The essential idea behind the ϵ-algorithm is to start off with a sequence of polynomial approximants for f and to rationalize these approximants by slightly increasing the degree of the denominator one step at the time until one reaches $[n/m]^f$. This idea is preserved in [30] but with a slightly harder rationalization process than in the ϵ-algorithm because of the generality of the approximant.

A similar generalization exists for the qd-algorithm that allows one to obtain Padé approximants $[n/m]^f$ in continued fraction form. Here the general multivariate Padé approximants $[N/D]_E^f$ are still obtained as convergents of an ordinary continued fraction [31], but the rhombus rules to compute the partial numerators and denominators in the continued fraction are more complicated as a consequence of the general formulation of the approximation problem.

4. Homogeneous Padé approximants.

4.1. Definition.

In order to avoid any confusion about the role of the degrees n and m, we switch to the use of ν and μ in the discussion of the homogeneous case. For the definition of the homogeneous multivariate Padé approximant $[\nu/\mu]^f$ we introduce the notations

$$\begin{aligned} A_\ell(x,y) &= \sum_{i+j=\nu\mu+\ell} a_{ij}x^iy^j \qquad & \ell &= 0,\ldots,\nu \\ B_\ell(x,y) &= \sum_{i+j=\nu\mu+\ell} b_{ij}x^iy^j \qquad & \ell &= 0,\ldots,\mu \\ C_\ell(x,y) &= \sum_{i+j=\ell} c_{ij}x^iy^j \qquad & \ell &= 0,1,2\ldots \end{aligned}$$

For chosen ν and μ the polynomials

$$\begin{aligned} p(x,y) &= \sum_{\ell=0}^{\nu} A_\ell(x,y) \\ q(x,y) &= \sum_{\ell=0}^{\mu} B_\ell(x,y) \end{aligned}$$

are then computed from the conditions

$$(fq-p)(x,y) = \sum_{i+j\geq\nu\mu+\nu+\mu+1} d_{ij}x^iy^j \tag{8}$$

which can be rewritten as

$$\begin{cases} C_0(x,y)B_0(x,y) = A_0(x,y) \\ C_1(x,y)B_0(x,y) + C_0(x,y)B_1(x,y) = A_1(x,y) \\ \quad\vdots \\ C_\nu(x,y)B_0(x,y) + \ldots + C_{\nu-\mu}(x,y)B_\mu(x,y) = A_\nu(x,y) \end{cases} \tag{9a}$$

$$\begin{cases} C_{\nu+1}(x,y)B_0(x,y) + \ldots + C_{\nu+1-\mu}(x,y)B_\mu(x,y) \equiv 0 \\ \vdots \\ C_{\nu+\mu}(x,y)B_0(x,y) + \ldots + C_\nu(x,y)B_\mu(x,y) \equiv 0 \end{cases} \tag{9b}$$

where $C_\ell(x,y) \equiv 0$ if $\ell < 0$. This is exactly the system of defining equations (1) for univariate Padé approximants if the term $c_\ell x^\ell$ in the univariate definition is substituted by

$$C_\ell(x,y) = \sum_{i+j=\ell} c_{ij} x^i y^j \qquad \ell = 0, 1, 2 \ldots$$

A simple count of unknowns and conditions in (9) shows that in the bivariate case the number of equations is one less than the number of unknowns, just like in the univariate case. But in the general multivariate case the system (9) is overdetermined. Nevertheless it has been proven that a nontrivial solution also exists in the multivariate case [71, pp. 60–62]. It is therefore unnecessary to consider the linear conditions (8) in a least squares sense. This inherent dependence among the homogeneous Padé approximation conditions is still not fully understood and may lead to new developments. The homogeneous analogue of equation (7c) is discussed in [82].

For the homogeneous Padé approximants we can also prove a multivariate analogon of theorem 2.1.

THEOREM 4.1 [71, pp. 14]:
If p_1 and q_1 as well as p_2 and q_2 satisfy condition (9), then

$$(p_1 q_2)(x,y) = (p_2 q_1)(x,y)$$

The homogeneous multivariate Padé approximant $[\nu/\mu]^f$ for $f(x,y)$ can then be defined as the unique irreducible form of a solution $p(x,y)/q(x,y)$ of (9). Several suitable normalizations are possible. This unicity of the irreducible form is a distinctive characteristic of the homogeneous approach.

4.2. Consistency property.

For the homogeneous Padé approximants the consistency property also holds.

THEOREM 4.2 [71, pp. 65]:
For an irreducible rational function

$$f(x,y) = \frac{g(x,y)}{h(x,y)} = \frac{\sum\limits_{i+j=0}^{\nu} g_{ij} x^i y^j}{\sum\limits_{i+j=0}^{\mu} h_{ij} x^i y^j}$$

with $h_{00} = 1$, the homogeneous Padé approximant $[\kappa/\lambda]^f$ with $\kappa \geq \nu$ and $\lambda \geq \mu$ is given by $[\kappa/\lambda]^f = [\nu/\mu]^f = g/h$.

This consistency property is an important advantage of the homogeneous multivariate Padé approximants over the general multivariate Padé approximants. Or rather, the unicity of the irreducible form, on which the property is based, is a big advantage.

4.3. Covariance properties.

Because of the close similarity between the homogeneous multivariate Padé approximants and the well-known univariate Padé approximants, a lot of classical properties remain valid.

THEOREM 4.3 [71, pp. 24]:
Let $f(0,0) \neq 0$ and let $[\nu/\mu]^f = p_{\nu,\mu}/q_{\nu,\mu}$. Then

$$[\mu/\nu]^{1/f} = \frac{q_{\nu,\mu}/f(0,0)}{p_{\nu,\mu}/f(0,0)}$$

THEOREM 4.4 [71, pp. 25]:
Let a, b, c and d be complex numbers with $cf(0,0) + d \neq 0$ and let $[\nu/\nu]^f = p_{\nu,\nu}/q_{\nu,\nu}$. Then

$$[\nu/\nu]^{(af+b)/(cf+d)} = \frac{(ap_{\nu,\nu} + bq_{\nu,\nu})/(c\,f(0,0)+d)}{(cp_{\nu,\nu} + dq_{\nu,\nu})/(c\,f(0,0)+d)}$$

4.4. Projection property.

The homogeneous multivariate Padé approximants satisfy a stronger projection property than the one given in theorem 3.3.

THEOREM 4.5 [65]:
Let $(x, y) = (\lambda_1 z, \lambda_2 z)$ with $\lambda_i \in \mathbb{C}$ for $i = 1, 2$ and let $f_{\lambda_1,\lambda_2}(z) = f(\lambda_1 z, \lambda_2 z)$. Then

$$[\nu/\mu]^{f_{\lambda_1,\lambda_2}}(z) = [\nu/\mu]^f(\lambda_1 z, \lambda_2 z)$$

4.5. Algorithms.

It must be clear from the above that the homogeneous Padé approximants are very similar to the univariate Padé approximants. This is even more apparent from the list of valid algorithms below. For instance, the univariate ϵ-algorithm can immediately be applied to the computation of homogeneous Padé approximants [68], after substituting the univariate starting values

$$\sum_{i=0}^{n} c_i z^i$$

by the multivariate homogeneous expressions

$$\sum_{i+j=0}^{n} c_{ij} x^i y^j$$

In the same way, the univariate qd-algorithm remains valid [72], after replacing the starting values

$$\frac{c_{n+1} z^{n+1}}{c_n z^n}$$

which in the univariate case reduce to $(c_{n+1}/c_n)z$, by

$$\frac{C_{n+1}(x,y)}{C_n(x,y)} = \frac{\sum_{i+j=0}^{n+1} c_{ij} x^i y^j}{\sum_{i+j=0}^{n} c_{ij} x^i y^j}$$

Both algorithms above are only useful in the multivariate case if one wants to compute the value of a homogeneous Padé approximant. They do not deliver the coefficients in numerator and denominator of the approximant unless polynomial arithmetic is used in the ϵ- and qd-algorithms. The denominator coefficients b_{ij} can also be obtained from (9b) as follows. When we introduce the vectors and matrices (^{T} denotes the transpose)

$$B_{\nu\mu+i} = (b_{\nu\mu+i,0}, b_{\nu\mu+i-1,1}, \ldots, b_{1,\nu\mu+i-1}, b_{0,\nu\mu+i})^T$$

$$H_{ij} = \begin{pmatrix} c_{i0} & 0 & \ldots & 0 \\ \vdots & \ddots & & \vdots \\ c_{0i} & & & 0 \\ 0 & & & c_{i0} \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & c_{0i} \end{pmatrix} \qquad (i+j+1) \times (j+1) \text{ matrix}$$

the system of homogeneous equations (9b) defining the denominator coefficients b_{ij} looks like

$$\begin{pmatrix} H_{\nu+1,\nu\mu} & H_{\nu,\nu\mu+1} & \ldots & H_{\nu+1-\mu,\nu\mu+\mu} \\ H_{\nu+2,\nu\mu} & \ldots & & \\ \vdots & & & \vdots \\ H_{\nu+\mu,\nu\mu} & \ldots & & H_{\nu,\nu\mu+\mu} \end{pmatrix} \begin{pmatrix} B_{\nu\mu} \\ \vdots \\ B_{\nu\mu+\mu} \end{pmatrix}$$

Since this coefficient matrix has displacement rank at most $\mu + 2$ [71, pp. 66–67], meaning that it is near-Toeplitz, the system can be solved in only $O[(\mu+2)(\#D-1)^2]$ operations where $\#D - 1$ is the system size, if we put, as in (4)

$$D = \{i + j \mid \nu\mu \le i + j \le \nu\mu + \mu\}$$

.

5. Symbolic-numeric Padé approximants.

5.1. Definition.

Given a bivariate function $f(x, y)$ in the form (2), this function is treated as a univariate function, with the remaining variables being parameters. The main publications on this approach are [84, 85, 89] with [90] being a variation of [84, 85, 89]. Let us therefore recall this definition. The series development (2) is rewritten as

$$\sum_{i=0}^{\infty} c_i(y) x^i \tag{10}$$

with

$$c_i(y) = \sum_{j=0}^{\infty} c_{ij} y^j$$

Of course the role of the variables x and y can be interchanged, with x being the parameter and y being the remaining variable. This is a drawback rather than an advantage, because the approximation process does not treat the variables of f in a symmetrical way. A univariate Padé approximant for (10) can be computed in the usual way. We denote by $[n/m]_x^f$ the irreducible form of p_x/q_x where

$$\begin{aligned} p_x(x,y) &= \sum_{i=0}^{n} a_i(y)x^i \\ q_x(x,y) &= \sum_{i=0}^{m} b_i(y)x^i \\ (fq_x - p_x)(x,y) &= \sum_{i\geq n+m+1} d_i(y)x^i \end{aligned} \tag{11}$$

If we develop $[n/m]_x^f$ into a series

$$[n/m]_x^f(x,y) = \sum_{i=0}^{\infty} \gamma_i(x)y^i \tag{12}$$

then the functions $\gamma_i(x)$ are rational functions of x. A univariate Padé approximant for (12) can be computed in the same way. Let us denote by $[\tilde{n}/\tilde{m}]_y^f \circ [n/m]_x^f$ the irreducible form of p_y/q_y where

$$\begin{aligned} p_y(x,y) &= \sum_{i=0}^{\tilde{n}} \tilde{a}_i(x)y^i \\ q_y(x,y) &= \sum_{i=0}^{\tilde{m}} \tilde{b}_i(x)y^i \\ ([n/m]_x^f q_y - p_y)(x,y) &= \sum_{i\geq n+m+1} \tilde{d}_i(x)y^i \end{aligned} \tag{13}$$

It is clear that

$$[\tilde{n}/\tilde{m}]_y^f \circ [n/m]_x^f \neq [n/m]_x^f \circ [\tilde{n}/\tilde{m}]_y^f$$

Also the numerator and denominator of $[\tilde{n}/\tilde{m}]_y^f \circ [n/m]_x^f$ are respectively of degree $\tilde{n}$ and $\tilde{m}$ in y, but in general not of degree n and m in x anymore. When rephrased in terms of the equation lattice approach, we can say that

$$(fq_y - p_y)(x,y) = \sum_{(i,j)\in \mathbb{N}^2\setminus E} d_{ij}x^iy^j$$

with

$$E \supset ([0, n+m] \cap \mathbb{N}) \times ([0, \tilde{n}+\tilde{m}] \cap \mathbb{N})$$

In [85] the author also explains the connection with the approximation technique using branched continued fractions: here too the variables are dealt with in an unsymmetrical way and a univariate approximation step is used per variable while the remaining variables at that time are treated as parameters.

5.2. Properties.

Not too many properties of symbolic-numeric Padé approximants can be found in the literature. The following covariance and projection property have been given in respectively [85] and [90].

Theorem 5.1:
Let $f(x,y) = g(\tilde{z})$ with $\tilde{z} = x(cy+d), c \neq 0$ or $\tilde{z} = (ax+b)y, a \neq 0$. Then

$$[n/m]_x^f \circ [n+m/0]_y^f(x,y) = [n/m]^g(\tilde{z})$$

and

$$[n/m]_y^f \circ [n+m/0]_x^f(x,y) = [n/m]^g(\tilde{z})$$

Theorem 5.2:
For $x = 0$ or $y = 0$

$$\begin{aligned} [n/m]_x^f \circ [\tilde{n}/\tilde{m}]_y^f(0,y) =& [\tilde{n}/\tilde{m}]^f(y) \\ =& [\tilde{n}/\tilde{m}]_y^f \circ [n/m]_x^f(0,y) \\ [\tilde{n}/\tilde{m}]_y^f \circ [n/m]_x^f(x,0) =& [n/m]^f(x) \\ =& [n/m]_x^f \circ [\tilde{n}/\tilde{m}]_y^f(x,0) \end{aligned}$$

A consistency property cannot be given in general. Not all rational functions can be reconstructed by this type of approximation process. The given rational function has to be of the same form as its approximant $[\tilde{n}/\tilde{m}]_y^f \circ [n/m]_x^f$. And we have already pointed out that the numerator and denominator degree of this symbolic-numeric Padé approximant do not depend in a straightforward way on the parameters $n, m, \tilde{n}$ and $\tilde{m}$.

5.3. Algorithms.

The approximants defined above can of course be computed using standard univariate techniques. The main difference is that one has to deal with the data in a symbolic way. The univariate algorithm also has to be called as many times as the number of variables.

In [91] a slight variation of the above definition is proposed, allowing the use of non-symbolic algorithms: the denominator coefficients of the symbolic-numeric Padé approximant are computed directly from linear systems arising from the univariate subproblems.

6. Convergence results and numerical example.

When discussing convergence results of Padé approximants, one compares a sequence of approximants in the Padé table with the given function f. The selection of an appropriate sequence is possible using information about f. If a univariate function has a fixed number of poles in a certain region, it makes sense to consider a sequence of Padé approximants with fixed denominator degree, in other words a column in the table. If the function has a countable number of singularities, it is wiser to consider a diagonal or ray in the table. We shall now list a number of famous theorems that have also been generalized to the multivariate case.

In comparing the results we have to distinguish between 'uniform' convergence, which is an overall convergence with the Chebyshev-norm of the error tending to zero, and convergence in 'measure' or 'capacity', where one has convergence except for an area of disruption of which the location is usually unknown but of which the size can be made arbitrarily small. In this text we restrict ourselves to the notion of measure only, to avoid the discussion of multivariate generalizations of the notion of capacity later on. If more general results hold however, we shall refer the reader to the literature. We denote

$$\begin{aligned}
B(0,r) &= \{z \in \mathbb{C} : |z| < r\} \\
\overline{B}(0,r) &= \{z \in \mathbb{C} : |z| \leq r\} \\
B((0,0),r) &= \{(x,y) \in \mathbb{C}^2 : ||(x,y)|| < r\} \\
\overline{B}((0,0),r) &= \{(x,y) \in \mathbb{C}^2 : ||(x,y)|| \leq r\} \\
B((0,0);r_1,r_2) &= \{(x,y) \in \mathbb{C}^2 : |x| < r_1, |y| < r_2\} \\
\overline{B}((0,0);r_1,r_2) &= \{(x,y) \in \mathbb{C}^2 : |x| \leq r_1, |y| \leq r_2\}
\end{aligned}$$

and Λ_4 for the Lebesgue-measure in $\mathbb{C}^2$.

THEOREM 6.1 [8]:
Let the function $f(z)$ be meromorphic in $\overline{B}(0,r)$ with poles z_i in $B(0,r)$ of total multiplicity M. Then the sequence $\{[n/M]^f\}_{n\in\mathbb{N}}$ converges uniformly to f on compact subsets of $\overline{B}(0,r) \setminus \{z_i\}$ with z_i attracting zeros of the Padé denominator according to its multiplicity:

$$\lim_{n\to\infty} ||[n/M]^f - f||_K = 0 \qquad \textit{compact } K \subset \overline{B}(0,r) \setminus \{z_1, \ldots, z_M\}$$

THEOREM 6.2A [12]:
Let the function $f(z)$ be meromorphic in $\overline{B}(0,r)$ with poles z_i in $B(0,r)$ of total multiplicity M. Then the sequence $\{n_k/m_k\}_{k\in\mathbb{N}}$ with $m_k \geq M$ and $\lim_{k\to\infty} n_k/m_k = \infty$, converges in $B(0,r)$ in measure to f:

$$\forall \epsilon, \delta,\ \exists \kappa : |f(z) - [n_k/m_k]^f(z)| < \epsilon \textit{ for } k \geq \kappa \textit{ and } z \in B(0,r) \setminus \mathcal{E} \textit{ with } \Lambda_2(\mathcal{E}) < \delta$$

THEOREM 6.2B [46]:
Let the function $f(z)$ be meromorphic in $\overline{B}(0,r)$ with poles z_i in $B(0,r)$ of total multiplicity M. Then for $m \geq M$ there exist points $\zeta_1, \ldots, \zeta_{m-M}$ in $\mathbb{C}$ and there exists a subsequence of $\{[n/m]^f\}_{n\in\mathbb{N}}$ that is uniformly convergent on compact subsets of $B(0,r) \setminus (\{z_1, \ldots, z_M\} \cup \{\zeta_1, \ldots, \zeta_{m-M}\})$.

In short, when one is approximating a meromorphic function and one chooses the denominator degree of the approximant equal to the total number of poles within a distance of at most r, then one can expect uniform convergence of the Padé approximants in that region. If one chooses the denominator degree slightly too large, then one can only expect convergence in measure (and capacity [2]) or one can only expect a subsequence to converge uniformly.

THEOREM 6.3 [9, 11]:
Let the function f be analytic in $\mathbb{C}$ except for a countable number of isolated poles and essential singularities. Then the sequence $\{[n_k/m_k]^f\}_{k\in\mathbb{N}}$ with $\lambda < n_k/m_k < 1/\lambda$ for $0 < \lambda < 1$, converges to f in measure on compact sets:

$$\forall \epsilon, r > 0 : \Lambda_2\left(\{z \in \overline{B}(0,r) : |f(z) - [n_k/m_k]^f(z)| \geq \epsilon\}\right) \longrightarrow_{k\to\infty} 0$$

This last theorem is a simpler version of the original one which proves convergence in capacity. Since the number of singularities of f is now countable, one has to let the denominator degree increase unboundedly, and hence column sequences make an inappropriate choice. The exceptional set that is excluded from the region of convergence is for instance caused by unwanted pole-zero combinations in the Padé approximant.

6.1. Results for the equation lattice and the symbolic-numeric approach.

The uniform convergence theorem of de Montessus de Ballore has been generalized both for the equation lattice and the symbolic-numeric approach. For each of the definitions that are a special case of the very general definition (6) or the symbolic-numeric approach (11) and (13), different versions of what can be called a multivariate de Montessus de Ballore theorem can be found in [39, 46, 51, 33, 86, 90]. We restrict ourselves here to outlining the differences between these theorems and the reason for the existence of these differences. This contributes much more to the understanding of multivariate Padé approximation than a dry list of results. In all generalizations locally uniform convergence is obtained for a function $f(x,y)$ that is such that there exists a multivariate polynomial $s(x,y)$ (not series, hence of finite degree) such that $(fs)(x,y)$ is analytic in some neighbourhood of the origin. The theorems differ in the specification of the additional constraints, which have to safeguard you from getting close to troublesome points in $\mathbb{C}^2$. These troublesome points are a direct consequence of the way the Padé approximant is defined: in other words, a direct consequence of the numerator and denominator polynomials (or for that matter the index sets) of the Padé approximant. The following two cases illustrate this.

Let us first look at the equation lattice uniform convergence theorems by presenting a typical case. In [39], the polynomial $s(x,y)$ describing the singularities of $f(x,y)$ in a polydisc $B((0,0);r_1,r_2)$ is of the form

$$s(x,y) = \sum_{i=0}^{m_1}\sum_{j=0}^{m_2} b_{ij}x^i y^j$$

Hence it is natural to look at approximants $[N/D]_E$ with denominator $q(x,y)$ indexed by $D = ([0,m_1] \times [0,m_2]) \cap \mathbb{N}^2$. For the numerator the index set $N = ([0,n_1] \times [0,n_2]) \cap \mathbb{N}^2$ is chosen. Without going into details we also mention that

$$E \supset \{(i,0) | 0 \le i \le n_1 + m_1\} \cup \{(0,j) | 0 \le j \le n_2 + m_2\}$$

This choice for E enforces a projection property of the multivariate Padé approximant on the x-axis and the y-axis. Consequently the poles of $s(x,0)$ and $s(0,y)$, and especially their moduli, play a crucial role in the formulation of the theorem and the specification of the region of uniform convergence. As a result the formulation of the conditions under which the theorem holds, are rather technical and depend very much on the form of the approximant.

If we look at [86] a similar conclusion holds. In this approach the variable y is treated as a parameter in the first step of the approximation process. Consequently the theorem includes the condition that $s(x,y_0)$ should not have multiple x-roots unless for a finite number of y_0-values. This is because $s(x,y)$, which is of the form

$$s(x,y) = x^M + \sum_{k=0}^{M-1} b_k(y)x^k$$

is also rewritten in the form

$$s(x,y) = (x - \sigma_1(y)) \dots (x - \sigma_M(y))$$

Moreover, in a certain polydisc $B((0,0);r_1,r_2)$, the polynomial $s(x,y_0)$ should have exactly M x-roots for every y_0. The number M determines the denominator degree of $[n/M]_x^f$ which is constructed in the first step of the symbolic-numeric Padé approximation process. In addition, for $x_0 \in B(0,r_1)$, the polynomial $s(x_0,y)$ should have at most a certain number of y-roots in $B(0,r_2)$. This last number determines the denominator degree $\tilde{M}$ of $[\tilde{n}/\tilde{M}]_y^f \circ [n/M]_x^f$ which is being constructed in the second step. It is clear that again the final formulation of the theorem depends very much on the nature of the computed numerator and denominator polynomial of the Padé approximant.

Whereas we will also have to stay away from a small set of troublesome points in the homogeneous Padé approximation approach, this set will clearly be unavoidable and will not depend so much on the construction of the homogeneous Padé approximant. It will contain points in $\mathbb{C}^2$ that are exceptional, even while the very strong projection property given in Theorem 4.5 is valid. For more information we refer to the discussion below.

Convergence results in measure have not been obtained for the symbolic-numeric approach. The oldest result for the equation lattice approach is only valid for a specific choice of the numerator, denominator and equation index sets N, D and E:

$$\begin{aligned} N :=& N(k) = \{(i,j) | 0 \leq i+j \leq k\} \\ D =& N \\ E :=& E(k) \supset \{(i,j) | 0 \leq i+j \leq \lfloor \sqrt{2}k \rfloor + 1\} \end{aligned}$$

THEOREM 6.4 [38]:
Let the function $f(x,y)$ be analytic in $\mathbb{C}^2 \setminus \mathcal{G}$ where the analytic set $\mathcal{G} = \{(x,y) \in \mathbb{C}^2 : g(x,y) = 0\}$ with $g(x,y)$ entire. Then the sequence $\{[N(k)/N(k)]^f_{E(k)}\}_{k \in \mathbb{N}}$ converges on compact sets in measure to f.

More recently results have been formulated for the general definition (6). We now respectively give generalizations of the Zinn-Justin convergence theorem and the Nuttall–Pommerenke convergence in capacity. After each theorem we translate the conditions to the univariate case, so that it becomes clear why those conditions are natural generalizations of the ones in the univariate theorems. This is also very helpful because the conditions under which the theorems hold, are again rather technical. For sequences of general index sets $\{N_k\}_{k \in \mathbb{N}}, \{D_k\}_{k \in \mathbb{N}}, \{E_k\}_{k \in \mathbb{N}}$ and an index set M we denote by

$$\begin{aligned} N_k * M =& \{(i,j) : i = i_1 + i_2, j = j_1 + j_2, (i_1, j_1) \in N_k, (i_2, j_2) \in M\} \\ i_{D_k} =& \max\{i : (i,j) \in D_k\} \\ j_{D_k} =& \max\{j : (i,j) \in D_k\} \\ \partial D_k =& \max\{i_{D_k}, j_{D_k}\} \\ \omega E_k =& \min\{i + j : (i,j) \in \mathbb{N}^2 \setminus E_k\} \end{aligned}$$

THEOREM 6.5 [34]:
Let the function $f(x,y)$ be meromorphic in the polydisc $B((0,0); r_1, r_2)$ in the sense that there exists a multivariate polynomial

$$s(x,y) = \sum_{(i,j) \in M} s_{ij} x^i y^j$$

such that fs is holomorphic in that polydisc. For N_k, D_k and E_k satisfying

$$\begin{aligned} & N_k * M \subset E_k \\ & \lim_{k \to \infty} \omega E_k / \partial D_k = \infty \end{aligned}$$

the sequence of approximants $\{[N_k/D_k]_{E_k}\}_{k \in \mathbb{N}}$ converges in $B((0,0); r_1, r_2)$ in measure to f.

In the univariate case the sets N_k, D_k and E_k equal

$$\begin{aligned} N_k =& \{0, \ldots, n_k\} \\ D_k =& \{0, \ldots, m_k\} \\ E_k =& \{0, \ldots, n_k + m_k\} \end{aligned}$$

Hence

$$\begin{aligned}\partial D_k =& m_k \\ \omega E_k =& n_k + m_k + 1\end{aligned}$$

and the conditions in the above theorem amount to

$$N_k * M \subset E_k \Longleftrightarrow n_k + M \le n_k + m_k \Longleftrightarrow m_k \ge M$$

$$\lim_{k\to\infty} \frac{\omega E_k}{\partial D_k} = \infty \Longleftrightarrow \lim_{k\to\infty} \frac{n_k}{m_k} = \infty$$

which are the standard univariate conditions.

THEOREM 6.6 [34]:
Let $f(x,y)$ be such that for each ρ there exists a polynomial $s_\rho(x,y)$ such that $(fs_\rho)(x,y)$ is analytic in the polydisc $B((0,0);\rho,\rho)$. Let $\ell_k = \max\{\partial N_k, \partial D_k\}$ and

$$C_{\lfloor k \rfloor} = \{(i,j) \in I\!N^2 : 0 \le i \le \lfloor k \rfloor, 0 \le j \le \lfloor k \rfloor\}$$

For N_k, D_k and E_k satisfying

$$\begin{aligned}&\lim_{k\to\infty} \ell_k = \infty \\ &N_k * C_{\lfloor \lambda \ell_k \rfloor} \subset E_k \\ &D_k * C_{\lfloor \lambda \ell_k \rfloor} \subset E_k\end{aligned}$$

with $0 < \lambda < 1$, the sequence of approximants $\{[N_k/D_k]_{E_k}\}_{k\in I\!N}$ converges on compact sets in measure to f.

In the univariate case these conditions translate to the following:

$$\begin{aligned}\ell_k &= \max\{n_k, m_k\} \\ N_k * C_{\lfloor \lambda \ell_k \rfloor} \subset E_k &\Longleftrightarrow n_k + \lambda \ell_k \le n_k + m_k \Longrightarrow \lambda n_k \le m_k \\ D_k * C_{\lfloor \lambda \ell_k \rfloor} \subset E_k &\Longleftrightarrow m_k + \lambda \ell_k \le n_k + m_k \Longrightarrow \lambda m_k \le n_k\end{aligned}$$

These last conditions amount to

$$\lambda \le n_k/m_k \le 1/\lambda$$

The last two theorems also hold if we replace the notion of measure by capacity as detailed in [34].

6.2. Convergence results for homogeneous Padé approximants.

Thanks to the projection property mentioned in Theorem 4.5 the following convergence results were obtained. We do not cite them in their most general form. For this the reader is referred to the original reference. We introduce for (λ_1, λ_2) in $\mathbb{C}^2$:

$$\begin{aligned}B_{(\lambda_1,\lambda_2)}(0,r) =& \{z \in \mathbb{C} : ||(\lambda_1 z, \lambda_2 z)|| < r\} \\ f_{(\lambda_1,\lambda_2)}(z) =& f(\lambda_1 z, \lambda_2 z)\end{aligned}$$

THEOREM 6.7 [75]:
Let the function $f(x,y)$ be meromorphic in the ball $B((0,0),r)$ in the sense that there exists a polynomial $s(x,y)$ of homogeneous degree M such that fs is holomorphic in $B((0,0),r)$. If we denote

$$\begin{aligned}W =& \{(\lambda_1,\lambda_2) : ||(\lambda_1,\lambda_2)|| = 1 \textit{ and} \\ & f_{(\lambda_1,\lambda_2)} \textit{ has less than } M \textit{ poles in } B_{(\lambda_1,\lambda_2)}(0,r)\} \\ \mathcal{S} =& \{(x,y) \in \mathbb{C}^2 : s(x,y) = 0\} \\ \mathcal{E} =& \{(\lambda_1 z, \lambda_2 z) : (\lambda_1,\lambda_2) \in W, z \in \mathbb{C}\}\end{aligned}$$

then the sequence $\{[\nu/M]_H^f\}_{\nu\in\mathbb{N}}$ converges uniformly on compact subsets of $B((0,0),r)$ not intersecting $\mathcal{E}\cup\mathcal{S}$. Outside W each zero of $s_{(\lambda_1,\lambda_2)}(z)$ attracts zeros of the projected Padé denominator according to its multiplicity.

The set W denotes the set of exceptional directions, meaning that for (λ_1,λ_2) in W the univariate convergence theorem of de Montessus de Ballore applies to a column different from that for the vectors outside W: for all vectors (λ_1,λ_2) outside W one has to consider the M-th column. Note than one does not have convergence in $(0,0)$, the point at which the series development for f was given, because it is always contained in $\mathcal{E}\cup\mathcal{S}$! The following example due to Lubinsky illustrates very well why this is the best one can expect.
Let h be an entire function and define

$$f(x,y)=h(x)+h(y)+\frac{y-x}{x-1}$$

It is easy to see that $f_{(\lambda_1,\lambda_2)}$ has poles of total multiplicity 1 unless $\lambda_1=\lambda_2$ or $\lambda_1=0$. So

$$\mathcal{E}=\{(x,x)\mid x\in\mathbb{C}\}\cup\{(0,x)\mid x\in\mathbb{C}\}$$

For $\lambda_1=\lambda_2$,

$$f_{(\lambda_1,\lambda_2)}(z)=2h(\lambda_1 z)$$

and for $\lambda_1=0$,

$$f_{(\lambda_1,\lambda_2)}(z)=h(\lambda_2 z)+h(0)-\lambda_2 z$$

Thus the $\{[\nu/1]_H^f\}_{\nu\in\mathbb{N}}$ Padé sequence to f will not converge locally uniformly in any neighbourhood of any point of $\mathcal{E}$ provided the ordinary Padé approximants to h do not converge locally uniformly in any neighbourhood of any point of $\mathbb{C}$. There are many well-known examples of such entire functions h, going back at least to [10].

THEOREM 6.8A [75]:
Let the function $f(x,y)$ be meromorphic in the ball $B((0,0),r)$ in the sense that there exists a polynomial $s(x,y)$ of homogeneous degree M such that fs is holomorphic in $B((0,0),r)$. Then for $m\ge M$ the sequence $\{[\nu/m]_H^f\}_{\nu\in\mathbb{N}}$ converges in $B((0,0),r)$ in measure to f.

This theorem nicely generalizes the univariate result obtained by Zinn-Justin while the next theorem generalizes the univariate result of Karlsson and Wallin. Both deal with a denominator choice that is again slightly too large. For the next convergence result we assume that the sequence $\{[\nu/m]_H^f\}_{\nu\in\mathbb{N}}$ with fixed $m\ge M$ has an infinite number of elements $[\nu_h/m]_H^f$ that are not singular at the origin. Remember that due to the term $\nu\mu$ in (8), the denominator of an approximant $[\nu/\mu]_H^f$ may evaluate to zero at the origin. We denote this subsequence of well-defined entries by $\{[\nu_h/m]_H^f\}_{h\in\mathbb{N}}$.

THEOREM 6.8B [67]:
Let the function $f(x,y)$ be meromorphic in the ball $B((0,0),r)$ in the sense that there exists a polynomial $s(x,y)$ of homogeneous degree M such that fs is holomorphic in that ball. Then for $m\ge M$ there exists an analytic set $\mathcal{T}\supset\mathcal{S}$ and there exists a subsequence of $\{[\nu_h/m]_H^f\}_{h\in\mathbb{N}}$ that converges uniformly to f on compact subsets of $B((0,0),r)\setminus\mathcal{T}$.

Let us now turn to a generalization of the Nuttall–Pommerenke result, for homogeneous Padé approximants.

Theorem 6.9 [74]:
Let the function $f(x,y)$ be analytic in $\mathbb{C}^2 \setminus \mathcal{G}$ where the analytic set $\mathcal{G} = \{(x,y) \in \mathbb{C}^2 : g(x,y) = 0\}$ with g entire. Then the sequence $\{[\nu_k/\mu_k]_H^f\}_{k\in\mathbb{N}}$ with $\lambda < \nu_k/\mu_k < 1/\lambda$ for $0 < \lambda < 1$, converges on compact sets in measure to f.

6.3. Numerical example.

When comparing the different definitions for multivariate Padé approximant on numerical examples, it is easy to come to the following two conclusions:

(1) the adaptiveness of the equation lattice approach to the data $c_{i_1,\ldots,i_p}$ is a clear advantage; if one of the variables in the multivariate problem is dominant, then one can adapt both the rational function given by (3) and (4) and the system of defining equations given by (6) to the situation;
(2) the homogeneous approach has such a strong projection property, delivering the well-known univariate Padé approximant on every one-dimensional slice $\{(\lambda_1 z, \ldots, \lambda_p z) | z \in \mathbb{C}\}$, that the quality of the approximation is comparable to that of the best tailor-made general multivariate Padé approximant.

The most important testproblem is probably the Beta-function because it has been used by almost all researchers active in multivariate Padé approximation theory and hence it allows an easy comparison of numerical results between the different generalizations. This function is defined by

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

where Γ is the Gamma-function. The Beta-function is meromorphic in $\mathbb{C}^2$ with poles at $x = -k$ and $y = -k$ and zeros at $y = -x - k$, all for $k = 0, 1, 2, \ldots$. The interested reader is referred to:

(1) [37, pp. 292] for the description of an optimally tailored general multivariate Padé approximant $[N/D]_E^{B(x,y)}$ to the Beta-function;
(2) [47] for numerical results using other general multivariate Padé approximants;
(3) [71, pp. 89–93] for the numerical calculation of homogeneous multivariate Padé approximants to the Beta-function;
(4) [41] for numerical results using the equation lattice definition given in [43] for the Beta-function;
(5) [5] for numerical results using interpolatory branched continued fractions for the Beta-function.