Acta Applicandae Mathematicae **33**: 165–194, 1993.

The Multipoint Padé Table and General Recurrences for Rational Interpolation

MARTIN H. GUTKNECHT
Interdisciplinary Project Center for Supercomputing, ETH Zurich, ETH-Zentrum, CH-8092 Zurich, Switzerland. email: mhg@ips.ethz.ch

(Received: December 1993)

Abstract. We first review briefly the Newton–Padé approximation problem and the analogous problem with additional interpolation conditions at infinity, which we call multipoint Padé approximation problem. General recurrence formulas for the Newton–Padé table combine either two pairs of Newton–Padé forms or one such pair and a pair of multipoint Padé forms. We show that, likewise, certain general recurrences for the multipoint Padé table compose two pairs of multipoint Padé forms to get a new pair of multipoint Padé forms. We also discuss the possibility of superfast, i.e., $O(n \log^2 n)$ algorithms for certain rational interpolation problems.

Mathematics Subject Classifications (1991): 41A21, 41A05.

Key words: Newton–Padé approximation, multipoint Padé approximation, rational interpolation, recurrences, fast algorithms.

0. Introduction

The rational interpolation problem, which can be formulated in various ways, is readily reduced to a particularly structured linear system of equations. For a rational interpolant with denominator degree n, one can derive, for example, a homogeneous $n \times (n+1)$ system with a divided-difference matrix (see [38], [45], [23]), which contains higher divided differences in each column, or a homogeneous system with a Löwner matrix [13], which contains first divided differences, or yet other structured systems. But, in any case, while the application of the Gauss algorithm would require $O(n^3)$ operations, there are recursive $O(n^2)$ algorithms for solving these structured systems. Most of the classical methods (see [36], [45], [22] for surveys), such as the determination of a Thiele continued fraction, only work under genericness assumptions, whereas the Kronecker algorithm [33], [45], [2] and the generalizations of the Thacher–Tukey algorithm and the Thiele fraction proposed by Graves-Morris and Hopkins [22], [21] and by Werner [46], [47] are generally applicable in exact arithmetic, i.e., are so-called reliable. The latter is also true for the more recent recurrences of Antoulas and Anderson [3] and the algorithms of Gutknecht [25], Van Barel and Bultheel [42], [44], and Beckermann [4], which are all closely related to Werner's algorithm. However, if subject to roundoff, all these methods are numerically unstable in cases of near-singularity of some linear systems that belong to rational interpolants of lower degrees constructed during the recursive process.

Therefore, in [23] we described new, more general recurrences that allow us to avoid ill-conditioned intermediate interpolants. These recurrences are modeled after those of the Cabay–Meleshko algorithm [10], [37] for computing Padé approximants, which was shown to be (weakly) stable. But whereas the recurrences of Cabay and Meleshko follow a diagonal of the Padé table, those in [23] are for the Newton–Padé table and allow us to move in any direction in which $m + n$ does not decrease. In particular, we can follow a pair of adjacent rows or antidiagonals of the table. While the general recurrence for a diagonal can be understood as combining two pairs of Newton–Padé forms, other paths require to combine a pair Newton–Padé forms with a pair of multipoint Padé forms. The latter represent multipoint Padé approximants, by which we mean rational interpolants with some interpolation data prescribed at infinity. (We need to point out that, in contrast to our terminology introduced in [26] and [23], the notion multipoint Padé approximant is sometimes used as equivalent to Newton–Padé approximant.) Here, we review in condensed form the background of this result, present general recurrences for both Newton–Padé forms and multipoint Padé forms, draw a number of conclusions, and discuss the merits and pitfalls of this approach, including the possibility of superfast $O(n \log^2 n)$ algorithms. For more information on the background we refer the reader to [23] and the many references cited there.

Such general recurrences are the basis of look-ahead algorithms for rational and Padé approximation, and for the solution of the corresponding structured linear systems. The notion of a look-ahead procedure came up in connection with the unsymmetric Lanczos algorithm [39], [27], [24], [14], which is closely related to a diagonal recurrence in the Padé table and the fast solution of a Hankel system. A look-ahead algorithm requires, in addition to a general recurrence, a look-ahead strategy, i.e., a rule for choosing the step size such that the recurrence remains stable. Although our theoretical results give some hints on how to choose such a strategy here, this topic is not yet discussed in detail. For the diagonal Padé recursion, a complete analysis was given in [10]. For various forms of Padé row recursions, a partial analysis is given in [28], [30], [29]. These papers also reference several articles on fast Hankel and Toeplitz solvers that make use of the look-ahead philosophy in a somewhat different way. The case of rational interpolation still needs further investigation. However, some first numerical results are presented in [7]. There are extensions of this approach to vector-valued and matrix-valued Padé approximation and rational interpolation; see, e.g., [5], [8], [9], [15], [41], [43], [40].

As in [25] and [23] we assume here for notational convenience that the interpolation data are given in Newton series form. This dispenses us of the need of using much more complicated notation for multiple interpolation points and the corresponding derivatives. However, we need to point out that our formulas, in particular the recurrences, are readily translated into formulas for different representations of the data. We do not consider the formal Newton series as a prime computational tool. In particular, if all interpolation points are distinct and well apart, it is much easier to represent the data as a set of function values than to transform these into divided differences. Only if some of the points are clustered, it is preferable to represent the data corresponding to these points by a Newton series, i.e., by a Newton interpolation polynomial if there are only finitely many data, as one can assume in practice.

When using directly the function values one is free to choose the order in which the data are taken into consideration. With an appropriate strategy, related to pivoting in Gaussian elimination, Graves-Morris improved the numerical stability of the Thiele–Thacher–Tukey–Werner interpolation process considerably; see [21], [20], [47]. This idea could be combined with our approach, but, for simplicity, is here left out of consideration.

1. Newton–Padé Approximation

In some applications of rational interpolation the data are given as quotients of values. Treating the numerators and denominators of these quotients separately does not make the problem much more difficult. In fact, then it becomes more symmetric. It is also no extra effort to allow both the data and the points to be complex. For notational convenience we assume that the data for the interpolation points z_j $(j = 0, 1, \ldots)$ are given as a quotient of two (finite or infinite) formal Newton series:

$$h(z) := -\frac{f(z)}{g(z)}, \quad f(z) := \sum_{k=0}^{\infty} \phi_k t_k(z), \quad g(z) := \sum_{k=0}^{\infty} \gamma_k t_k(z), \tag{1.1}$$

where

$$t_0(z) :\equiv 1, \quad t_k(z) := \prod_{j=0}^{k-1} (z - z_j), \quad k = 1, 2, \ldots. \tag{1.2}$$

The two formal Newton series must be relatively prime in the sense that they do not vanish simultaneously at any z_k. While a single formal Newton series can be used to represent a series of finite complex function values and derivatives at arbitrary points in the complex plane, we can represent also poles of h with the representation (1.1). It is well known that the coefficients ϕ_k and γ_k in (1.1) are the kth *divided differences* at $z_0, \ldots, z_k$ of f and g, respectively.

We let $\mathcal{N}_Z$ be the set of all formal Newton series for the fixed sequence $Z := \{z_j\}_{j=0}^{\infty}$ of interpolation points. $\mathcal{N}_Z$ becomes a commutative algebra if addition, scalar multiplication, and multiplication are defined pointwise (using the Leibniz rule for the derivatives of the product) [45]. By $O(t_l)$ we denote any element of $\mathcal{N}_Z$ that has zeros at $z_0, \ldots z_{l-1}$, multiple ones being understood as zeros of formal derivatives. In other words, $g = O(t_l)$ if and only if the Newton series of g starts with the t_l-term. If $l \le 0$, the symbol $O(t_l)$ is understood as a void condition.

$\mathcal{P}_m$ is the space of complex *polynomials* p of exact degree $\partial p \le m$, and $\mathcal{R}_{m,n}$ is the set of *rational functions* $r = p/q$ with $p \in \mathcal{P}_m$, $q \in \mathcal{P}_n$, $q \ne 0$. If $r = p/q$ with relatively prime polynomials p and q of exact degrees ∂p and ∂q, r is said to be of *exact type* $(\partial p, \partial q)$ and to have the *defect*

$$\delta := \delta_{m,n} := \min\{m - \partial p, n - \partial q\} \tag{1.3}$$

in $\mathcal{R}_{m,n}$. If $\delta > 0$, r is called *degenerate* in $\mathcal{R}_{m,n}$. In particular, if $p = 0$, $\partial p := -\infty$, and hence $r = 0$ has exact type $(-\infty, 0)$ and defect n in $\mathcal{R}_{m,n}$.

We augment $\mathcal{R}_{m,n}$ by the constant ∞, which can be represented as $1/0$ and has exact type $(0, -\infty)$ and defect m in $\mathcal{R}_{m,n}$. Let $\overline{\mathcal{R}}_{m,n} := \mathcal{R}_{m,n} \cup \{\infty\}$. Note that then $r \in \overline{\mathcal{R}}_{m,n}$ if and only if $1/r \in \overline{\mathcal{R}}_{n,m}$.

Given a relatively prime pair $(f, g) \in \mathcal{N}_Z \times \mathcal{N}_Z$, the double sequence of *linearized rational interpolation problems* or *Newton–Padé approximation problems* for (f, g) consists in finding, for each pair $(m, n) \in \mathbf{N} \times \mathbf{N}$ the pairs $(p, q) \in \mathcal{P}_m \times \mathcal{P}_n$ for which

$$g(z)p(z) + f(z)q(z) = O(t_{m+n+1}(z)). \tag{1.4}$$

Any such pair $(p, q) \in \mathcal{P}_m \times \mathcal{P}_n$ is called an (m, n) *Newton–Padé form* (NPF) of (f, g). The associated rational function $r_{m,n} := p/q$ is an (m, n) *Newton–Padé approximant* (NPA) of (f, g) [17]. When writing p/q we do not assume that p and q are relatively prime, but we think of p/q as the rational function obtained after cancellation of common factors of p and q. Therefore p/q has to be distinguished from (p, q). Actually, there may be no relatively prime solution pair.

The pair (p, q), which only depends on the $(m+n)$th partial sums of the Newton series f and g, is clearly never uniquely determined. But it is easy to prove (see, e.g., [23]) that for fixed m and n all solutions (p, q) of (1.4) yield the same rational function p/q, i.e., the (m, n) NPA $r_{m,n}$ is well defined. We call it a *true* rational interpolant if $r_{m,n}$ interpolates $h = -f/g$ at the first $m + n + 1$ data points, which is the case unless there exists no relatively prime pair (p, q).

The set of all NPFs for a particular interpolation problem is characterized in the following fundamental theorem, which is essentially due to Maehly and Witzgall [34]. A simple proof is given in [23].

THEOREM 1. *The general solution* $(p, q) \in \mathcal{P}_m \times \mathcal{P}_n$ *of* (1.4) *is*

$$(p, q) = (\hat{p}_{m,n}\, s_{m,n}\, w,\ \hat{q}_{m,n}\, s_{m,n}\, w), \tag{1.5}$$

where $\hat{p}_{m,n}$, $\hat{q}_{m,n}$, *and* $s_{m,n}$ *are polynomials that are uniquely determined up to a common scaling factor of* $\hat{p}_{m,n}$ *and* $\hat{q}_{m,n}$, *while* $w \in \mathcal{P}_{\delta_{m,n} - \partial s_{m,n}}$ *is arbitrary.* $\hat{p}_{m,n}$ *and* $\hat{q}_{m,n}$ *are relatively prime,* $s_{m,n}$ *is a monic divisor of* t_{m+n+1} *of degree* $\partial s_{m,n} \le \delta_{m,n}$, *and* $\delta_{m,n}$ *is the defect of* $r_{m,n} = \hat{p}_{m,n}/\hat{q}_{m,n}$ *in* $\overline{\mathcal{R}}_{m,n}$. *The zeros of* $s_{m,n}$ *are the points* z_k *where the pair* $(\hat{p}_{m,n}, \hat{q}_{m,n})$ *does not satisfy the interpolation conditions.*

From this theorem we learn in particular that if $\partial s_{m,n} > 0$, then $(\hat{p}_{m,n}, \hat{q}_{m,n})$ is not a solution of the linearized problem, and, hence, $r_{m,n} = \hat{p}_{m,n}/\hat{q}_{m,n}$ is not a true interpolant of h. The points where $r_{m,n}$ does not interpolate up to the required order are called *unattainable*. These points are exactly the zeros of $s_{m,n}$. For this reason $s_{m,n}$ is called the *deficiency polynomial*. For a summary of the connections between the nonlinear rational interpolation problem and the Newton–Padé approximation problem, see [23] and the references cited there.

Among the other easy consequences of Theorem 1 there is the following Characterization Theorem.

THEOREM 2 (NPA Characterization Theorem). *The function $r \in \overline{\mathcal{R}}_{m,n}$ with defect δ is the (m, n) NPA of $h = -f/g$ if and only if r interpolates h in at least $m + n + 1 - \delta$ data points.*

Associated with a double sequence of Newton–Padé problems defined by (1.4) there is a *Newton–Padé table* which covers a quarter of the (m, n)-plane and contains as its (m, n) entry the NPA $r_{m,n}$. For the recursive computation of rational interpolants one often follows some path in the Newton–Padé table. For example, this path may be a diagonal, an antidiagonal, or a staircase. (By a 'diagonal' we mean the main diagonal or any of its infinitely many upper or lower codiagonals.) However, in such algorithms special measures need to be taken if two interpolants are identical on the path. Hence, it is important to understand in which situations identical entries can occur. Moreover, for floating-point computations it is equally important to deal appropriately with nearly identical interpolants.

Sets of more than one equal entry are called *singular blocks* of the Newton–Padé table. The block structure of the Newton–Padé table was first investigated by Claessens [11]. The following formulation of the Block Structure Theorem was given in [25]. Additional results from [25] omitted here concern the characterization of the blocks of the zero and the infinity function, and of the infinite block occurring when h is rational; see Figure 1.

THEOREM 3 (NP Block Structure Theorem). *Let $r = \hat{p}/\hat{q} \neq 0, \infty$ be an NPA of $h = -f/g$ of exact type $(\partial\hat{p}, \partial\hat{q})$. Then the block of r in the Newton–Padé table of h is a (finite or infinite) union of squares whose upper left corners lie at or below the location $(\partial\hat{p}, \partial\hat{q})$ on the diagonal passing through this location.*

It was also shown in [25] that this statement is sharp, i.e., any block satisfying the above description can actually occur. In general, a singular block in the Newton–Padé table needs not even be connected. It may consist of a finite or even infinite number of disjoint components, see Figure 1. There is an easy interpretation for the form of these blocks. Every antidiagonal of the table is associated with an interpolation condition: e.g., the one through $m + n = k = \text{const}$ is associated with the data $(z_k, h(z_k))$ if $z_k \neq z_i$ $(\forall i < k)$. If r is an NPA whose block contains at least one point of the above antidiagonal, then this block becomes broader or

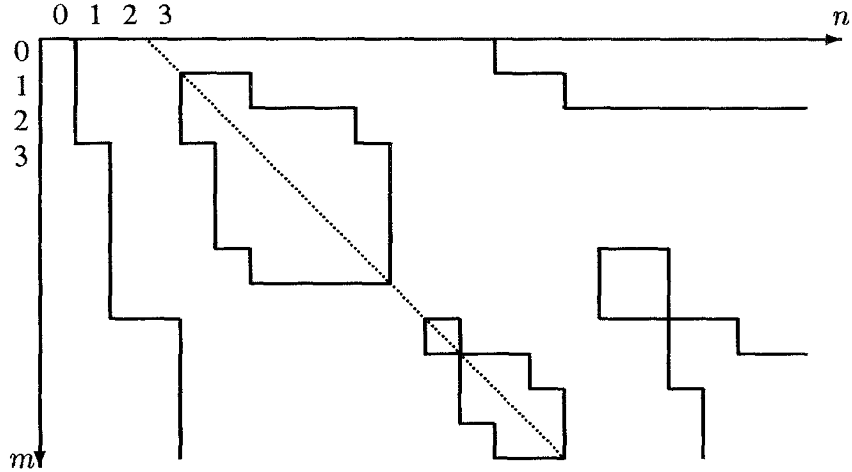

Fig. 1. A Newton–Padé table with four singular blocks: at left a block of the constant ∞, at the top right a block of the constant 0, in the middle a singular block consisting of three components (with the common diagonal dotted), and at the bottom right the infinite block of h (assumed to be rational).

narrower at this antidiagonal depending on whether the new linearized interpolation condition is satisfied by r or not. For a more detailed discussion of this see [25]. Once the block gets narrower, r is certainly no longer a true interpolant. But it may even happen that an NPA r is for no index pair (m, n) a true interpolant. In fact this is true if and only if the $(\partial\hat{p}, \partial\hat{q})$ entry does not belong to the block of r. For a particular $h = -f/g$ and r, this situation can be avoided by reordering the data. However, this means considering another formal Newton series representation and, hence, another Newton–Padé table.

2. Multipoint Padé Approximation

While the rational interpolation problem of the previous section includes the case of prescribed poles at finite interpolation points, it does not allow any interpolation condition at infinity, except the prescription of the degrees m and n, which determine whether the NPA r has a pole, a zero, or a finite nonzero value at infinity. In this section we include interpolation conditions at infinity.

Again, we want to consider a double sequence of approximants. However, since conditions at infinity fix the difference between numerator and denominator degree, one of the two parameters has now a new meaning: the first, μ, determines the bias between using information at infinity and at finite interpolation points; the second, n, still denotes the nominal denominator degree of the rational function, i.e., the maximum number of finite poles (with account of their multiplicity).

The data at the finite interpolation points are again assumed to be given in the Newton series form (1.1), although, as mentioned before, we suggest to use in practice the direct representation by function values and derivatives unless the interpolation points are clustered. Assuming that the function does not have a multiple zero at ∞, we prescribe its behavior there by a quotient of two formal power series in z^{-1},

$$\hat{h}(z) := -\frac{\hat{f}(z)}{\hat{g}(z)}, \quad \hat{f}(z) := \sum_{k=-\infty}^{\iota} \hat{\phi}_k z^k, \quad \hat{g}(z) := \sum_{k=-\infty}^{0} \hat{\gamma}_k z^k, \tag{2.1}$$

with $\iota \geq -1$ and $\hat{\phi}_\iota \neq 0$ or $\hat{\gamma}_0 \neq 0$. Of course, we could choose $\iota = 0$ or $\iota = -1$ without restricting generality, but we prefer to keep the parameter ι (iota) since we will encounter data of the above form for different values of ι, and, a priori, we will only know that $\hat{\phi}_\iota$ or $\hat{\gamma}_0$ does not vanish, but not which one.

It has been shown in [23] and will be further discussed below that for those general recurrences for the Newton–Padé table that do not just follow a diagonal, one has to solve rational interpolation problems whose data are of the form (1.1), (2.1).

We let $\mathcal{L}$ denote the linear space of *formal Laurent series* $y(z) = \sum_{k=-\infty}^{\infty} \eta_k z^k$ with complex coefficients. We need also subspaces of $\mathcal{L}$ that are of the form

$$\begin{aligned} \mathcal{L}_{l:m} &:= \{y \in \mathcal{L};\ \eta_k = 0 \text{ if } k < l \text{ or } k > m\}, \\ \mathcal{L}_l &:= \mathcal{L}_{l:\infty} = \{y \in \mathcal{L};\ \eta_k = 0 \text{ if } k < l\}, \\ \mathcal{L}_m^\star &:= \mathcal{L}_{-\infty:m} = \{y \in \mathcal{L};\ \eta_k = 0 \text{ if } k > m\}. \end{aligned}$$

In particular, $\mathcal{L}_0$ and $\mathcal{L}_0^\star$ are the sets of *formal power series* in z and $1/z$, respectively. When $y \in \mathcal{L}_m^\star$, we write $y(z) = O_-(z^m)$, and if additionally $y \notin \mathcal{L}_{m-1}^\star$, we may express this as $\partial y = m$. The set of all polynomials (i.e., of formal power series with finitely many terms) is denoted by $\mathcal{P}$.

As in [23] we consider now the following double sequence of *multipoint Padé approximation problems*: given $(f, g) \in \mathcal{N}_Z \times \mathcal{N}_Z$, as in (1.1), and $(\hat{f}, \hat{g}) \in \mathcal{L}_\iota^\star \times \mathcal{L}_0^\star$ as in (2.1), let, for each index pair $[\mu; n] \in \mathbf{Z} \times \mathbf{N}$, the nonnegative integer $m := m(\mu; n)$ be defined by

$$m := \max\{\iota - \mu - 1, \iota + n, \mu\}, \tag{2.2}$$

and determine the pairs

$$(u, v) \in \begin{cases} \mathcal{L}_{\mu+n+1:\iota+n} \times \mathcal{P}_n = \mathcal{L}_{\iota+n-m:\iota+n} \times \mathcal{P}_n \\ \qquad \text{if } \mu \leq -n-1, \\ \mathcal{P}_{\iota+n} \times \mathcal{P}_n = \mathcal{P}_m \times \mathcal{P}_n \\ \qquad \text{if } -n-1 \leq \mu \leq \iota + n, \\ \mathcal{P}_\mu \times \mathcal{P}_n = \mathcal{P}_m \times \mathcal{P}_n \\ \qquad \text{if } \mu \geq \iota + n, \end{cases} \tag{2.3}$$

which satisfy the conditions $(u, v) \neq (0, 0)$ and

$$\hat{g}(z)u(z) + \hat{f}(z)v(z) = O_-(z^\mu), \tag{2.4a}$$

$$g(z)u(z) + f(z)v(z) = O(t_{\mu+n+1}(z)), \tag{2.4b}$$

where $t_j(z) \equiv 1$ if $j \leq 0$. Any such pair (u, v) is here called a $[\mu; n]$ *multipoint Padé form (MPF)* of $(\hat{f}, \hat{g}; f, g)$, and

$$r_{\mu;n}(z) := \frac{u(z)}{v(z)} \tag{2.5}$$

is referred to as the corresponding $[\mu; n]$ *multipoint Padé approximant (MPA)*.

It was shown in [23] that in analogy to the situation in Padé and Newton–Padé approximation, for fixed μ and n, the $[\mu; n]$ MPFs (u, v) are equivalent in the sense that they yield a uniquely determined rational function u/v. Hence, also the MPA $r_{\mu;n}$ is well defined.

Let us discuss the three cases in (2.3). In the first one, where $n \leq -\mu-1 = m-\iota$, (2.4b) is void and (2.4a) implies $\iota+n-\mu = m+n+1$ conditions at $z = \infty$ for the pair (u, v) with a total of $m + n + 2$ parameters. Hence, we have a one-point Padé approximation problem at infinity. In the third case, $\iota+n \leq \mu = m$, (2.4a) is void, while (2.4b) is identical to the condition (1.4) for an (m, n) NPA. Finally, in the second case, both (2.4a) and (2.4b) are in effect, and we have a *proper* multipoint Padé problem with $\iota + n - \mu$ conditions at $z = \infty$ and $\mu + n + 1$ conditions at finite interpolation points. Hence, there is a total of $\iota + 2n + 1$ conditions for $(u, v) \in \mathcal{P}_{\iota+n} \times \mathcal{P}_n$.

In summary, the three cases in (2.3) have the following interpretation:

$\mu \leq -n-1$:	Padé problem at $z = \infty$,	$m = \iota - \mu - 1$,
$-n-1 < \mu < \iota+n$:	proper multipoint Padé problem,	$m = \iota + n$,
$\iota+n \leq \mu$:	Newton–Padé problem,	$m = \mu$.

The MPAs for a data set $(\hat{f}, \hat{g}; f, g)$ can be collected in a *multipoint Padé table* $\{r_{\mu;n}\ ;\ [\mu; n] \in \mathbf{Z} \times \mathbf{N}\}$ that covers a half-plane. We let the n-axis point to the right and the μ-axis point to the bottom. If $\iota = 0$, the proper MPAs lie in a 90° sector with horizontal axis of symmetry between $\mu = -1$ and $\mu = 0$; cf. Figure 2. Above this sector the table contains Padé approximants at infinity, below there are NPAs. If $\iota > 0$, the lower border of this sector lies farther down. If $\iota = -1$, it is farther up by one entry. In this case $r_{-1;0}(z) \equiv 0$, as is seen from (2.3)–(2.4b).

Note that n is the nominal denominator degree of $r_{\mu;n}$, but μ is *not* the numerator degree, except in the NPA sector of the table, where $\mu = m$. The number of parameters in the numerator is always $m + 1$.

Again, one can describe the set of solutions of the defining equations (2.2)–(2.4b). This result was given in [26]. We cite it from [23], where it was reformulated.

THEOREM 1'. *In the case* $-n-1 < \mu < \iota+n$ *of proper multipoint Padé approximation the general form of the MPFs* (u, v) *defined by* (2.2)–(2.4b) *is*

$$(u, v) = (\hat{u}_{\mu;n}\, s_{\mu;n}\, w,\ \hat{v}_{\mu;n}\, s_{\mu;n}\, w), \tag{2.6}$$

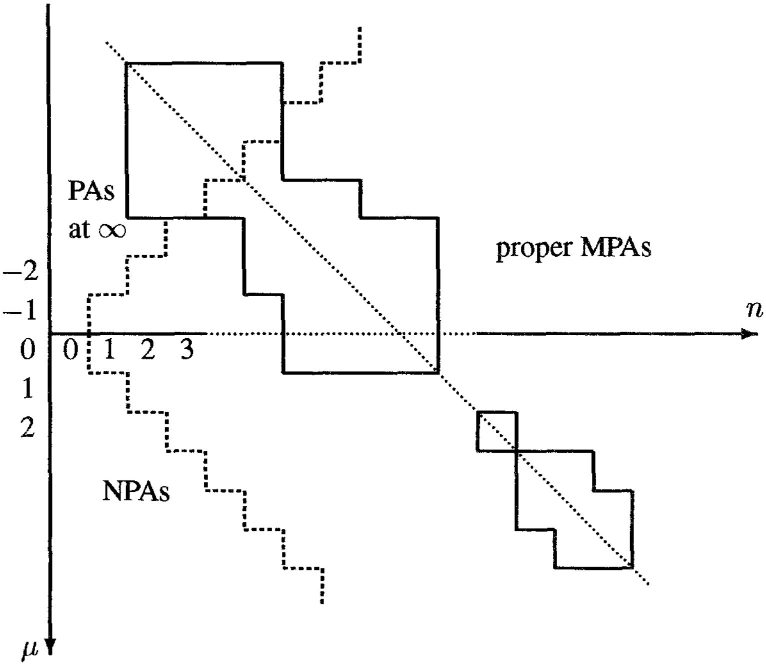

Fig. 2. The case $\iota = 0$ of a multipoint Padé table with its 90° sector of proper multipoint Padé approximants and the two adjoining sectors of Padé approximants at infinity and Newton–Padé approximants, which are separated by dashed staircases. Also shown is a typical singular block with its dotted diagonal.

where $\hat{u}_{\mu;n}, \hat{v}_{\mu;n}$ are up to scaling uniquely determined relatively prime polynomials, and $s_{\mu;n}$ is a monic divisor of $t_{\mu+n+1}$ of degree $\partial s_{\mu;n} \leq \min\{\delta, \nu\}$, with $\delta := \delta_{\mu;n}$ (≥ 0) being the defect of $r_{\mu;n} := \hat{u}_{\mu;n}/\hat{v}_{\mu;n}$ in $\overline{\mathcal{R}}_{\iota+n,n}$, and

$$\nu := \nu_{\mu;m} := \mu - \partial(\hat{g}\hat{u}_{\mu;n} + \hat{f}\hat{v}_{\mu;n}) \quad (\geq 0) \tag{2.7}$$

being the number of extra (linearized) interpolation conditions fulfilled at ∞ by $(\hat{u}_{\mu;n}, \hat{v}_{\mu;n})$. Finally, w is an arbitrary nonzero polynomial of degree

$$\partial w \leq \min\{\delta, \nu\} - \partial s_{\mu;n}. \tag{2.8}$$

The zeros of the deficiency polynomial $s_{\mu;n}$ are the finite interpolation points z_k (with appropriate multiplicity) at which the pair $(\hat{u}_{\mu;n}, \hat{v}_{\mu;n})$ does not satisfy the interpolation conditions. The deficiency at ∞,

$$\hat{\sigma} := \hat{\sigma}_{\mu;n} := \begin{cases} 0 & \text{if } \nu \geq \delta, \\ \delta - \nu & \text{if } \nu < \delta, \end{cases} \tag{2.9}$$

indicates the number of (linearized) interpolation conditions at ∞ that are not fulfilled by a pair (u, v) of the form (2.6) with $\partial s_{\mu;n} + \partial w = \delta$. It satisfies

$$0 \leq \hat{\sigma} \leq \min\{n - \mu, \delta - \partial s_{\mu;n}\}. \tag{2.10}$$

To avoid the treatment of several cases we restricted Theorem 1' to proper MPFs, i.e., to $-n-1 < \mu < \iota+n$. However, when the appropriate interpretations of void conditions are made, the theorem remains true for $\mu \leq -n-1$ and $\mu \geq \iota+n$. In fact, then one has a one-point Padé approximation problem at ∞ or a Newton–Padé approximation problem, respectively, and one can therefore just refer to Theorem 1.

In view of a general MPA characterization theorem, it is useful to define the defect δ for all three types of MPAs. Assume r is any rational function that can be written as $r = \hat{u}/\hat{v}$ with $(\hat{u}, \hat{v})$ satisfying (2.3) and $\hat{u}$, $\hat{v}$ being relatively prime. In the case $\mu \leq -n-1$, where $m = \iota - \mu - 1$ and thus $\iota + n - m \leq 0$, we call $(\hat{u}, \hat{v})$ relatively prime if the polynomials $z^{\iota+m-n}\hat{u}(z)$ and $\hat{v}(z)$ are relatively prime. We define the $[\mu; n]$-*defect* of r by

$$\delta := \delta_{\mu;n} := \max\{\mu - \partial\hat{u}, \iota + n - \partial\hat{u}, n - \partial\hat{v}\}. \tag{2.11}$$

For proper MPAs (i.e., if $-n-1 \leq \mu \leq \iota+n$), this is just the defect of r in $\overline{\mathcal{R}}_{\iota+n,n}$. For NPAs (i.e., for $\mu \geq \iota+n$, where $m = \mu$), it is the defect of r in $\overline{\mathcal{R}}_{m,n}$. Finally, for PAs at ∞ (i.e., for $\mu \leq -n-1$, where $m = \iota - \mu - 1$), $\delta_{\mu;n}$ is equal to the defect of the function $z \mapsto r(1/z)$ in $\overline{\mathcal{R}}_{m,n}$.

With this definition, the following analogue of Theorem 2 holds [23].

THEOREM 2' (MPA Characterization Theorem). *Let* $[\mu; n] \in \mathbf{Z} \times \mathbf{N}$, *let* (u, v) *satisfy* (2.3), *and let* δ *be the* $[\mu; n]$*-defect of* $r = u/v$. *Then* r *is the* $[\mu; n]$ *MPA if it satisfies at least* $\iota + 2n + 1 - \delta$ *of the* $\iota + 2n + 1$ *nonlinear interpolation conditions implied by* (2.4a)–(2.4b).

If the finite interpolation points z_k all coalesce at $z = 0$, our proper multipoint Padé approximation problem becomes a two-point Padé approximation problem. Then g and f are formal power series, and the term $O(t_{\mu+n+1})$ in (2.4b) becomes $O(z^{\mu+n+1})$. Such two-point Padé approximation problems for two formal power series $\hat{h} := -\hat{f}/\hat{g}$ and $h := -f/g$ in $1/z$ and z, respectively, have been investigated by several authors; see [23] for references and comments. The two-point analog of our multipoint Padé table is called *M-table*. It was shown by Cooper, Magnus, and McCabe [12] that in the M-table all singular blocks are either square or infinite, exactly as in the case of the classical Padé table. The only new feature is the possibility of a new type of an infinite block, a half-plane $\{(\mu, n) \in \mathbf{Z}\times\mathbf{N};\ n \geq n_0\}$. We showed in [23] that also in the multipoint Padé table the typical singular blocks have the same form as in the Newton–Padé table, except that the portion in the Padé approximation sector must be square; see Figure 2. For the formulation of this Block Structure Theorem it is important to have the following characterization for the diagonal of a singular block [23].

LEMMA . *Let* $r = \hat{u}/\hat{v}$ *be an MPA. Assume that* $\hat{u} \neq 0$ *and* $\hat{v} \neq 0$ *are relatively prime. Then those entries in the block of* r *where the* $[\mu; n]$*-defect* δ *satisfies*

$$\hat{g}(z)\hat{u}(z) + \hat{f}(z)\hat{v}(z) \equiv O_{-}(z^{\mu-\delta}) \tag{2.12}$$

lie on a particular diagonal of the multipoint Padé table. The case $\mu - \delta = -\infty$ *is permitted and means that the upper edge of the square block and hence also its diagonal are at infinity.*

We call the constant difference $\mu - \delta$ in (2.12) the *order at* ∞ of $(\hat{u}, \hat{v})$ and denote it by μ_∞. With this definition, the following theorem holds [23].

THEOREM 3' (MP Block Structure Theorem). *Let the data* $(\hat{f}, \hat{g}; f, g)$ *be of the form* (1.1) *and* (2.1), *with* $\iota = 0$. *Let* $r = \hat{u}/\hat{v}$ *be an MPA of* $(\hat{f}, \hat{g}; f, g)$, *and let* μ_∞ *be the order at* ∞ *of* $(\hat{u}, \hat{v})$. *Assume that* $\hat{u} \neq 0$ *and* $\hat{v} \neq 0$ *are relatively prime.*

Then the block of r *in the multipoint Padé table of* $(\hat{f}, \hat{g}; f, g)$ *is a finite or infinite union of squares whose upper left corners lie at or below the location* $[\mu_\infty, \partial\hat{v}]$ *on the diagonal passing through this location. If not empty, the intersection of the block with the sector* $\mu < -n - 1$ *is equal to the intersection of a square with this sector; and the upper left corner of this square is then at* $[\mu_\infty, \partial\hat{v}]$.

As for Theorem 3 the strength of this result is that it is best possible: again, one can prove that any block of the described form can appear; see [23].

3. Recursive Computation of NPAs and MPAs: Basic Ideas

The general recurrence relations for Newton–Padé approximants (NPAs) and multipoint Padé approximants (MPAs) that we want to discuss in the rest of this paper are modeled after the recurrences applied by Cabay and Meleshko [37], [10] for the stable computation of a diagonal sequence of Padé approximants. Under more restrictive assumptions these recurrences were used before by Gragg, Gustavson, Warner, and Yun [18], who modified the recurrences of Brent, Gustavson, and Yun [6]. Here we consider NPAs and MPAs instead of Padé approximants, and a more general type of recurrence that allows us to proceed in other directions too. Whereas some of the other algorithms that have been proposed for singular tables take special small steps to follow the border of the block when encountering a singular one, here the idea is to take a large step crossing the block (as in [25]) or even several blocks. The option of jumping over several blocks allows us to avoid ill-conditioned interpolants that could occur as intermediate results. This is a necessity for a numerically stable algorithm. Moreover, this option allows us to apply recursive doubling and opens up the possibility of superfast $O(n \log^2 n)$ algorithms. Such algorithms were suggested in [6] for staircase Padé sequences, in [8] for diagonal matrix Padé sequences, and in [28] for row Padé sequences. Here we will point out that for a certain class of rational interpolation problems we can still attain the same low complexity.

A crucial tool is the usage of *basic pairs* of NPFs and MPFs. According to Theorems 1 and 1', NPFs and MPFs are determined uniquely up to scaling if and only if in (1.5) and (2.6), respectively, the polynomial w is of degree 0, i.e., a constant. It can be seen that w is a constant if and only if, in the respective table, we are either at a normal entry or at a position at the border of a singular block. In view of the possible form of singular blocks, the latter conditions are fulfilled if the previous or the following entry on the same diagonal is different. Hence, we are going to use pairs of NPFs (and MPFs, respectively) with the property that they are upper left and lower right neighbors, respectively, of each other, but belong to different blocks. This then implies that these NPFs are unique up to scaling, and that the denominators (and also the numerators) are relatively prime except for the possibility of a common zero at an interpolation point, in which case either the corresponding NPAs are not true interpolants or the data function g (or f, respectively) has a zero there also. The second member of these pairs will be called *weakly regular*. (We would call it regular if we knew additionally that the NPA is a true interpolant, but we will not have this knowledge in general.) Instead of basic pairs consisting of a regular NPA and its upper left neighbor, one could also use basic pairs consisting of an NPA (then called weakly column-regular) and its upper neighbor, or an NPA (then called weakly row-regular) and its left neighbor, or an NPA and its lower left neighbor. The last type of basic pairs appears implicitly in the classical Kronecker algorithm [33], [45], [2].

Basic pairs appear at many places in the literature, at least implicitly; see, e.g., [1], [3] (Assumption (3.8)), [4], [42], [44]. In the Padé case they can be traced back to the Euclidean algorithm [6], [18] and are related to Bézoutiants, the Christoffel–Darboux formula, and inversion formulas for Hankel and Toeplitz matrices; see, e.g., [16], [19], [31], [32]. Heinig and Rost [31] call the denominators of the basic pairs *fundamental solutions*. Implicitly, the basic pairs considered here also come up in one version of the methods of Werner [46], [47] and Gutknecht [25] that produces a generalization of the Thiele fraction called a diagonal G-fraction in [25], because its convergents are the distinct entries on a particular diagonal of the Newton–Padé table. In all these situations, one proceeds from one basic pair to the next one, so that the two pairs have a common member. However, the full algorithmic power of basic pairs becomes apparent in conjunction with the mentioned recurrences that allow us to jump over arbitrary many blocks, like in [37], [10], [28], [23].

Antoulas and Anderson, who consider rational interpolants with $\max\{\partial p, \partial q\}$ chosen minimal, give in Lemma 3.2 of [3] a recurrence that is said to allow one to go from any pair of such interpolants to any new interpolant satisfying the union of the previously fulfilled interpolation conditions, plus, possibly, some additional ones. However, the proof of the formula is not given, and, in any case, its application becomes complicated if many new data are added at the same interpolation point, as one then needs to compute higher derivatives of a quotient. In their Theorem 3.9, whose proof is said to serve as a guideline for the one of their Lemma 3.2, the authors turn then to what we call a basic pair and add only one new interpolation condition at a time.

Briefly, the general recurrence from [23] (stated as Theorem 9 below) says that a basic pair of NPFs of types $(m-1, n-1)$ and (m, n) can be updated to a basic pair with types $(m+\kappa-1, n+k-1)$ and $(m+\kappa, n+k)$ satisfying additional

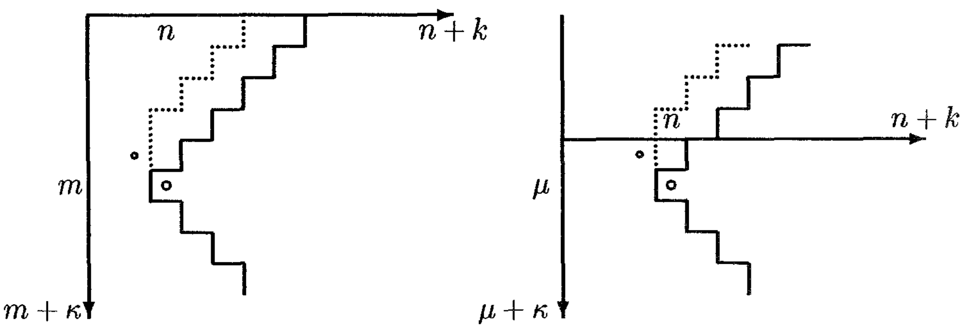

Fig. 3. At left: the index range of the new $(m+\kappa, n+k)$ NPA (solid line) and its upper left neighbor (dotted line) that can be constructed according to Theorem 9 below. At right: the index range of the new $[\mu+\kappa; n+k]$ MPA (solid line) and its upper left neighbor (dotted line) that can be constructed according to Theorem 9'.

$\kappa+k$ interpolation conditions. For such a step of length $\kappa+k$ one needs to compute the so-called residuals of the previous pair, and a pair of $[\kappa; k-1]$ and $[\kappa+1; k]$ MPFs. Here, κ is restricted to $-k \leq \kappa \leq k$. Hence, viewed from the weakly regular (m, n) NPF (p_n, q_n) the newly constructed weakly regular $(m+\kappa, n+k)$ NPF (p_{n+k}, q_{n+k}) lies in the Newton–Padé table in a 90° sector with horizontal axis, see Figure 3. If we want to move instead in a 90° sector with vertical axis we can apply the same theorem, but with (f, g), (p_n, q_n), (m, n) replaced by (g, f), (q_n, p_n), (n, m). In the case where $\kappa = k$ (and also when $\kappa = k-1$), the two MPFs are NPFs. If, moreover, all interpolation points coalesce at 0, the MPFs are just Padé forms, and the recurrence can be seen to be equivalent to the one applied by Cabay and Meleshko [10].

For the multipoint Padé table we present in Theorem 9' below a completely analogous recurrence: starting from a basic pair consisting of a $[\mu-1, n-1]$ and a $[\mu, n]$ MPF for $(\hat{f}, \hat{g}; f, g)$, we can construct a basic pair consisting of a $[\mu+\kappa-1, n+k-1]$ and a $[\mu+\kappa, n+k]$ MPF for $(\hat{f}, \hat{g}; f, g)$ by computing a basic pair consisting of a $[\kappa, k-1]$ and a $[\kappa+1, k]$ MPF for another multipoint Padé problem that is made up of the residuals of the first pair. Again, κ is restricted to $-k \leq \kappa \leq k$, so that we can move in the multipoint Padé table in a 90° sector with horizontal axis whose cusp is at $[\mu; n]$; see Figure 3.

The natural way to apply these recurrences consists in computing NPFs (or MPFs) with slowly increasing n, leaving out values of n where the NPF (or MPF) and its upper left neighbor are not numerically well conditioned or not sufficiently independent. However, there is the other option of applying *recursive doubling*. Assume we want to compute some (n, n) NPA, where n is large. For simplicity, let n be a power of 2 and assume that all the entries on the main diagonal of the Newton–Padé table are distinct, and, hence, regular. According to the NPA recurrence, we can split up the problem and first compute a pair consisting of a $(n/2-1, n/2-1)$

and a $(n/2, n/2)$ NPF and then, for a multipoint Padé problem that depends on these solutions, compute a pair consisting of a $[n/2; n/2-1]$ MPF and a $[n/2+1; n/2]$ MPF. (Actually, here these MPFs can be understood as NPFs.) To fulfill these two different tasks we can again apply, for each one, the same kind of splitting, using the NPA recurrence for the first task and the MPA recurrence for the second one. This can be continued until we get to problems where $n = 1$, and thus the solution becomes trivial. Unfortunately, this application of the *divide-and-conquer* strategy has a weak point: since the data for each subproblem depend on the solution of the previous problem, the subproblems cannot be solved in parallel. Hence, there is no gain from the point of view of parallel computing. The divide-and-conquer strategy is here not applied in its form amenable to parallel computation, but in its original sense, where an enemy is split up and its sections are conquered one after another. Nevertheless there can be some gain here too, since the original problem is split up repeatedly until each subproblem becomes trivial. If the work for the reduction is sufficiently small, the computational cost for solving the original problem may be smaller than when applying the straightforward recurrence. We will return to this question in §6.

4. General Recurrence Relations for NPAs

Assuming that, for the sequence of NPFs we are going to construct, the numerator degree $m = m(n)$ is a function of the denominator degree n, we write $\{(p_n, q_n)\}$ instead of $\{(p_{m(n),n}, q_{m(n),n})\}$. For example, $m(n) = n + l$, with fixed l, for a diagonal sequence, and $m(l) = l$, with fixed l, for a row sequence in the Newton–Padé table. (Note that m has here another meaning than in §2.) Actually, there is normally no point in computing the full sequence $\{(p_n, q_n)\}$; only the subsequence of weakly regular NPFs, or rather, the subsequence of well-conditioned weakly regular NPFs will be constructed, together with their upper left neighbors

$$(\grave{p}_n, \grave{q}_n) := (p_{m(n)-1,n-1}, q_{m(n)-1,n-1}). \tag{4.1}$$

(p_n, q_n) and $(\grave{p}_n, \grave{q}_n)$ can be normalized by letting the sum of the squares of the coefficients of each pair be 1. Other options for normalization follow from Lemma 4 below.

When $n = 0$ or $m(n) = 0$, the upper left neighbor $(\grave{p}_n, \grave{q}_n)$ lies outside the Newton–Padé table, but the following definitions prove to be useful:

$$\begin{array}{llll} \grave{p}_0(z) := t_{m-1}(z), & \grave{q}_0(z) :\equiv 0, & \text{if } n = 0, m > 0; \\ \grave{p}_0(z) :\equiv 0, & \grave{q}_0(z) := t_{n-1}(z), & \text{if } n > 0, m = 0; \\ \grave{p}_0(z) :\equiv 1, & \grave{q}_0(z) :\equiv 0, & \text{if } n = m = 0 \text{ and } q_0 \neq 0; \\ \grave{p}_0(z) :\equiv 0, & \grave{q}_0(z) :\equiv 1, & \text{if } n = m = 0 \text{ and } q_0 = 0; \end{array} \tag{4.2}$$

We could use the fourth definition whenever $p_0 \neq 0$ (and, actually, we should when $|q_0|$ is small). That would leave a free choice between the third and the fourth definition when $p_0 \neq 0$ and $q_0 \neq 0$.

These special definitions allow the following interpretation: we can extend the Newton–Padé table to the left by adding an infinite block of the constant ∞, and we can extend it above by adding an infinite block of the constant 0, as suggested in [19] for the Padé table. The quadrant $m < 0, n < 0$ can be filled with either 0 or ∞, but if $r_{0,0} = 0$ (or ∞), it is preferable to choose ∞ (or 0, respectively).

We call the NPF $(p_n, q_n) := (p_{m(n),n}, q_{m(n),n})$ *weakly regular* if

$$\frac{p_n}{q_n} \neq \frac{\grave{p}_n}{\grave{q}_n}, \quad i.e., \quad \grave{p}_n q_n - p_n \grave{q}_n \neq 0 \in \mathcal{P}. \tag{4.3}$$

Likewise, we call (p_n, q_n) *weakly row-regular* if the $(m(n), n-1)$ and the $(m(n), n)$ NPA differ, and we call it *weakly column-regular* if the $(m(n) - 1, n)$ and the $(m(n), n)$ NPA differ. According to the Block Structure Theorem, (p_n, q_n) is weakly regular if and only if it is weakly row-regular or weakly column-regular.

The *residual* e_n of the NPF (p_n, q_n) is defined by

$$g(z)p_n(z) + f(z)q_n(z) = t_{m(n)+n+1}(z)e_n(z). \tag{4.4}$$

This residual can be written as a formal Newton series for the points $z_{m(n)+n+i}$, $i = 1, 2, \ldots$, i.e., as a series of the polynomials

$$t_0^{(m+n)}(z) :\equiv 1, \quad t_k^{(m+n)}(z) := \prod_{i=1}^{k}(z - z_{m(n)+n+i}), \quad k = 1, 2, \ldots. \tag{4.5}$$

In practice, this Newton series can be replaced by some other representation of the data determining the residual. The definitions of the residual e_n of (p_n, q_n) and the residual $\grave{e}_n$ of $(\grave{p}_n, \grave{q}_n)$ can be summarized in

$$[g \;\; f] \begin{bmatrix} \grave{p}_n & p_n \\ \grave{q}_n & q_n \end{bmatrix} = t_{m(n)+n-1}[\grave{e}_n \;\; \tau_n e_n], \tag{4.6}$$

where

$$\tau_n(z) := (z - z_{m(n)+n-1})(z - z_{m(n)+n}) = t_2^{(m+n-2)}(z). \tag{4.7}$$

When $n = 0$ and $m > 0$, (4.6) still holds with

$$\grave{e}_0 = g, \quad e_0 = (gp_0 + fq_0)/t_{m+1}, \tag{4.8}$$

where q_0 is a nonzero constant that can be normalized to 1. Moreover, (4.3) holds when $q_0 \neq 0$. Analogous statements can be made for the other special cases from (4.2).

The following three lemmas were proved in [28] for $m, n > 0$. Here, we cover also the cases $n = 0, m > 0$ and $n > 0, m = 0$. The latter always follows from the former by symmetry. We leave out the case $m = n = 0$ because it requires special treatment.

LEMMA 4. *Let* (p_n, q_n) *be an* $(m(n), n)$ *NPF (with* $n > 0$ *or* $m(n) > 0$*), and let* $\grave{e}_n$ *be the residual of an* $(m(n)-1, n-1)$ *NPF* $(\grave{p}_n, \grave{q}_n)$*. Then the following statements are equivalent:*

(i) (p_n, q_n) *is weakly regular, i.e.,* (4.3) *holds;*

(ii) $\grave{e}_n(z_{m(n)+n}) \neq 0$ *and* $(p_n(z_{m(n)+n}) \neq 0$ *or* $q_n(z_{m(n)+n}) \neq 0)$*;*

(iii) $\grave{e}_n(z_{m(n)+n-1}) \neq 0$ *and* $(p_n(z_{m(n)+n-1}) \neq 0$ *or* $q_n(z_{m(n)+n-1}) \neq 0)$*.*

Proof for $n = 0$, $m > 0$. In this case all three statements hold if and only if $q_0 \neq 0$. First, $\grave{p}_0 q_0 - p_0 \grave{q}_0 = t_{m-1} q_0$. Second, $\grave{e}_0 = g$, and $g(z_{m-1}) = 0$ or $g(z_m) = 0$ imply by (1.4) that $q_0 = 0$. Conversely, $q_0 = 0$ implies by (1.4) that, for $i = 0$ and $i = 1$ either $g(z_{m-i}) = 0$ or $p(z_{m-i}) = 0$. Q.E.D.

LEMMA 5. *If* (p_n, q_n) *is a weakly regular* $(m(n), n)$ *NPF (with* $n > 0$ *or* $m(n) > 0$*), then*

(i) $\partial q_n = n$ *and* $\partial \grave{p}_n = m(n) - 1$

or

(ii) $\partial p_n = m(n)$ *and* $\partial \grave{q}_n = n - 1$.

Proof for $n = 0$, $m > 0$. From the previous proof we know that $q_0 \neq 0$. But then, in view of (4.2), statement (i) is clearly true. Q.E.D.

LEMMA 6. *Let* (p_n, q_n) *be an* $(m(n), n)$ *NPF (with* $n > 0$ *or* $m(n) > 0$*); then*

$$\det \begin{bmatrix} \grave{p}_n & p_n \\ \grave{q}_n & q_n \end{bmatrix} = \delta_n t_{m(n)+n-1}, \tag{4.9}$$

where δ_n *is the leading coefficient of* $\grave{p}_n q_n - \grave{q}_n p_n \in \mathcal{P}_{m(n)+n-1}$*, which does not vanish if and only if* (p_n, q_n) *is weakly regular.*

Proof for $n = 0$, $m > 0$. The determinant in (4.9) is equal to $t_{m-1} q_0$; hence, (4.9) holds with $\delta_0 = q_0$, which in this case is also the leading coefficient of $\grave{p}_0 q_0 - \grave{q}_0 p_0 \in \mathcal{P}_{m-1}$. Q.E.D.

COROLLARY 7. *If the* $(m(n), n)$ *NPF* (p_n, q_n) *(with* $n > 0$ *or* $m(n) > 0$*) is weakly regular, then any common zero of the polynomials* $\grave{p}_n$ *and* p_n *is a zero of* f *or an unattainable interpolation point of* p_n/q_n*. In particular, if* p_n/q_n *is a true* $(m(n), n)$ *interpolant and* f *does not vanish at any interpolation point, then* $\grave{p}_n$ *and* p_n *are relatively prime. These statement remain true when* $\grave{p}_n$*,* p_n*, and* f *are replaced by* $\grave{q}_n$*,* q_n*, and* g*, respectively.*

Proof. Assume $m, n > 0$ first. By (4.9), if (p_n, q_n) is weakly regular, any common zero of $\grave{p}_n$ and p_n must be a zero of $t_{m(n)+n-1}$. On the other hand, since $\partial w = 0$ in the factorization (1.5) of (p_n, q_n) (because of uniqueness), such a zero must be either a zero of $s_{m,n}$, i.e., an unattainable point, or a zero of $\hat{p}_{m,n}$, in which case it must be a zero of f; see (1.4). The statement about exchanging $\grave{p}_n$, p_n, and f for $\grave{q}_n$, q_n, and g holds by symmetry. If $n = 0$, a zero of $\grave{p}_n = t_{m-1}$ is by (4.4) also a zero of $gp_0 - fq_0$. Hence, a common zero of $\grave{p}_0$ and p_0 must be a zero of f unless $q_0 = 0$, in which case a zero of p_0 is an unattainable point. Q.E.D.

Finally, we need to justify the later usage of the pair of residuals $(e_n, \grave{e}_n)$ as data of a Newton–Padé or multipoint Padé problem.

LEMMA 8. *Assume that* (p_n, q_n) *is a weakly regular* $(m(n), n)$ *NPF* (*with* $n > 0$ *or* $m(n) > 0$), *and let* $Z(n) := \{z_j \in Z\,;\, j \geq m(n) + n - 1\}$. *Then* $\grave{e}_n$ *and* $\tau_n e_n$ *are relatively prime elements of* $\mathcal{N}_{Z(n)}$.

Proof. In [23] we established this lemma under the additional assumption of distinct interpolation points. Here we give a short proof that does not require this assumption. Let $n > 0$, and let (p_n, q_n) be a weakly regular NPF. Then, after normalizing an appropriate coefficient to 1, the other coefficients of these two polynomials and likewise those of $\grave{p}_n$ and $\grave{q}_n$ are determined by a nonsingular linear system that depends on f and g. Clearly, under small perturbations of the data, the coefficients and values of these polynomials and of t_{m+n-1} depend continuously on the data, and, hence, the same is true for $\grave{e}_n$ and $\hat{\tau}_n e_n$. (Note that under such a perturbation t_{m+n-1} remains a factor of $g\grave{q}_n + f\grave{p}_n$, and t_{m+n+1} remains a factor of $gq_n + fp_n$.) By such small perturbations we can make that none of the points $z_j \in Z(n)$ is a zero of t_{m+n-1}. (Actually, we could keep the zeros of t_{m+n-1} fixed; but this is not required.) Applying Cramer's rule to (4.6) and making use of (4.9) we get, after canceling t_{m+n-1},

$$g = \frac{1}{\delta_n} \det \begin{bmatrix} \grave{e}_n & \tau_n e_n \\ \grave{q}_n & q_n \end{bmatrix}, \quad f = \frac{1}{\delta_n} \det \begin{bmatrix} \grave{p}_n & p_n \\ \grave{e}_n & \tau_n e_n \end{bmatrix}. \tag{4.10}$$

These formulas remain true if some points z_j converge to zeros of t_{m+n-1}. Therefore, clearly, at any point where both $\grave{e}_n$ and e_n vanish, f and g must vanish also, in contrast to the assumption that f and g are relatively prime. Finally, if $n = 0$ and $m > 0$, then, for $i \in \mathbb{N}^+$, $0 = \grave{e}_0(z_{m+i}) = g(z_{m+i})$ implies that $gp_0 + fq_0$ does not vanish at z_{m+i} since f and g are relatively prime and q_0 is a nonzero constant. Q.E.D.

Note that the formulas (4.10) can be summarized as

$$[g \ \ f] = \frac{1}{\delta_n}[\grave{e}_n \ \ \tau_n e_n] \begin{bmatrix} q_n & -p_n \\ -\grave{q}_n & \grave{p}_n \end{bmatrix}. \tag{4.11}$$

Both for the general NPF recurrence and, of course, for the MPF recurrence to be discussed later, we need analogous definitions and statements for MPFs. For a $[\mu(n); n]$ MPF $(u_n, v_n) := (u_{\mu(n);n}, v_{\mu(n);n})$ of $(\hat{f}, \hat{g}; f, g)$, we let

$$(\grave{u}_n, \grave{v}_n) := (u_{\mu(n)-1;n-1}, v_{\mu(n)-1;n-1}) \tag{4.12}$$

denote its upper left neighbor. We call the MPF (u_n, v_n) *weakly regular* if

$$\frac{u_n}{v_n} \neq \frac{\grave{u}_n}{\grave{v}_n}, \quad \text{i.e.,} \quad \grave{u}_n v_n - u_n \grave{v}_n \neq 0 \in \mathcal{L}. \tag{4.13}$$

The *residual* $(\hat{e}_n, e_n)$ of the MPF (u_n, v_n) consists of a formal Laurent series $\hat{e}_n \in \mathcal{L}_0^\star$ and a formal Newton series e_n for the points $z_{\mu(n)+n+i}$, $i = 1, 2, \ldots$, defined implicitly by

$$\hat{g}(z)u_n(z) + \hat{f}(z)v_n(z) = z^{\mu(n)}\,\hat{e}_n(z), \tag{4.14a}$$
$$g(z)u_n(z) + f(z)v_n(z) = t_{\mu(n)+n+1}(z)\,e_n(z). \tag{4.14b}$$

The definition of this residual and of the residual $(\grave{\hat{e}}_n, \grave{e}_n)$ of $(\grave{u}_n, \grave{v}_n)$ can be summarized in

$$\begin{aligned}\begin{bmatrix} \hat{g} & \hat{f} \\ g & f \end{bmatrix}\begin{bmatrix} \grave{u}_n & u_n \\ \grave{v}_n & v_n \end{bmatrix} &= \begin{bmatrix} z^{\mu(n)-1}\,\grave{\hat{e}}_n & z^{\mu(n)}\,\hat{e}_n \\ t_{\mu(n)+n-1}\,\grave{e}_n & t_{\mu(n)+n+1}\,e_n \end{bmatrix} \\ &= \begin{bmatrix} z^{\mu(n)-1} & 0 \\ 0 & t_{\mu(n)+n-1} \end{bmatrix}\begin{bmatrix} \grave{\hat{e}}_n & z\hat{e}_n \\ \grave{e}_n & \hat{\tau}_n e_n \end{bmatrix},\end{aligned} \tag{4.15}$$

where

$$\hat{\tau}_n(z) := (z - z_{\mu+n-1})(z - z_{\mu+n}) = t_2^{(\mu+n-2)}(z). \tag{4.16}$$

However, in the following we first consider $[\kappa + 1; k]$ MPFs ($k \geq 1$); i.e., we have to replace μ and n by $\kappa + 1$ and k in all formulas. We let $(p_n, q_n) := (p_{m(n),n}, q_{m(n),n})$ be a fixed weakly regular (m, n) NPF. We assume that its upper left neighbor $(\grave{p}_n, \grave{q}_n)$ and the residuals e_n and $\grave{e}_n$ of this pair are also at our disposal. Then we can consider the following $[\kappa + 1; k]$ multipoint Padé problem for the polynomials $u_k^{(n)}$ and $v_k^{(n)}$, which will play the role of "recurrence coefficients":

$$z^{-m+1}[\grave{p}_n(z)u_k^{(n)}(z) + p_n(z)v_k^{(n)}(z)] = O_-(z^{\kappa+1}), \tag{4.17a}$$
$$\grave{e}_n(z)u_k^{(n)}(z) + \tau_n(z)e_n(z)v_k^{(n)}(z) = O(t_{\kappa+k+2}^{(m+n-2)}(z)). \tag{4.17b}$$

The polynomial $t_{\kappa+k+2}^{(m+n-2)}$ is still defined by (4.5). In our notation from §2 we have for this problem to substitute $(\hat{f}, \hat{g}; f, g)$ by

$$\begin{aligned}(z^{-m+1}p_n(z), z^{-m+1}\grave{p}_n(z);\ \tau_n(z)e_n(z), \grave{e}_n(z)) \in \\ \mathcal{L}_1^\star \times \mathcal{L}_0^\star \times \mathcal{N}_{Z(n)} \times \mathcal{N}_{Z(n)}.\end{aligned} \tag{4.18}$$

(Recall that $Z(n)$ was defined in Lemma 8.) By Lemma 5 we know that $\partial(z^{-m+1}\grave{p}_n) = 0$ if (p_n, q_n) is weakly row-regular, and that $\partial(z^{-m+1}p_n) = 1$ if (p_n, q_n) is weakly column-regular. Hence, the assumption of §2 that $\hat{\gamma}_0 \neq 0$ or $\hat{\phi}_\iota \neq 0$ is fulfilled if we let $\iota = 1$. In addition to this $[\kappa + 1; k]$ MPF $(u_k^{(n)}, v_k^{(n)})$ we consider its upper left neighbor in the multipoint Padé table for the data (4.18), denoting it by $(\grave{u}_k^{(n)}, \grave{v}_k^{(n)})$.

We restrict κ to $-k \leq \kappa \leq k$, so that according to (2.3)

$$(\grave{u}_k, \grave{v}_k) \in \mathcal{P}_k \times \mathcal{P}_{k-1}, \quad (u_k^{(n)}, v_k^{(n)}) \in \mathcal{P}_{k+1} \times \mathcal{P}_k. \tag{4.19}$$

According to (4.15) the residual $(\hat{e}_k^{(n)}, e_k^{(n)})$ of $(u_k^{(n)}, v_k^{(n)})$ and the residual $(\grave{\hat{e}}_k^{(n)}, \grave{e}_k^{(n)})$ of $(\grave{u}_k^{(n)}, \grave{v}_k^{(n)})$ satisfy

$$\begin{bmatrix} z^{-m+1}\grave{p}_n & z^{-m+1}p_n \\ \grave{e}_n & \tau_n e_n \end{bmatrix} \begin{bmatrix} \grave{u}_k^{(n)} & u_k^{(n)} \\ \grave{v}_k^{(n)} & v_k^{(n)} \end{bmatrix} = \begin{bmatrix} z^{\kappa} & 0 \\ 0 & t_{\kappa+k}^{(m+n-2)} \end{bmatrix} \begin{bmatrix} \grave{\hat{e}}_k^{(n)} & z\,\hat{e}_k^{(n)} \\ \grave{e}_k^{(n)} & \tau_{n+k}\, e_k^{(n)} \end{bmatrix}, \tag{4.20}$$

where

$$\tau_{n+k}(z) := (z - z_{\kappa+k+m+n-1})(z - z_{\kappa+k+m+n}) = t_2^{(\kappa+k+m+n-1)}(z). \tag{4.21}$$

Note that τ_n is a factor of $u_k^{(n)}$ since $\kappa + k \geq 0$, and a factor of $\grave{u}_k^{(n)}$ if $\kappa + k \geq 2$. In particular, if $\kappa = k$ (≥ 1), then this multipoint Padé problem with $\iota = 1$ can be reduced to one with $\iota = -1$, which means reducing the degree of freedom by 2.

In the rest of this section, f and g will again refer to the data for the Newton–Padé problem for which (p_n, q_n) is an (m, n) NPF.

The general recurrence formula for NPFs from [23], which is stated next, assumes that (p_n, q_n), $(\grave{p}_n, \grave{q}_n)$, and $(e_n, \grave{e}_n)$ are known, and relies on the pair of MPFs just considered.

THEOREM 9 (General NPF recurrence). *Let a relatively prime pair $(f, g) \in \mathcal{N}_Z \times \mathcal{N}_Z$ be given. Let $(m, n) \in \mathbf{N}^+ \times \mathbf{N}$, and let $[\kappa; k] \in \mathbf{Z} \times \mathbf{N}^+$ be such that $-k \leq \kappa \leq k$, $m + \kappa > 0$. Assume that (p_n, q_n) is a weakly regular (m, n) NPF of (f, g) with residual e_n, and let $(\grave{p}_n, \grave{q}_n)$ be an $(m - 1, n - 1)$ NPF of (f, g) with residual $\grave{e}_n$. Moreover, let $(u_k^{(n)}, v_k^{(n)})$ be a $[\kappa + 1; k]$ MPF with residual $(\hat{e}_k^{(n)}, e_k^{(n)})$ of the data (4.18), and let $(\grave{u}_k^{(n)}, \grave{v}_k^{(n)})$ be a $[\kappa; k - 1]$ MPF with residual $(\grave{\hat{e}}_k^{(n)}, \grave{e}_k^{(n)})$ for the same data. Then, the formulas*

$$\begin{bmatrix} \grave{p}_{n+k} & p_{n+k} \\ \grave{q}_{n+k} & q_{n+k} \end{bmatrix} := \begin{bmatrix} \grave{p}_n & p_n \\ \grave{q}_n & q_n \end{bmatrix} \begin{bmatrix} \grave{u}_k^{(n)} & u_k^{(n)} \\ \grave{v}_k^{(n)} & v_k^{(n)} \end{bmatrix} \tag{4.22}$$

*and**

$$t_{m+n+k+\kappa-1}\begin{bmatrix} \grave{e}_{n+k} & \tau_{n+k}e_{n+k} \end{bmatrix} := t_{m+n-1}\begin{bmatrix} \grave{e}_n & \tau_n e_n \end{bmatrix} \begin{bmatrix} \grave{u}_k^{(n)} & u_k^{(n)} \\ \grave{v}_k^{(n)} & v_k^{(n)} \end{bmatrix} \tag{4.23}$$

yield an $(m + \kappa, n + k)$ NPF (p_{n+k}, q_{n+k}) and an $(m + \kappa - 1, n + k - 1)$ NPF $(\grave{p}_{n+k}, \grave{q}_{n+k})$ of (f, g), as well as the corresponding residuals e_{n+k} and $\grave{e}_{n+k}$, which are equal to $e_k^{(n)}$ and $\grave{e}_k^{(n)}$, respectively. The new NPF (p_{n+k}, q_{n+k}) is weakly regular if and only if also $(u_k^{(n)}, v_k^{(n)})$ is weakly regular.

* On the left-hand side of Eq. (5.20) in [23], τ_n should be replaced by τ_{n+k} too.

Proof. The proof of this crucial result from [23] is so short that we can repeat it here. Since $(p_n, q_n) \in \mathcal{P}_m \times \mathcal{P}_n$ and $(\grave{p}_n, \grave{q}_n) \in \mathcal{P}_{m-1} \times \mathcal{P}_{n-1}$, (4.19) indicates that the pair (p_{n+k}, q_{n+k}) defined by (4.22) lies in $\mathcal{P}_{m+k} \times \mathcal{P}_{n+k}$. However, by definition of $(u_k^{(n)}, v_k^{(n)})$ as a $[\kappa + 1; k]$ MPF of (4.18), we have actually $p_{n+k} \in \mathcal{P}_{m+\kappa}$, cf. (4.17a)–(4.17b). Analogously, it can be verified that $(\grave{p}_{n+k}, \grave{q}_{n+k})$ defined by (4.22) lies in $\mathcal{P}_{m+\kappa-1} \times \mathcal{P}_{n+k-1}$. Moreover, by (4.6) and (4.20),

$$\begin{aligned}
[g \;\; f] \begin{bmatrix} \grave{p}_{n+k} & p_{n+k} \\ \grave{q}_{n+k} & q_{n+k} \end{bmatrix} &= [g \;\; f] \begin{bmatrix} \grave{p}_n & p_n \\ \grave{q}_n & q_n \end{bmatrix} \begin{bmatrix} \grave{u}_k^{(n)} & u_k^{(n)} \\ \grave{v}_k^{(n)} & v_k^{(n)} \end{bmatrix} \\
&= t_{m+n-1} [\grave{e}_n \;\; \tau_n e_n] \begin{bmatrix} \grave{u}_k^{(n)} & u_k^{(n)} \\ \grave{v}_k^{(n)} & v_k^{(n)} \end{bmatrix} \\
&= t_{m+n-1} \left[t_{k+\kappa}^{(m+n-2)} \grave{e}_k^{(n)} \;\; t_{k+\kappa+2}^{(m+n-2)} e_k^{(n)} \right] \\
&= t_{m+n+k+\kappa-1} \left[\grave{e}_k^{(n)} \;\; \tau_{n+k} e_k^{(n)} \right].
\end{aligned}$$

By (4.23), the expression on the second line is equal to $t_{m+n+k+\kappa-1} \left[\grave{e}_{n+k} \;\; \tau_{n+k} e_{n+k} \right]$. Altogether, this shows that $(\grave{p}_{n+k}, \grave{q}_{n+k})$ and (p_{n+k}, q_{n+k}) are an $(m + \kappa - 1, n + k - 1)$ and an $(m + \kappa, n + k)$ NPF of (f, g), respectively, and that $\grave{e}_{n+k}$ and e_{n+k} defined by (4.23), as well as $\grave{e}_k^{(n)}$ and $e_k^{(n)}$ defined by (4.20), are equal to their residuals. Finally, if and only if both (p_n, q_n) and $(u_n^{(k)}, v_n^{(k)})$ are weakly regular, the determinants* of both matrices on the right-hand side of (4.22) do not vanish identically, hence the same is true for their product, which means that (p_{n+k}, q_{n+k}) is weakly regular. Q.E.D.

As pointed out in [23], Theorem 9 has an interpretation in terms of continued fractions. For example, for a diagonal sequence of weakly regular NPFs, p_n/q_n is a convergent (i.e., a 'partial sum') of the *diagonal G-fraction* [25] of $-f/g$, and $e_n/\grave{e}_n$ is a representation of the corresponding tail of the G-fraction. Hence, to get p_{n+k}/q_{n+k} we just combine the nth convergent of the continued fraction with the kth convergent of its tail. In the Padé case, the diagonal G-fraction becomes a (diagonal) P-fraction [35]. This interpretation is not restricted to diagonal sequences, however.

5. General Recurrence Relations for Proper MPAs

Let us return to the multipoint Padé problem of §2. Again, consider a $[\mu(n); n]$ MPF $(u_n, v_n) := (u_{\mu(n);n}, v_{\mu(n);n})$ of $(\hat{f}, \hat{g}; f, g)$ and its upper left neighbor $(\grave{u}_n, \grave{v}_n)$. Recall from (4.13)–(4.15) the definitions of weak regularity and of the residuals of this pair of MPFs. First, we want to establish analogues of the Lemmas 4–6.

---

* In the proof in [23] we also meant the determinants of the matrices, not the matrices themselves.

LEMMA 4'. *Let* (u_n, v_n) *be a proper* $[\mu(n); n]$ *MPF (with* $n > |\mu_0| \geq 0$*), and let* $(\grave{e}_n, \grave{e}_n)$ *be the residual of an* $[\mu(n) - 1; n - 1]$ *MPF* $(\grave{u}_n, \grave{v}_n)$*. Then the following statements are equivalent:*

(i) (u_n, v_n) *is weakly regular, i.e.,* (4.13) *holds;*

(ii) $\grave{e}_n(z_{\mu(n)+n}) \neq 0$ *and* $(u_n(z_{\mu(n)+n}) \neq 0$ *or* $v_n(z_{\mu(n)+n}) \neq 0)$*;*

(iii) $\grave{e}_n(z_{\mu(n)+n-1}) \neq 0$ *and* $(u_n(z_{\mu(n)+n-1}) \neq 0$ *or* $v_n(z_{\mu(n)+n-1}) \neq 0)$.

Proof. If (u_n, v_n) is weakly regular, one can conclude from the definition (2.4a)–(2.4b) or from Theorem 1' that $(\grave{u}_n, \grave{v}_n)$ does not have $z_{\mu+n-1}$ or $z_{\mu+n}$ as an extra interpolation point. If, say, $z_{\mu+n}$ were one, then $(u_n, v_n) := ((z - z_{\mu+n-1})\grave{u}_n, (z - z_{\mu+n-1})\grave{v}_n)$ would be a $[\mu; n]$ MPF, possibly with $z_{\mu+n-1}$ as an unattainable point, and consequently, (4.13) would not hold. Similarly we can conclude that u_n and v_n cannot vanish simultaneously at $z_{\mu+n-1}$ or $z_{\mu+n}$. Because, if they did, say at $z_{\mu+n}$, then we could obtain a $[\mu - 1; n - 1]$ MPF $(\grave{u}_n, \grave{v}_n)$ by canceling the common factor $z - z_{\mu+n}$ in (u_n, v_n). The reverse directions of these two conclusions are verified easily too. For assume that (u_n, v_n) is not weakly regular and, say, $z_{\mu+n}$ is not an extra interpolation point of $(\grave{u}_n, \grave{v}_n)$. Then, $z_{\mu+n}$ is an unattainable point of u_n/v_n, and hence, by Theorem 1', $z - z_{\mu+n}$ is a factor of $s_{\mu;n}$, and therefore a common factor of u_n and v_n. Q.E.D.

LEMMA 5'. *If* (u_n, v_n) *is a weakly regular proper* $[\mu(n); n]$ *MPF (with* $n > |\mu_0| \geq 0$*), then*

(i) $\partial v_n = n$ *and* $\partial \grave{u}_n = \mu(n) - 1$

or

(ii) $\partial u_n = \mu(n)$ *and* $\partial \grave{v}_n = n - 1$.

Proof. Assume (u_n, v_n) is weakly regular. Then not only the upper left neighbor of the $[\mu; n]$ entry $r_n := u_n/v_n$, but also the $[\mu - 1; n]$ entry or the $[\mu; n - 1]$ entry (or both) are different from r_n. If, say, the $[\mu - 1; n]$ entry differs, then clearly $\partial u_n = \mu$; otherwise (u_n, v_n) would qualify as an $[\mu - 1; n]$ MPF. Hence, (ii) holds unless $\partial \grave{v}_n < n - 1$, in which case the pair $((z - z_{\mu+n-1})\grave{u}_n, (z - z_{\mu+n-1})\grave{v}_n)$ qualifies as $[\mu; n - 1]$ MPF, which means that also the left neighbor differs from r_n; consequently, $\partial v_n = n$. Therefore, (i) holds unless $\partial \grave{u}_n < \mu - 1$. In the latter case we could conclude that the same pair is a $[\mu - 1; n]$ MPF, and that the $[\mu - 1; n - 1]$, the $[\mu; n - 1]$, and the $[\mu - 1; n]$ MPA are all the same; this would imply that they agree with the $[\mu; n]$ MPA, in contrast to the weak regularity of the latter. Q.E.D.

LEMMA 6'. *Let* (u_n, v_n) *be a proper* $[\mu(n); n]$ *MPF (with* $n > |\mu_0| \geq 0$*); then*

$$\det \begin{bmatrix} \grave{u}_n & u_n \\ \grave{v}_n & v_n \end{bmatrix} = \hat{\delta}_n t_{\mu(n)+n-1}, \tag{5.1}$$

where $\hat{\delta}_n$ *is the leading coefficient of* $\grave{u}_n v_n - \grave{v}_n u_n \in \mathcal{P}_{\mu(n)+n-1}$*, which does not vanish if and only if* (u_n, v_n) *is weakly regular.*

Proof. The determinant is equal to $\grave{u}_n v_n - \grave{v}_n u_n \in \mathcal{P}_{\iota+2n-1}$, which, by the definition of weak regularity, is identically zero if and only if the MPF (u_n, v_n) is not weakly regular. Multiplying (4.14a)–(4.14b) by $\grave{v}_n$ and $\grave{u}_n$, and their analogues for $(\grave{u}_n, \grave{v}_n)$ by v_n and u_n, we obtain by pairwise subtraction

$$\hat{g}(\grave{u}_n v_n - \grave{v}_n u_n) = O_-(z^{\mu+n-1}), \qquad g(\grave{u}_n v_n - \grave{v}_n u_n) = O(t_{\mu+n-1}),$$
$$\hat{f}(\grave{u}_n v_n - \grave{v}_n u_n) = O_-(z^{\iota+\mu+n-1}), \qquad f(\grave{u}_n v_n - \grave{v}_n u_n) = O(t_{\mu+n-1}).$$

Since f and g are relatively prime by assumption, and since $\hat{\phi}_\iota \neq 0$ or $\hat{\gamma}_0 \neq 0$ in (2.1), it follows that

$$\grave{u}_n v_n - \grave{v}_n u_n = O_-(z^{\mu+n-1}), \quad \grave{u}_n v_n - \grave{v}_n u_n = O(t_{\mu+n-1}).$$

In other words, $\grave{u}_n v_n - \grave{v}_n u_n \in \mathcal{P}_{\mu+n-1}$, and $t_{\mu+n-1}$ is a polynomial factor of it, the quotient being a scalar $\hat{\delta}_n$, which, since $t_{\mu+n-1}$ is monic, is equal to the leading coefficient of $\grave{u}_n v_n - \grave{v}_n u_n$ in $\mathcal{P}_{\mu+n-1}$. Q.E.D.

COROLLARY 7'. *If the $[\mu(n); n]$ MPF (u_n, v_n) (where $n > |\mu_0| \geq 0$) is weakly regular, then any common zero of the polynomials $\grave{u}_n$ and u_n is a zero of f or an unattainable interpolation point of u_n/v_n. In particular, if u_n/v_n is a true $[\mu(n); n]$ interpolant and f does not vanish at any interpolation point, then $\grave{u}_n$ and u_n are relatively prime. These statements remain true when $\grave{u}_n$, u_n, and f are replaced by $\grave{v}_n$, v_n, and g, respectively.*

Proof. By (5.1), if (u_n, v_n) is weakly regular, any common zero of $\grave{u}_n$ and u_n must be a zero of $t_{\mu(n)+n-1}$. On the other hand, since $\partial w = 0$ in the factorization (2.6) of (u_n, v_n) (because of uniqueness), such a zero must be either a zero of $s_{\mu;n}$, i.e., an unattainable point, or a zero of $\hat{u}_{\mu;n}$, in which case it must be a zero of f; see (2.4b). The statement about exchanging $\grave{u}_n$, u_n, $\hat{f}$, and f for $\grave{v}_n$, v_n, $\hat{g}$, and g holds by symmetry. Q.E.D.

Finally, we need to justify the later usage of the pair of residuals $(e_n, \grave{e}_n)$ as data of another multipoint Padé problem. Recall the definition of $\hat{\tau}_n$ from (4.16).

LEMMA 8'. *Assume that (u_n, v_n) is a weakly regular $[\mu(n); n]$ MPF (where $n > |\mu_0| \geq 0$), and let $\hat{Z}(n) := \{z_j \in Z\,;\, j \geq \mu(n) + n - 1\}$. Then $\grave{e}_n$ and $\hat{\tau}_n e_n$ are relatively prime elements of $\mathcal{N}_{\hat{Z}(n)}$, and* $\max\{\partial \grave{e}_n, \partial \grave{e}_n\} = 0$.

Proof. By complete analogy to the proof of Lemma 8 we can conclude from (4.15) that the formulas

$$g = \frac{1}{\hat{\delta}_n} \det \begin{bmatrix} \grave{e}_n & \hat{\tau}_n e_n \\ \grave{v}_n & v_n \end{bmatrix}, \quad f = \frac{1}{\hat{\delta}_n} \det \begin{bmatrix} \grave{u}_n & u_n \\ \grave{e}_n & \hat{\tau}_n e_n \end{bmatrix}. \tag{5.2}$$

hold, even if some of the points z_j converge to zeros of $t_{\mu+n-1}$. Moreover, by applying Cramer's rule to the first row of (4.15) we get

$$\hat{g} = \frac{\ell_{\mu;n}}{\hat{\delta}_n} \det \begin{bmatrix} \grave{e}_n & z\hat{e}_n \\ \grave{v}_n & v_n \end{bmatrix}, \quad f = \frac{\ell_{\mu;n}}{\hat{\delta}_n} \det \begin{bmatrix} \grave{u}_n & u_n \\ \grave{e}_n & z\hat{e}_n \end{bmatrix}, \tag{5.3}$$

where $\ell_{\mu;n}$ denotes the Laurent series at ∞ of the function $z^{\mu-1}/t_{\mu+n-1}(z)$. This Laurent series belongs to $\mathcal{L}^{\star}_{-n}$ and has exact degree $-n$, i.e., the coefficient of z^{-n} does not vanish. In view of $u_n \in \mathcal{P}_{\iota+n}$, $v_n \in \mathcal{P}_n$, $\grave{u}_n \in \mathcal{P}_{\iota+n-1}$, $\grave{v}_n \in \mathcal{P}_{n-1}$, $z\hat{e}_n \in \mathcal{L}^{\star}_1$, and $\grave{\hat{e}}_n \in \mathcal{L}^{\star}_0$, we see that the first determinant is in $\mathcal{L}^{\star}_n$ and the second one in $\mathcal{L}^{\star}_{\iota+n}$, so that the formulas confirm that $\hat{g} \in \mathcal{L}^{\star}_0$ and $\hat{f} \in \mathcal{L}^{\star}_\iota$. Clearly, if the leading coefficients of $\grave{\hat{e}}_n$ and $\hat{e}_n$ both vanished, those of $\hat{g}$ and $\hat{f}$ would vanish also, contrary to the assumption for (2.1). Q.E.D.

The Eqs. (5.2) and (5.3) can be combined into

$$\begin{bmatrix} \hat{g} & \hat{f} \\ g & f \end{bmatrix} = \frac{1}{\hat{\delta}_n} \begin{bmatrix} \ell_{\mu;n} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \grave{\hat{e}}_n & z\hat{e}_n \\ \grave{e}_n & \hat{\tau}_n e_n \end{bmatrix} \begin{bmatrix} v_n & -u_n \\ -\grave{v}_n & \grave{u}_n \end{bmatrix}. \tag{5.4}$$

Now assume that (u_n, v_n) is weakly regular and that, moreover, $(\grave{u}_n, \grave{v}_n)$ and the residuals $(\grave{\hat{e}}_n, \hat{e}_n)$ and $(\grave{e}_n, e_n)$ are at our disposal, so that we can additionally consider $[\kappa+1; k]$ MPFs of the following multipoint Padé problem for the "recurrence coefficients" $u_k^{(n)}$ and $v_k^{(n)}$:

$$\grave{\hat{e}}_n(z) u_k^{(n)}(z) + z\hat{e}_n(z) v_k^{(n)}(z) = O_-(z^{\kappa+1}), \tag{5.5a}$$

$$\grave{e}_n(z) u_k^{(n)}(z) + \hat{\tau}_n(z) e_n(z) v_k^{(n)}(z) = O(t_{\kappa+k+2}^{(\mu+n-2)}(z)). \tag{5.5b}$$

Hence, here the data are

$$(z\hat{e}_n(z), \grave{\hat{e}}_n(z); \hat{\tau}_n(z) e_n(z), \grave{e}_n(z)) \in \mathcal{L}^{\star}_1 \times \mathcal{L}^{\star}_0 \times \mathcal{N}_{\hat{Z}(n)} \times \mathcal{N}_{\hat{Z}(n)}. \tag{5.6}$$

($\hat{Z}(n)$ was defined in Lemma 8', and the polynomial $t_{\kappa+k+2}^{(\mu+n-2)}$ is still defined by (4.5).) From Lemma 5' we recall that $\partial(\grave{\hat{e}}_n) = 0$ or $\partial(z\hat{e}_n) = 1$. Hence, the assumption of §2 that $\hat{\gamma}_0 \neq 0$ or $\hat{\phi}_\iota \neq 0$ is again fulfilled for $\iota = 1$.

In addition to this $[\kappa+1; k]$ MPF $(u_k^{(n)}, v_k^{(n)})$ we consider its upper left neighbor, denoting it by $(\grave{u}_k^{(n)}, \grave{v}_k^{(n)})$. Restricting κ to $-k \leq \kappa \leq k$, we see that (4.19) still holds. The residuals $(\hat{e}_k^{(n)}, e_k^{(n)})$ of $(u_k^{(n)}, v_k^{(n)})$ and $(\grave{\hat{e}}_k^{(n)}, \grave{e}_k^{(n)})$ of $(\grave{u}_k^{(n)}, \grave{v}_k^{(n)})$ now satisfy

$$\begin{bmatrix} \grave{\hat{e}}_n & z\hat{e}_n \\ \grave{e}_n & \hat{\tau}_n e_n \end{bmatrix} \begin{bmatrix} \grave{u}_k^{(n)} & u_k^{(n)} \\ \grave{v}_k^{(n)} & v_k^{(n)} \end{bmatrix} = \begin{bmatrix} z^\kappa & 0 \\ 0 & t_{\kappa+k}^{(\mu+n-2)} \end{bmatrix} \begin{bmatrix} \grave{\hat{e}}_k^{(n)} & z\,\hat{e}_k^{(n)} \\ \grave{e}_k^{(n)} & \hat{\tau}_{n+k}\, e_k^{(n)} \end{bmatrix}, \tag{5.7}$$

where

$$\hat{\tau}_{n+k}(z) := (z - z_{\kappa+k+\mu+n-1})(z - z_{\kappa+k+\mu+n}) = t_2^{(\kappa+k+\mu+n-2)}(z). \tag{5.8}$$

THEOREM 9' (General proper MPF recurrence). *Let a relatively prime pair* $(f, g) \in \mathcal{N}_Z \times \mathcal{N}_Z$ *and a pair* $(\hat{f}, \hat{g}) \in \mathcal{L}^{\star}_\iota \times \mathcal{L}^{\star}_0$ *with* $\iota \geq -1$ *and* $\hat{\phi}_\iota \neq 0$ *or* $\hat{\gamma}_0 \neq 0$ *be given. Let* $[\mu; n]$, $[\kappa; k] \in \mathbf{Z} \times \mathbf{N}^+$ *be such that* $-n-1 \leq \mu \leq \iota + n$ *and*

$-k \le \kappa \le k$. Assume that (u_n, v_n) is a weakly regular $[\mu; n]$ MPF of $(\hat{f}, \hat{g}; f, g)$ with residual $(\hat{e}_n, e_n)$, and let $(\grave{u}_n, \grave{v}_n)$ be a $[\mu-1; n-1]$ MPF of $(\hat{f}, \hat{g}; f, g)$ with residual $(\grave{\hat{e}}_n, \grave{e}_n)$.

Moreover, let $(u_k^{(n)}, v_k^{(n)})$ be a $[\kappa+1; k]$ MPF with residual $(\hat{e}_k^{(n)}, e_k^{(n)})$ of the data (5.6), and let $(\grave{u}_k^{(n)}, \grave{v}_k^{(n)})$ be a $[\kappa; k-1]$ MPF with residual $(\grave{\hat{e}}_k^{(n)}, \grave{e}_k^{(n)})$ for the same data. Then, the formulas

$$\begin{bmatrix} \grave{u}_{n+k} & u_{n+k} \\ \grave{v}_{n+k} & v_{n+k} \end{bmatrix} := \begin{bmatrix} \grave{u}_n & u_n \\ \grave{v}_n & v_n \end{bmatrix} \begin{bmatrix} \grave{u}_k^{(n)} & u_k^{(n)} \\ \grave{v}_k^{(n)} & v_k^{(n)} \end{bmatrix} \tag{5.9}$$

and

$$\begin{bmatrix} z^{\kappa} & 0 \\ 0 & t_{k+\kappa}^{(\mu+n-2)} \end{bmatrix} \begin{bmatrix} \grave{\hat{e}}_{n+k} & z\hat{e}_{n+k} \\ \grave{e}_{n+k} & \hat{\tau}_{n+k} e_{n+k} \end{bmatrix} := \begin{bmatrix} \grave{\hat{e}}_n & z\hat{e}_n \\ \grave{e}_n & \hat{\tau}_n e_n \end{bmatrix} \begin{bmatrix} \grave{u}_k^{(n)} & u_k^{(n)} \\ \grave{v}_k^{(n)} & v_k^{(n)} \end{bmatrix} \tag{5.10}$$

yield a $[\mu+\kappa; n+k]$ MPF (u_{n+k}, v_{n+k}) and a $[\mu+\kappa-1; n+k-1]$ MPF $(\grave{u}_{n+k}, \grave{v}_{n+k})$ of $(\hat{f}, \hat{g}; f, g)$, as well as the corresponding residuals $(\hat{e}_{n+k}, e_{n+k})$ and $(\grave{\hat{e}}_{n+k}, \grave{e}_{n+k})$, which are equal to $(\hat{e}_k^{(n)}, e_k^{(n)})$ and $(\grave{\hat{e}}_k^{(n)}, \grave{e}_k^{(n)})$, respectively.

The new MPF (u_{n+k}, v_{n+k}) is weakly regular if and only if also $(u_k^{(n)}, v_k^{(n)})$ is weakly regular.

Proof. Since $(u_n, v_n) \in \mathcal{P}_{\iota+n} \times \mathcal{P}_n$ and $(\grave{u}_n, \grave{v}_n) \in \mathcal{P}_{\iota+n-1} \times \mathcal{P}_{n-1}$, (4.19) implies that the pairs (u_{n+k}, v_{n+k}) and $(\grave{u}_{n+k}, \grave{v}_{n+k})$ defined by (5.9) lie in $\mathcal{P}_{\iota+n+k} \times \mathcal{P}_{n+k}$ and $\mathcal{P}_{\iota+n+k-1} \times \mathcal{P}_{n+k-1}$, respectively. Moreover, by (5.9), (4.15), and (5.7),

$$\begin{aligned} \begin{bmatrix} \hat{g} & \hat{f} \\ g & f \end{bmatrix} \begin{bmatrix} \grave{u}_{n+k} & u_{n+k} \\ \grave{v}_{n+k} & v_{n+k} \end{bmatrix} &= \begin{bmatrix} \hat{g} & \hat{f} \\ g & f \end{bmatrix} \begin{bmatrix} \grave{u}_n & u_n \\ \grave{v}_n & v_n \end{bmatrix} \begin{bmatrix} \grave{u}_k^{(n)} & u_k^{(n)} \\ \grave{v}_k^{(n)} & v_k^{(n)} \end{bmatrix} \\ &= \begin{bmatrix} z^{\mu-1} & 0 \\ 0 & t_{\mu+n-1} \end{bmatrix} \begin{bmatrix} \grave{\hat{e}}_n & z\hat{e}_n \\ \grave{e}_n & \hat{\tau}_n e_n \end{bmatrix} \begin{bmatrix} \grave{u}_k^{(n)} & u_k^{(n)} \\ \grave{v}_k^{(n)} & v_k^{(n)} \end{bmatrix} \\ &= \begin{bmatrix} z^{\kappa+\mu-1} & 0 \\ 0 & t_{\kappa+k+\mu+n-1} \end{bmatrix} \begin{bmatrix} \grave{\hat{e}}_k^{(n)} & z\,\hat{e}_k^{(n)} \\ \grave{e}_k^{(n)} & \hat{\tau}_{n+k}\, e_k^{(n)} \end{bmatrix}, \end{aligned}$$

while applying (5.7) instead of the last step yields

$$\begin{bmatrix} z^{\kappa+\mu-1} & 0 \\ 0 & t_{\kappa+k+\mu+n-1} \end{bmatrix} \begin{bmatrix} \grave{\hat{e}}_{n+k} & z\,\hat{e}_{n+k} \\ \grave{e}_{n+k} & \hat{\tau}_{n+k} e_{n+k} \end{bmatrix}.$$

This shows that $(\grave{u}_{n+k}, \grave{v}_{n+k})$ and (u_{n+k}, v_{n+k}) are a $[\mu+\kappa-1; n+k-1]$ and a $[\mu+\kappa; n+k]$ MPF of $(\hat{f}, \hat{g}; f, g)$, and that $(\grave{\hat{e}}_{n+k}, \grave{e}_{n+k})$ and $(\hat{e}_{n+k}, e_{n+k})$ defined by (5.10) as well as $(\grave{\hat{e}}_k^{(n)}, \grave{e}_k^{(n)})$ and $(\hat{e}_k^{(n)}, e_k^{(n)})$ defined by (5.7) are equal to their residuals.

Again, if and only if also $(u_n^{(k)}, v_n^{(k)})$ is weakly regular, the determinants of both matrices on the right-hand side of (5.9) do not vanish identically, and thus the same is true for their product, which means that (u_{n+k}, v_{n+k}) is weakly regular. Q.E.D.

6. Comments and Conclusions

6.1. PRODUCT REPRESENTATION OF BASIC PAIRS

For the recursive construction of a regular NPF, say (p_{n_J}, q_{n_J}), we apply Theorem 9 repeatedly. Not only yields it such an NPF, but it also provides us with a representation of (p_{n_J}, q_{n_J}) and its upper left neighbor in terms of a product of 2×2 matrices whose elements are polynomials:

$$\begin{bmatrix} \grave{p}_{n_J} & p_{n_J} \\ \grave{q}_{n_J} & q_{n_J} \end{bmatrix} = \begin{bmatrix} \grave{p}_0 & p_0 \\ \grave{q}_0 & q_0 \end{bmatrix} \prod_{j=0}^{J-1} \begin{bmatrix} \grave{u}_{k_j}^{(n_j)} & u_{k_j}^{(n_j)} \\ \grave{v}_{k_j}^{(n_j)} & v_{k_j}^{(n_j)} \end{bmatrix}. \tag{6.1}$$

Here, $k_j := n_{j+1} - n_j$ and $n_0 := 0$. The index j in the product increases from left to right. In exact arithmetic we can choose k_j in each step as small as possible, namely just such that the determinant of the new factor is not identically zero. If $h = -f/g$ does not represent a rational function, we can let n_J go to infinity. The formal infinite product that is then obtained by proceeding, say, along a diagonal (i.e., by choosing $\kappa = k$) is a matrix representation of the diagonal G-fraction.

In floating-point arithmetic, one has to avoid those factors in (6.1) that are "nearly singular". For special cases, the question what "nearly singular" means, has been addressed in [10], [37], [28], [30], [29], but it still requires further investigation.

6.2. A CONNECTION BETWEEN NPFS AND MPFS

By the close analogy between the theory of weakly regular NPFs and MPFs, and, in particular, between the general recurrences of Theorems 9 and 9' most of what can be said on the computation of NPFs holds with minor modifications for the computation of MPFs. In particular, the product representation (6.1) has an MPF analogue. This is no surprise since the 2×2 polynomial matrices that are used to update the pair of NPFs at the same time update a pair of MPFs, $(u_{N_J}^{(0)}, v_{N_J}^{(0)})$ and $(\grave{u}_{N_J}^{(0)}, \grave{v}_{N_J}^{(0)})$:

$$\begin{bmatrix} \grave{p}_{n_J} & p_{n_J} \\ \grave{q}_{n_J} & q_{n_J} \end{bmatrix} = \begin{bmatrix} \grave{p}_0 & p_0 \\ \grave{q}_0 & q_0 \end{bmatrix} \begin{bmatrix} \grave{u}_{N_J}^{(0)} & u_{N_J}^{(0)} \\ \grave{v}_{N_J}^{(0)} & v_{N_J}^{(0)}. \end{bmatrix}. \tag{6.2}$$

The data of the corresponding multipoint Padé problem are

$$\begin{aligned}(z^{-m+1}p_0(z), z^{-m+1}t_{m-1}(z);\ \tau_0(z)e_0(z), \grave{e}_0(z)) \in \\ \mathcal{L}_1^\star \times \mathcal{L}_0^\star \times \mathcal{N}_{Z(0)} \times \mathcal{N}_{Z(0)},\end{aligned} \tag{6.3}$$

where $\grave{e}_0 = g$, $\tau_0 e_0 = (gp_0 + fq_0)/t_{m-1}$.

One might guess that most results for the MPF recurrence can be proved by determining first a Newton–Padé problem such that (6.2) holds. However, the data (6.3) do not look general enough so that any data $(\hat{f}, \hat{g}; f, g)$ admitted for a multipoint Padé problem could be interpreted in this way. But we could establish the

MPF results in this way for any such quadruple with $\iota = -1$, *finite* series $\hat{f}$ and $\hat{g}$ with at most m terms and $\gamma_0 := \hat{g}(\infty) = 1$, and a Newton series f satisfying $f(z_m) = f(z_{m+1}) = 0$. Setting

$$\begin{aligned} p_0(z) &:= z^{m-1}\hat{f}(z), & \grave{p}_0(z) &:= t_{m-1}(z) := z^{m-1}\hat{g}(z), \\ q_0(z) &:\equiv 1, & \grave{q}_0(z) &:\equiv 0, \\ e_0(z) &:= f(z)/\tau_0(z), & \grave{e}_0(z) &:= g(z). \end{aligned} \tag{6.4}$$

we could find from (4.11) data $(f^{\mathrm{NP}}, g^{\mathrm{NP}})$ [not to be mixed up with the given multipoint Padé data $(\hat{f}, \hat{g}; f, g)$] for a Newton–Padé problem related to the given multipoint Padé problem by (6.3).

6.3. LEVINSON-TYPE AND SCHUR-TYPE ALGORITHMS

As long as a small step size k is used in Theorem 9, the recurrence coefficients only depend on few initial coefficients of the Newton series for the residuals (or equivalent data) and the few highest coefficients of the numerators p_n and $\grave{p}_n$. There are two ways to deal with the data of the initial problem. In each step we either transform all the remaining data, i.e., compute a full representation of the residuals (either as Newton series or as functions values and, possibly, derivatives), or we have to compute the data needed for the next step from the definition (4.6) of the residuals. Adopting the terminology used for fast Toeplitz solvers, which correspond to recurrences along two adjacent rows in the Padé table, we can say that in the first case we obtain a Schur-type algorithm, and in the second case a Levinson-type algorithm. The superfast algorithms discussed next combine the merits of both approaches.

Again, these remarks also apply to the multipoint Padé problem, except that the numerators $\grave{p}_n$ and p_n are replaced by the Laurent series residuals $\grave{e}_n$ and e_n.

6.4. SUPERFAST RATIONAL INTERPOLATION

The straightforward way of evaluating the product (6.1) consists in multiplying it out from left to right. In other words, the partial product represented by the polynomials $\grave{p}_{n_j}$, $\grave{q}_{n_j}$, p_{n_j}, and q_{n_j} (whose degrees are growing) is multiplied by the next factor with the polynomials $\grave{u}_{k_j}^{(n_j)}$, $u_{k_j}^{(n_j)}$, $\grave{v}_{k_j}^{(n_j)}$, and $v_{k_j}^{(n_j)}$ (whose degrees are normally small). In the generic situation, where the Newton–Padé table is normal, i.e., has no singular blocks, we do $N := N_J$ steps of total cost $O(\sum_{n=1}^{N} n) = O(N^2)$.

The fastest way of evaluating the product (6.1) is by a binary tree. Assuming a normal table and $N = 2^M$, we compute at the lowest level $\frac{1}{2}N$ matrix products at cost $O(1)$ each. At the highest level, there is just one matrix product involving polynomials of degree $\sim \frac{1}{2}N$. Using the fast Fourier transform (FFT) this last product costs $O(N \log N)$. At an intermediate level j, there are $2^{-j}N$ products costing $O(2^j \log 2^j)$ each; so the cost of this level is $O(Nj)$. Since j runs from 1 to $M = \log_2 N$, total cost is $O(N \log^2 N)$.

However, before we can evaluate product (6.1), we need to determine its factors. The individual factors are found by solving small, particularly structured linear systems. In the generic case they have order 2, and since there are N such factors, the total cost for solving these systems is $O(N)$. But before we can solve them, we need to prepare the corresponding data, and this is the most cost-critical part of the whole recursive process. The *divide-and-conquer strategy* allows us to reduce the given data according to a binary tree: whenever we have solved a problem of size $n = 2^j$, we can use the definitions (4.20) or (5.7), respectively, to find the corresponding new residuals of which we need to compute $O(n)$ terms in order to have the data for a new step of size $k = n$ ready. (The exact number of terms needed depends on the path taken in the Newton–Padé table.) Unless $k = 1$, this step will be divided further, but this is of no concern. The cost of computing residuals clearly depends on how the data and the residuals are represented. If this is by Newton series, then one can not expect, in general, to compute n coefficients of a residual in less than $O(n^2)$ operations. In contrast, in the case of Padé approximation or two-point Padé approximation, where data and residuals are represented by power and Laurent series, FFTs can be applied and reduce the cost to $O(n \log^2 n)$ operations. This gives rise to algorithms of total complexity $O(N \log^2 N)$. This principle was first explored by Brent, Gustavson, and Yun [6] for Padé recursions along a staircase, which lead to a superfast Hankel solver; see also [18]. Since 1980, it has appeared in many forms in superfast Hankel and Toeplitz solvers; for references see, e.g., [8], [28], [30], [31]. It is a natural question to ask whether this operation count persists for certain rational interpolation problems involving more than two interpolation points. (Any two finite points could be mapped into 0 and ∞ by a fractional linear transformation that leaves $\mathcal{R}_{n,n}$ invariant.) Of particular interest are the diagonal recurrences ($m = n - 1$ or $m = n$, $\kappa = k$), which generalize the Hankel solvers, and the row recurrences ($m = \frac{1}{2}N - 1$, $\kappa = 0$), which generalize the Toeplitz solvers.

For interpolation in a few, say L, points $\hat{z}_\ell$ that all carry the same amount of data to be taken into account in a cyclic manner, a superfast rational interpolation algorithm indeed exists. Let in each point the data be given as a finite power series in $z - z_\ell$. In all our computation we want to store L different representations of the NPFs and MPFs involved, namely as power series in $z - z_\ell$. To switch between them would be more expensive than to build them up in parallel. Since L is independent of N, this does not increase the order of complexity. Assuming a normal Newton–Padé table, we can concentrate on step sizes where $\kappa + k$ is a multiple of L. Then the polynomials $t_{\kappa+k}^{(m+n-2)}$ and $t_{\kappa+k}^{(\mu+n-2)}$ in (4.20) and (5.7), by which we need to divide, are powers of $t_L(z) = (z - z_0) \cdots (z - z_{L-1})$. The FFT can be applied to build up these powers by recursive doubling and to perform the division.

Rational interpolation at the $2N$th roots of unity by $r \in \mathcal{R}_{N-1,N}$ is another interesting case that can be solved in $O(N \log^2 N)$ operations. At the lowest level of the row recurrence we interpolate at one new root (i.e., $\kappa = 0$, $k = 1$). For the diagonal recurrence we add at the lowest level two new roots (i.e., $\kappa = k = 1$) that lie on opposite sides of the unit circle; they are zeros of $z^2 = z_\ell$, where z_ℓ is any zero of $z_\ell^{N/2} = -1$. At the jth level of the row recurrence, and at the $(j-1)$th level of the diagonal recurrence, we interpolate additionally at the $n = 2^{j-1}$ zeros of $z^n = z_\ell$, where z_ℓ is any zero of $z_\ell^{N/n} = -1$. Since all interpolation points are distinct, the data will normally be given as function values. The computation of the values of the residuals at all nth roots of z_ℓ, which is done according to the definitions (4.20) and (5.7), requires the evaluation of the NPFs or MPFs at these points. Once more, with FFTs of length $2n$ this is accomplished in $O(n \log^2 n)$ operations. Altogether one obtains again the bound $(N/n)\,O(n \log n) = O(N \log n) = O(Nj)$ for level j, and an overall complexity of $O(N \log^2 N)$ for the algorithm.

An alternative is to apply the Kronecker algorithm [33], [45], [2] to this problem: first, one has to compute the interpolation polynomial of degree $2N-1$ by an FFT of length $2N$; then one applies the $O(N \log^2 N)$ version of the Euclidean algorithm [6] to this polynomial and $z^{2N} - 1$.

The treatment of non-normal tables and the inclusion of look-ahead for avoiding unstable intermediate results makes these superfast algorithms more complicated, but the complexity remains $O(N \log^2 N)$ as long as the look-ahead step size remains bounded independent of N.

References

1. A. C. Antoulas, On recursiveness and related topics in linear systems, *IEEE Trans. Automatic Control* **AC-31** (1986), pp. 1121–1135.
2. A. C. Antoulas, Rational interpolation and the Euclidean algorithm, *Linear Algebra Appl.* **108** (1988), pp. 157–171.
3. A. C. Antoulas, and B. D. O. Anderson, On the scalar rational interpolation problem, *IMA J. Math. Control Inform.* **3** (1986), pp. 61–88.
4. B. Beckermann, A reliable method for computing M-Padé approximants on arbitrary staircases, *J. Comput. Appl. Math.* **40** (1992), pp. 19–42.
5. B. Beckermann, and G. Labahn, A uniform approach to the fast computation of matrix-type Padé approximants, *SIAM J. Matrix Anal. Appl.* (to appear).
6. R. P. Brent, F. G. Gustavson, and C. Y. Y. Yun, Fast solution of Toeplitz systems of equations and computation of Padé approximants, *J. Algorithms* **1** (1980), pp. 259–295.
7. S. Cabay, M. H. Gutknecht, and R. Meleshko, Stable rational interpolation?, in *Proceedings MTNS-'93, Regensburg* (to appear).
8. S. Cabay, and G. Labahn, A superfast algorithm for multi-dimensional Padé systems, *Numerical Algorithms* **2** (1992), pp. 201–224.
9. S. Cabay, G. Labahn, and B. Beckermann, On the theory and computation of non-perfect Padé-Hermite approximants, *J. Comput. Appl. Math.* **39** (1992), pp. 295–313.
10. S. Cabay, and R. Meleshko, A weakly stable algorithm for Padé approximants and the inversion of Hankel matrices, *SIAM J. Matrix Anal. Appl.* **14** (1993), pp. 735–765.
11. G. Claessens, On the Newton-Padé approximation problem, *J. Approx. Theory* **22** (1978), pp. 150–160.
12. S. C. Cooper, A. Magnus, and J. H. McCabe, On the non-normal two-point Padé table, *J. Comput. Appl. Math.* **16** (1986), pp. 371–380.
13. M. Fiedler, Hankel and Loewner matrices, *Linear Algebra Appl.* **58** (1984), pp. 75–95.
14. R. W. Freund, M. H. Gutknecht, and N. M. Nachtigal, An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices, *SIAM J. Sci. Comput.* **14** (1993), pp. 137–158.

15. P. A. Fuhrmann, A matrix Euclidean algorithm and matrix continued fractions, *System Control Letters* **3** (1983), pp. 263–271.
16. P. A. Fuhrmann, On the partial realization problem and the recursive inversion of Hankel and Toeplitz matrices, *Contemporary Math.* **47** (1985), pp. 149–161.
17. M. A. Gallucci, and W. B. Jones, Rational approximations corresponding to Newton series (Newton-Padé approximants), *J. Approx. Theory* **17** (1976), pp. 366–392.
18. W. Gragg, F. Gustavson, D. Warner, and D. Yun, On fast computation of superdiagonal Padé fractions, *Math. Programming Stud.* **18** (1982), pp. 39–42.
19. W. B. Gragg, The Padé table and its relation to certain algorithms of numerical analysis, *SIAM Rev.* **14** (1972), pp. 1–62.
20. P. Graves-Morris, Efficient reliable rational interpolation, in *Padé Approximation and its Applications*, Amsterdam 1980, M. B. de Bruin and H. van Rossum, eds., Lecture Notes in Mathematics **888**, Springer-Verlag, 1981, pp. 28–63.
21. P. R. Graves-Morris, Practical, reliable, rational interpolation, *J. Inst. Maths. Applics.* **25** (1980), pp. 267–286.
22. P. R. Graves-Morris, and T. R. Hopkins, Reliable rational interpolation, *Numer. Math.* **36** (1981), pp. 111–128.
23. M. H. Gutknecht, Block structure and recursiveness in rational interpolation, in *Approximation Theory VII* (E. W. Cheney, C. K. Chui, and L. L. Schumaker, eds.), Academic Press, Boston, 1992 (to appear).
24. M. H. Gutknecht, A completed theory of the unsymmetric Lanczos process and related algorithms, Part II, *SIAM J. Matrix Anal. Appl.* (to appear).
25. M. H. Gutknecht, Continued fractions associated with the Newton-Padé table, *Numer. Math.* **56** (1989), pp. 547–589.
26. M. H. Gutknecht, The rational interpolation problem revisited, *Rocky Mountain J. Math.* **21** (1991), pp. 263–280.
27. M. H. Gutknecht, A completed theory of the unsymmetric Lanczos process, and related algorithms, Part I, *SIAM J. Matrix Anal. Appl.* **13** (1992), pp. 594–639.
28. M. H. Gutknecht, Stable row recurrences in the Padé table and generically superfast lookahead solvers for non-Hermitian Toeplitz systems, *Linear Algebra Appl.* **188/189** (1993), pp. 351–421.
29. M. H. Gutknecht, and M. Hochbruck, Look-ahead Levinson- and Schur-type recurrences in the Padé table, (in preparation).
30. M. H. Gutknecht, and M. Hochbruck, Look-ahead Levinson and Schur algorithms for non-Hermitian Toeplitz systems, *IPS Research Report* 93-11, IPS, ETH-Zürich, August 1993.
31. G. Heinig, and K. Rost, *Algebraic Methods for Toeplitz-like Matrices and Operators*, Akademie-Verlag, Berlin, Birkhäuser, Basel/Stuttgart, 1984.
32. U. Helmke, and P. A. Fuhrmann, Bezoutians, *Linear Algebra Appl.* **122/123/124** (1989), pp. 1039–1097.
33. L. Kronecker, Zur Theorie der Elimination einer Variabel aus zwei algebraischen Gleichungen, *Monatsber. Königl. Preuss. Akad. Wiss.* (1888), pp. 535–600.
34. H. Maehly, and C. Witzgall, Tschebyscheff-Approximationen in kleinen Intervallen II, Stetigkeitssätze für gebrochen rationale Approximationen, *Numer. Math.* **2** (1960), pp. 293–307.
35. A. Magnus, Certain continued fractions associated with the Padé table, *Math. Z.* **78** (1962), pp. 361–374.
36. J. Meinguet, On the solubility of the Cauchy Interpolation problem, in *Approximation Theory* (A. Talbot, ed.), Academic Press, 1970, pp. 137–163.
37. R. J. Meleshko, A stable algorithm for the computation of Padé approximants, PhD thesis, Department of Computing Science, University of Alberta (Canada) (1990).
38. G. Opitz, Steigungsmatrizen, *Z. Ang. Math. Mech.* **44** (1964), pp. T52–T54.
39. B. N. Parlett, D. R. Taylor, and Z. A. Liu, A look-ahead Lanczos algorithm for unsymmetric matrices, *Math. Comp.* **44** (1985), pp. 105–124.
40. M. Van Barel, and A. Bultheel, The "look-ahead" philosophy applied to matrix rational interpolation problems, in *Proceedings MTNS-'93, Regensburg* (to appear).

41. M. Van Barel, and A. Bultheel, A matrix Euclidean algorithm for minimal partial realization, *Linear Algebra Appl.* **121** (1989), pp. 674–682.
42. M. Van Barel, and A. Bultheel, A new approach to the rational interpolation problem, *J. Comput. Appl. Math.* **32** (1990), pp. 281–289.
43. M. Van Barel, and A. Bultheel, The computation of non-perfect Padé-Hermite approximants, *Numerical Algorithms* **1** (1991), pp. 285–304.
44. M. Van Barel, and A. Bultheel, A new formal approach to the rational interpolation problem, *Numer. Math.* **62** (1992), pp. 87–122.
45. D. D. Warner, Hermite Interpolation with Rational Functions, PhD thesis, University of California at San Diego (1974).
46. H. Werner, A reliable method for rational interpolation, in *Padé Approximation and its Applications* (L. Wuytack, ed.), Lecture Notes in Mathematics **765** , Springer-Verlag, 1979, pp. 257–277.
47. H. Werner, Algorithm 51: A reliable and numerically stable program for rational interpolation in Lagrange data, *Computing* **31** (1983), pp. 269–286.