Vectors of type II Hermite–Padé approximations and a new linear independence criterion

Raffaele Marcovecchio

Received: 17 July 2020 / Accepted: 27 March 2021 / Published online: 29 April 2021
© Fondazione Annali di Matematica Pura ed Applicata and Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract
We propose a linear independence criterion, and outline an application of it. Down to its simplest case, it aims at solving this problem: given three real numbers, typically as special values of analytic functions, how to prove that the \( \mathbb{Q} \)-vector space spanned by 1 and those three numbers has dimension at least 3, whenever we are unable to achieve full linear independence, by using simultaneous approximations, i.e. those usually arising from Hermite–Padé approximations of type II and their suitable generalizations. It should be recalled that approximations of type I and II are related, at least in principle: when the numerical application consists in specializing actual functional constructions of the two types, they can be obtained, one from the other, as explained in a well-known paper by Mahler (1968) Compos Math 19: 95–166. That relation is reflected in a relation between the asymptotic behavior of the approximations at the infinite place of \( \mathbb{Q} \). Rather interestingly, the two viewpoints split away regarding the asymptotic behaviors at finite places (i.e. primes) of \( \mathbb{Q} \), and this makes the use of type II more convenient for particular purposes. In addition, sometimes we know type II approximations to a given set of functions, for which type I approximations are not known explicitly. Our approach can be regarded as a dual version of the standard linear independence criterion, which essentially goes back to Siegel.

Keywords  Linear independence · Matrices and determinants · Orthogonal polynomials · Difference equations · Simultaneous approximations

Mathematics Subject Classification  Primary 11J72 · Secondary 11B83, 11C20, 15B36, 33C45, 39A60, 41A28, 42C05

Dedicated to Wadim Zudilin, with warm wishes, on the occasion of his 50th birthday.
1 Introduction

The systematically creative search of new irrationality (in the broad sense that includes linear independence) criteria is a flowering topic in number theory; see [4, 6, 9, 10, 12, 14, 17–19, 22, 32, 39, 40, 45, 57–59, 67], among many others. Their reviews on MathSciNet and Zentralblatt MATH provide an historically and methodologically well informed context. The development of this research topic is stimulated by a rich amount of problems, each of which carries intrinsic geometric aspects that require slightly, or sometimes substantially, new approaches. We believe that a novelty, whether substantial or just formal, is connected to suitable aspects of the established theory that are reflected in the way it should be communicated and interpreted. This introduction, as well as the many remarks we have placed all along in this paper, suggest a reception of the text according to the described tradition-oriented spirit, and it is no surprise that geometry, possibly in the modern algebra-focused meaning of the word, is again at the core of the investigation. We recommend the notes of course written by Waldschmidt [63], for a very useful introduction to the seminal work of Hermite, Lindemann, Weierstrass, Padé, Gel’fond, Schneider, Siegel, Mahler,... on this topic.

1.1 Background

Suppose we have $\gamma_1, \ldots, \gamma_m$ real numbers, and we wish to prove that the $m+1$ numbers $1, \gamma_1, \ldots, \gamma_m$ are linearly independent over the field $\mathbb{Q}$ of all rational numbers. One efficient method, when successful, is to construct $m$ sequences of small linear forms satisfying the following properties ($\mu = 1, \ldots, m$):

(i) $\varepsilon_{n}(\mu) = q_n\gamma_\mu - p_n(\mu) \in \mathbb{Z}\gamma_\mu + \mathbb{Z};$

(ii) $\varepsilon_{n}(\mu) \to 0$ for $n \to \infty;$

(iii) for all $\lambda \in \mathbb{Z}^m \setminus \{0\}$ the inner product $\lambda \cdot \varepsilon_{n}$ is nonzero for infinitely many $n.$

In a reasonably convenient situation, the assumption (ii) is achieved by standard, although possibly difficult to apply, analytic methods, while the requirement in (iii) looks elusive, and is indeed, at least sometimes, subtle to handle.

More precisely, those analytic methods are applied to suitable linear forms $\tilde{\varepsilon}_{n}(\mu) = \tilde{q}_n\gamma_\mu - \tilde{p}_n(\mu)$ with rational coefficients, in such a way that $\varepsilon_{n}(\mu) = D_n\tilde{\varepsilon}_{n}(\mu)$ and that the asymptotic behavior of the sequence of integers $D_n$ is controlled through a careful application of the Prime Number Theorem, using refined tools of different kinds, pioneered in papers like [16, 55], on one hand, and [53], on the other hand; see [61] for a historical perspective, and [37] for a tentative unification of different approaches.

In some cases, one can replace (iii) by both of

(iii.a) $1, \gamma_1, \ldots, \gamma_{m-1}$ are linearly independent over $\mathbb{Q};$

(iii.b) for all $\lambda \in \mathbb{Z}^{m-1}, |\lambda_1\varepsilon_{n}(1) + \cdots + \lambda_{m-1}\varepsilon_{n}(m-1)| < \varepsilon_{n}(m)$ for infinitely many $n.$

This is the strategy followed in [28, Sect. 5], [62, Lemma 5.1] and [54, (4-6)-(4-7)]. The methods in [54] can be, in principle, extended to the situation in [37], by playing the ‘wild card’: the $\mathbb{C}^N$-saddle point method nicely developed in [47].
1.2 The new ingredient

Let us switch to those awkward situations where (ii) does not hold, in spite of our effort to arrange such a crucial requirement. Accordingly, we confine our goal to merely proving that

$$\dim_\mathbb{Q}(\mathbb{Q}y_1 + \cdots + \mathbb{Q}y_m) \geq m,$$  \hspace{1cm} (1)

say, which, at least, is less demanding. If, for instance, $m = 3$, $y_1 = \text{Li}_1\left(\frac{1}{z}\right) = -\text{Li}_1\left(-\frac{1}{1-z}\right)$, $y_2 = \text{Li}_2\left(\frac{1}{z}\right)$, $y_3 = \text{Li}_2\left(-\frac{1}{1-z}\right)$, then $2(y_2 + y_3) = -y_1^2$, and $y_1$ is transcendental for all algebraic $z \neq 0, 1$. In several interesting cases (1) is an open problem: the numbers $y_i$ can be chosen to be values of polylogarithms at several points [15], or multiple polylogarithms [20, 25]; or your favorite to-be-proved-linearly-independent ‘constants’.

The key idea is to deal with (1) when there is no available analytic construction of functional approximations to $m - 1$ functions, that is to say: we are not able to identify a subspace of dimension $m$. We are about to suggest that a strategy to achieve (1), maybe for some other numbers unrelated to specific examples above, can still be quite well akin to (i)–(iii). We should indeed construct (virtually: see 1.3 below) $2m$ sequences of linear forms such that:

(iv) $\varepsilon^{(\mu, \nu)}_n = q_n^{(\nu)}y_\mu - p_n^{(\mu, \nu)} \in \mathbb{Z}y_\mu + \mathbb{Z}$ ($\mu = 1, \ldots, m; \nu = 1, 2$);

(v) $\Delta_n^{(\mu_1, \mu_2)} = \varepsilon^{(\mu_1, 1)}_n \varepsilon^{(\mu_2, 2)}_n - \varepsilon^{(\mu_1, 2)}_n \varepsilon^{(\mu_2, 1)}_n \to 0$ for $n \to \infty$ ($\mu_1, \mu_2 = 1, \ldots, m$);

(vi) for all $\lambda \in \mathcal{M}(2, m; \mathbb{Z})$ with rank $2$, the matrix $\lambda \cdot \varepsilon_n \in \mathcal{M}(2, 2; \mathbb{Z})$ is non-singular for infinitely many $n$.

If suitably related sequences $\tilde{\varepsilon}_n^{(\mu, \nu)}$ (see 2.5 below for more precision) are holonomic, in the sense of [64], then the sequences of determinants in $\nu$ are holonomic as well, so that the linear recurrence equation satisfied by the determinants can be computed from the equation satisfied by their entries, and $\nu$ can be achieved through the Poincaré-Perron-Pituk’s theorems. Moreover,

$$\Delta_n^{(\mu_1, \mu_2)} = (q_n^{(1)}y_{\mu_1} - p_n^{(\mu_1, 1)})(q_n^{(2)}y_{\mu_2} - p_n^{(\mu_2, 2)}) - (q_n^{(2)}y_{\mu_1} - p_n^{(\mu_2, 1)})(q_n^{(1)}y_{\mu_2} - p_n^{(\mu_1, 2)})$$

$$= (p_n^{(\mu_1, 1)}q_n^{(1)}y_{\mu_1} - p_n^{(\mu_1, 1)}p_n^{(\mu_2, 2)}y_{\mu_2}) + (p_n^{(\mu_2, 1)}q_n^{(1)}y_{\mu_1} - p_n^{(\mu_2, 1)}p_n^{(\mu_1, 2)}y_{\mu_2}).$$

is just a linear form in $1, y_{\mu_1}, y_{\mu_2}$; for this reason, the case $m = 2$ simply resolves into an application of Hermite–Padé approximations of type I. The simplest new case is therefore $m = 3$. In the general situation below, we replace the lower bound $m$ for the dimension with the lower bound $m + 2 - l$ in (1), and consider $lm$ linear forms, instead of $m$ as in (i)–(iii), or $2m$ as in (iv)–(vi). We handle the main theorems in Sect. 2, with some variations on the same theme and a refinement on the main result that should incorporate most of potential applications.

1.3 Auxiliary results between modern and neoclassical

Section 3 is devoted to an auxiliary tool: in many concrete situations, the sequences $q_n, p_n^{(\mu)}, \varepsilon^{(\mu)}_n$ in (i)–(iii) satisfy the same Poincaré-Perron-Pituk-type linear recurrence
equation of order \( m + 1 \), see [36, (2.1), (4.4) and (4.6)] for examples with \( m \) arbitrary. In such a special situation it is convenient to simply take

\[ q_n^{(\nu)} = q_{n+\nu-1}, \quad p_n^{(\mu, \nu)} = p_{n+\nu-1}^{(\mu)} \]

in (iv)–(vi), where \( q_n \) and \( p_n^{(\mu)} \) are the sequences in (i)–(iii). To achieve our program, one possible strategy is to find a recurrence equation satisfied by the minors involved. In the special case \( m = 3 \) and \( l = 2 \), this new recurrence equation has order 3, and, roughly, is (up to a suitable normalization) the adjoint of the recurrence equation for \( \epsilon_n^{(\mu)} \). For general \( 1 \leq l \leq m \), the linear recurrence equation for the determinants \( \Delta_n^{(\mu_1, \ldots, \mu_l)} \) similar to \( \Delta_n^{(\mu_1, \mu_2)} \) in (v) has order \( \binom{m}{l} \). However, as we shall show, there is a better way to achieve the main goal of this program, which is to find a technical tool that helps us to deal with (v) and (vi) above, and their generalizations to lower dimensions \( m + 2 - l \) in (1). We find very precisely the asymptotic behavior of the minors \( \Delta_n^{(\mu_1, \ldots, \mu_l)} \) in quite a general setting, which should suffice for applications. Most of the material in this section is well-known. We apply a Golden Oldie, the Sylvester–Franke Theorem, to the compound matrix of the Casoratian matrix of a difference equation, and combine that with a general result by Pituk on the asymptotic behavior of solutions of a Poincaré-Perron-type difference system.

A completely different problem is to find a recurrence equation satisfied by \( q_n, p_n^{(\mu)}, \epsilon_n^{(\mu)} \) (of course, if it exists in the first place), and how to determine the roots of the characteristic polynomial of that equation. This can be achieved in different ways, see e.g. [36] and [66].

There is a small inconsistency between the notations in Sects. 2 and 3. When we consider a recurrence of order \( m \) in Sect. 3, we aim at applications to the linear independence of \( m \) numbers \( 1, \gamma_1, \ldots, \gamma_{m-1} \) (not of \( m + 1 \) numbers), in the sense of a lower bound for the dimension, as explained above. We hope that this will not trouble the reader, and we anticipate that the expositions in the two sections are quite autonomous from each other.

### 1.4 Two examples to begin with

In Sect. 4 we sketch a possible application of our criterion. We give two examples to illustrate the results from Sect. 2 and from Sect. 3 separately. We do so building, in both cases, on the clever construction introduced in [15]. In principle, the results from Sects. 2 and 3 could be combined together in the second example, but the implementation seems to be difficult.

The simplest example we can produce involves five numbers, namely

\[ 1, Li_1\left(\frac{1}{q}\right), Li_2\left(\frac{1}{q}\right), Li_1\left(\frac{2}{q}\right), Li_2\left(\frac{2}{q}\right). \]

It is still open to find an example involving only four numbers. Our second example is with the \( km + 1 \) numbers

\[ 1, Li_1\left(\frac{1}{q}\right), \ldots, Li_k\left(\frac{1}{q}\right), \ldots, Li_1\left(\frac{1}{mq}\right), \ldots, Li_k\left(\frac{1}{mq}\right). \]

We hope that further, and perhaps more interesting, applications will be provided in some future.

### 2 Main results

Throughout this paper, we use the following abbreviations for a matrix with \( u \) rows and \( v \) columns:
\[
\rho = [\rho^{(i,j)}]_{i = 1, \ldots, u}^{j = 1, \ldots, v} = \begin{bmatrix}
\rho^{(1,1)} & \cdots & \rho^{(1,v)} \\ \\
\vdots & \ddots & \vdots \\ \\
\rho^{(u,1)} & \cdots & \rho^{(u,v)}
\end{bmatrix}.
\]

We often deal with a sub-matrix of \(\rho\):

\[
\rho^{(\mu,v)} = [\rho^{(\mu,v)}]_{i = 1, \ldots, k}^{j = 1, \ldots, l} = \begin{bmatrix}
\rho^{(\mu_1,v_1)} & \cdots & \rho^{(\mu_1,v_k)} \\ \\
\vdots & \ddots & \vdots \\ \\
\rho^{(\mu_l,v_1)} & \cdots & \rho^{(\mu_l,v_k)}
\end{bmatrix}.
\]

Sometimes we replace \(\mu\) (resp.: \(v\)) with \(k\) (resp.: \(l\)) when we select the first \(k\) rows (resp.: the first \(l\) columns). A sub-matrix of \(\rho\) with \(u\) rows (resp.: with \(v\) columns) is denoted by \(\rho^{(\mu,-v)}\) (resp.: by \(\rho^{(\mu,-\nu)}\)), while the dash is used, after the square brackets and in place of the range for the rows (or the columns), when the concerned matrix has one row (or one column). We write \(I_{(a,b)}\) for the identity matrix, and \(0_{(a,b)}\) for the zero matrix. We denote by \(\rho^{T}\) the transpose of \(\rho\). We also consider matrices divided by blocks, for example \(\rho = [\rho^{(-1,\ldots,k)}] \cup \rho^{(-1,\ldots,l-1)}\).

If \(\gamma_1, \ldots, \gamma_m\) are real numbers to be approximated, we write them as a column vector \(\gamma\). We also denote the approximations \(q_n^{(v)}\gamma - p_n^{(\mu,v)}\) as the matrix \(\gamma q_n - p_n\), where \(q_n\) is a row vector. Clearly, \(q_n\) and \(p_n\) share the same number of columns, so that we often write them together in a matrix whose first row is \(q_n\) and the remaining rows are those in \(p_n\). As usual, our proofs go by contradiction, assuming the existence of more linear relations than those allowed by our claim; in matrix notation, those relations are written as \(\sigma + \omega \gamma = 0\), where \(\omega\) is a column vector. We also often use the notation \([\sigma | \omega]\) to represent a matrix whose first column is \(\sigma\), and the remaining columns are those of \(\omega\). To sum up, the vectors \(\gamma\) and \(\omega\) are columns, while \(q_n\) is a row, and the products \(\sigma q_n\) and \(\gamma q_n\) are not scalar products, except for \(\sigma q_n\) in the special case recalled by (i)–(iii) in the introduction, where \(q_n = q_n\) and \(\omega = \omega\) have length 1.

### 2.1 The main criterion

Let \(\gamma_1, \ldots, \gamma_m\) be real numbers, and let \(l \in \{1, \ldots, m\}\). Let also \(q_n^{(v)}\), \(p_n^{(1,v)}\), \ldots, \(p_n^{(m,v)}\), with \(v = 1, \ldots, l\), be \((m+1)\) sequences of integers (hereafter: elements of \(\mathbb{Z}\)), and put

\[
e^{(\mu,v)}_n := q_n^{(v)}\gamma - p_n^{(\mu,v)} \quad (\mu = 1, \ldots, m; v = 1, \ldots, l).
\]

**Theorem 2.1** Suppose that for all choices of \(l\) (distinct) indices \(\mu = (\mu_1, \ldots, \mu_l)\) from 1 to \(m\),

\[
\det e^{(\mu,-\nu)}_n = \det e^{(\mu,v)}_n \quad (\mu = \mu_1, \ldots, \mu_l \to 0 \quad (n \to \infty).
\]

Furthermore, suppose that for all \(\lambda^{(i,j)} \in \mathbb{Z} \ (i = 1, \ldots, l; j = 1, \ldots, m)\) such that the matrix

\[
\lambda = \begin{bmatrix}
\lambda^{(i,j)} \\
\end{bmatrix} \quad i = 1, \ldots, l \quad \in \mathcal{M}(l,m;\mathbb{Z})
\]

has rank \(l\), the square matrix
is non-singular for infinitely many \( n \).

Then

\[
\dim_{\mathbb{Q}}(\mathbb{Q} + \mathbb{Q} \gamma_1 + \cdots + \mathbb{Q} \gamma_m) \geq 2 + m - l.
\]  

(4)

**Proof** Arguing by contradiction, thus allowing

\[
\dim_{\mathbb{Q}}(\mathbb{Q} + \mathbb{Q} \gamma_1 + \cdots + \mathbb{Q} \gamma_m) \leq 1 + m - l,
\]

let \( \omega^{(i)}, \omega^{(i,j)} \in \mathbb{Z} \) be such that

\[
\omega^{(i)} + \omega^{(i-1)} \gamma_1 + \cdots + \omega^{(i,m)} \gamma_m = 0 \quad (i = 1, \ldots, l),
\]

i.e. \( \omega + \omega \gamma = 0 \), and that \( [\omega|\omega] \) and \( \omega \) have rank \( l \). We shorten (2) by \( \varepsilon_n = \gamma q_n - p_n \), thus obtaining

\[
\omega \varepsilon_n = \omega \gamma q_n - \omega p_n = -(\omega q_n + \omega p_n) = -[\omega|\omega] \left[ \begin{array}{c} q_n \\ -p_n \end{array} \right].
\]

By our assumptions, this must be a non-singular matrix, so that its determinant has to be a nonzero integer, for infinitely many \( n \). On the other hand, by the Binet–Cauchy formula for the determinant of a product of two matrices,

\[
\det \omega \varepsilon_n = \sum_{\mu} \det \omega^{(-\mu)} \det \varepsilon_n^{(\mu,-)},
\]

where the sum is over all multi-indices \( \mu \) such that \( 1 \leq \mu_1 < \cdots < \mu_l \leq m \). Therefore, using (3), \( \det \omega \varepsilon_n \to 0 \) as \( n \to \infty \). This contradiction ends the proof of (4). \( \square \)

**Remark 2.1** By the Binet–Cauchy formula, the assumption (3) is equivalent to

\[
\det \varepsilon_n \varepsilon_n = \sum_{\mu} \det^2 \varepsilon_n^{(\mu,-)} \to 0 \quad (n \to \infty),
\]

with an interesting interpretation of the determinant as the square of the \( l \)-dimensional volume of the parallelootope generated by the \( l \) columns of \( \varepsilon_n \) in \( \mathbb{R}^m \). The non-vanishing assumption has the following geometric meaning: for all \( l \)-dimensional subspaces \( V \) of \( \mathbb{R}^m \) generated by \( l \) vectors of \( \mathbb{Z} \) (or of \( \mathbb{Q} \)) the orthogonal projection of the above parallelootope over \( V \) has nonzero \( l \)-volume for infinitely many \( n \).

Also, the validity of the non-vanishing assumption is checked more easily if we have a prior partial information on the linear independence of some numbers among \( 1, \gamma_1, \ldots, \gamma_m \), in analogy to the situation outlined in the introduction.

**Remark 2.2** It’s worth noticing that each \( \det \varepsilon_n^{(\mu,-)} \) is a linear combination of \( 1, \gamma_{\mu_1}, \cdots, \gamma_{\mu_l} \), because
Vectors of type II Hermite–Padé approximations and a new linear…

\[
\begin{align*}
\det \epsilon_n^{(\mu,-)} &= \det [\gamma q_n - p_n]^{(\mu,-)} = \det \begin{bmatrix}
\gamma | p_n \\
- & - \\
1 | q_n
\end{bmatrix}^{(\hat{\mu},-)},
\end{align*}
\]

where \(\hat{\mu} = (\mu_1, \ldots, \mu_l, l + 1)\), as is easily seen from

\[
\begin{bmatrix}
\gamma q_n - p_n | \gamma \\
\theta^{(i,j)} | 1
\end{bmatrix} = \begin{bmatrix}
\gamma | p_n \\
- & - \\
1 | q_n
\end{bmatrix}^{(\mu,-)} = \begin{bmatrix}
\gamma | p_n \\
- & - \\
1 | q_n
\end{bmatrix} - f^{(i,j)} q^{(i,1)}.
\]

This generalizes an observation we made in Sect. 1.2 for the case \(l = 2\).

### 2.2 First variation

By repeating the same proof as in Theorem 2.1, we have

**Theorem 2.2** Suppose that for all choices of \(l\) (distinct) indices \(\mu = (\mu_1, \ldots, \mu_l)\) from 1 to \(m\),

\[
\det \epsilon_n^{(\mu,-)} = \det \epsilon_n^{(\mu,\nu)} | \mu = \mu_1, \ldots, \mu_l \to 0 \quad (n \to \infty).
\]

Furthermore, suppose that for all \(\theta^{(i)}, \lambda^{(i)} \in \mathbb{Z} \quad (i = 1, \ldots, l; j = 1, \ldots, m)\) such that the matrix

\[
[\theta | \lambda] = \begin{bmatrix}
\theta^{(i)} & i = 1, \ldots, l \\
\lambda^{(i,j)} & j = 1, \ldots, m
\end{bmatrix}
\]

has rank \(l\), the square matrix

\[
[\theta | \lambda] q_n = \theta q_n + \lambda p_n
\]

\[
= \begin{bmatrix}
\theta^{(i)} & i = 1, \ldots, l \\
\lambda^{(i,j)} & j = 1, \ldots, m
\end{bmatrix} q_n - \begin{bmatrix}
\theta^{(i)} & i = 1, \ldots, l \\
\lambda^{(i,j)} & j = 1, \ldots, m
\end{bmatrix} p_n^{(\mu,\nu)} \mu = 1, \ldots, m
\]

is non-singular for infinitely many \(n\).

Then

\[
\dim_{\mathbb{Q}} (\mathbb{Q} + q_{\gamma_1} + \cdots + q_{\gamma_m}) \geq 2 + m - l.
\]

### 2.3 Second variation

It may be appropriate to have alternative versions of the above criterion, for use in different situations. Therefore, let us change our setting a little bit. We still have \(m\) real numbers
\(\gamma_1, \ldots, \gamma_m\) and \(1 \leq l \leq m\). Now, let \(q_n^{(v)}, q_n^{(1,v)}, \ldots, q_n^{(m,v)}\) \((v = 0, \ldots, m)\) be \((m + 1)^2\) sequences in \(\mathbb{Z}\). Let us extend the notation in (2) accordingly:

\[
\epsilon_n^{(\mu,v)} := q_n^{(v)} \gamma_\mu - p_n^{(\mu,v)} \quad (\mu = 1, \ldots, m; v = 0, \ldots, m).
\]

**Theorem 2.3** Suppose that for all choices of \(l\) (distinct) indices \(\mu = (\mu_1, \ldots, \mu_l)\) from 1 to \(m\) and \(v = (v_1, \ldots, v_l)\) from 0 to \(m\),

\[
\det \epsilon_n^{(\mu,v)} \rightarrow 0 \quad (n \rightarrow \infty).
\]

Furthermore, suppose that

\[
\det \begin{bmatrix} q_n \\ -p_n \end{bmatrix} \neq 0 \quad (n = 0, 1, 2, \ldots).
\]

Then

\[
\dim_{\mathbb{Q}}(\mathbb{Q} + \mathbb{Q} \gamma_1 + \cdots + \mathbb{Q} \gamma_m) \geq 2 + m - l.
\]

**Proof** As above, we argue by contradiction, and, like in the previous proof, we have

\[
\omega \epsilon_n = \omega \gamma q_n - \omega p_n = -(\omega q_n + \omega p_n) = -[\omega | \omega] \begin{bmatrix} q_n \\ -p_n \end{bmatrix},
\]

but now \(\omega \epsilon_n\) is a matrix with \(l\) rows and \(m + 1\) columns. Since \([\omega | \omega]\) has rank \(l\) and \([q_n | p_n]\) is non-singular, their product, which is \(\omega \epsilon_n\), has rank \(l\). By the pigeonhole principle, there exists a square sub-matrix \(\omega \epsilon_n^{(-\nu)}\), with \(\nu = (v_1, \ldots, v_l)\) and \(0 \leq v_1 < \cdots < v_l \leq m\), that is non-singular for infinitely many \(n\). For that sub-matrix we have \(\omega \epsilon_n^{(-\nu)} = -(\omega q_n^{(-\nu)} + \omega p_n^{(-\nu)})\), so that its determinant must be a non-zero integer for infinitely many \(n\). For concluding the proof we apply the Binet–Cauchy formula, to obtain

\[
\det \omega \epsilon_n^{(-\nu)} = \sum_\mu \det \omega^{(-\mu)} \det \epsilon_n^{(\mu,v)} \rightarrow 0 \quad (n \rightarrow \infty).
\]

\(\blacksquare\)

**Remark 2.3** The last theorem is designed to be used when \(q_n^{(v)}, p_n^{(\mu,v)}\) are specializations of a system of type II Hermite–Padé approximations to \(m\) functions, see e.g. [34]. In this case, indeed, the non-vanishing of the determinant follows, more or less routinely, by analytic properties that characterize the polynomials involved.

2.4 A refinement

The above criteria can be refined following an idea I learned from Amoroso [3], see also [12] and Remark 2.6 below. A trickier use of this idea led in [22] to a refinement of Nesterenko’s criterion with very interesting applications to the linear independence of zeta values and related numbers. As in (2), let
Vectors of type II Hermite–Padé approximations and a new linear…

but here $p_n^{(\mu,\nu)}$ are rational numbers, and $q_n^{(\nu)}$ is an integer with a large divisor. In order to cope with this subtler situation, let $D_n^{(1)}, \ldots, D_n^{(m)}, \delta_n^{(1)}, \ldots, \delta_n^{(l)}$ be positive integers, and suppose that

$$ \frac{q_n^{(\nu)}}{\delta_n^{(\nu)}} \in \mathbb{Z}, \quad \frac{D_n^{(\mu_2)}}{\delta_n^{(\mu_2)}} \frac{p_n^{(\mu_1,\nu)}}{\delta_n^{(\mu_1,\nu)}} \in \mathbb{Z} \quad (1 \leq \mu_1 \leq \mu_2 \leq m; \nu = 1, \ldots, l). \quad (5) $$

We have the following

**Theorem 2.4** (Refinement of Theorem 2.1) Besides the assumptions above, suppose that for all choices of $l$ (distinct) indices $\mu = (\mu_1, \ldots, \mu_l)$ from 1 to $m$,

$$ \frac{D_n^{(m)}}{\delta_n^{(1)}} \cdots \frac{D_n^{(m-l+1)}}{\delta_n^{(l)}} \det \epsilon_n^{(\mu,-)} \to 0 \quad (n \to \infty). $$

Furthermore, suppose that for all $\lambda^{(i,j)} \in \mathbb{Z} (i = 1, \ldots, l; j = 1, \ldots, m)$ such that the matrix

$$ \lambda = [\lambda^{(i,j)}] \quad i = 1, \ldots, l \\
\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \-quarterspringer
by $D_{n}^{(m-l+i)}$, dividing its $j$th column by $\delta_{n}^{(i)}$ and using (5).

Remark 2.4 In Theorem 2.4 we can change the non-vanishing assumption, identical to that in Theorem 2.1, by replacing it with the non-vanishing assumption in Theorem 2.2. Also, as in the setting of Theorem 2.3 we can enlarge the range for $\nu$ allowing $\nu = 0, \ldots, m$ (i.e.: we have more sequences at our disposal), suppose that

$$\frac{D_{n}^{(m)}}{\delta_{n}^{(v_1)}} \cdots \frac{D_{n}^{(m-l+1)}}{\delta_{n}^{(v_l)}} \det \epsilon_{n}^{(\mu, \nu)} \to 0 \quad (n \to \infty),$$

and that the non-vanishing assumption in Theorem 2.3 holds. Then the conclusion on the dimension of the vector space over $\mathbb{Q}$ spanned by $1, \gamma_1, \ldots, \gamma_m$ holds all the same.

Remark 2.5 Theorem 2.4 above is equivalent to its special case where $\delta_{n}^{(v)} = 1, \nu = 1, \ldots, l$ (simply put $\tilde{q}_{n}^{(v)} = q_{n}^{(v)} / \delta_{n}^{(v)}$ and $\tilde{p}_{n}^{(\mu, \nu)} = p_{n}^{(\mu, \nu)} / \delta_{n}^{(v)}$). We decided to present it in that form in order to stress its meaning in the context outlined in the introduction, in which the sequences $D_{n}^{(\mu)}$ represent the rough estimate of the denominators of the approximations, while the sequences $\delta_{n}^{(v)}$ represent the arithmetical correction arising, e.g., from the permutation group method, or from other methods. On the other hand, the sequences $\epsilon_{n}^{(\mu, \nu)}$ come from a purely analytic construction, without any direct consideration of the denominators. Next section is devoted to obtaining an estimate of $\det \epsilon_{n}^{(\mu, -)}$ in a special situation.

Remark 2.6 The key player in the above theorem are Grassmann’s (or Plücker’s) coordinates (i.e.: the maximal order minors) of the matrix

$$\begin{bmatrix}
q_n \\
p_n
\end{bmatrix}^{(\xi, -)},$$

our assumptions just ensure that they become integers, after multiplication by $\tilde{D}_{n} \in \mathbb{Z}$, and, at the same time, $\tilde{D}_{n} \det \epsilon_{n}^{(\mu, -)} \to 0$. In other words, the last theorem implicitly involves an height of the matrix (6). This height is central in Diophantine geometry: see [5], and Amoroso’s proof of the Nesterenko criterion in [12]. Our criterion can be easily extended, as usual, to the linear independence over an imaginary quadratic extension of $\mathbb{Q}$, and a generalization of it to arbitrary number fields, as in [35, Proposition 4.1], would arguably involve this height.

Moreover, it would be interesting to obtain a quantitative version of our criterion, yielding a linear independence measure.
3 Minors of the Casoratian matrix

3.1 Notation and purpose

Let \( \alpha_n^{(0)}, \ldots, \alpha_n^{(m)} \) be sequences of complex numbers, and we generally assume that \( \alpha_n^{(0)} \alpha_n^{(m)} \neq 0 \). For most applications we have in mind, we also require

\[
\lim_{n \to \infty} \alpha_n^{(j)} = \alpha^{(j)} (j = 0, \ldots, m),
\]

in which case we also assume \( \alpha_0 \alpha_m \neq 0 \). Let \( x_n \) be a sequence of complex numbers satisfying

\[
\alpha_n^{(m)} x_{n+m} + \alpha_n^{(m-1)} x_{n+m-1} + \cdots + \alpha_n^{(0)} x_n = 0.
\]

The coefficients \( \alpha_n^{(m)} \) and \( \alpha_n^{(0)} \) are said to be the highest order and lowest order coefficients of \( (8) \). It is well known that the set of solutions of \( (8) \) is a vector space, and that \( (8) \) can be written as a first-order linear recurrence system. Given \( m \) solutions \( x_n^{(1)}, \ldots, x_n^{(m)} \), they are linearly independent if and only if the Casoratian matrix \[8\], sometimes also called Wronskian by analogy with the differential equation setting,

\[\det x_n\]

is non-singular for some \( n \) (and therefore for any \( n \)). In such a case, any solution of \( (8) \) is a linear combination of \( x_n^{(1)}, \ldots, x_n^{(m)} \) with constant coefficients (i.e.: independent of \( n \)), and \( x_n^{(1)}, \ldots, x_n^{(m)} \) is said to be a basis of solutions of \( (8) \). Also, \( \det x_n \) satisfies the discrete Abel formula (see [1, Problem 2.16.21])

\[
\det x_{n+1} = (-1)^m \frac{\alpha_n^{(0)}}{\alpha_n^{(m)}} \det x_n.
\]

In other words, \( \det x_n \) satisfies a first-order linear recurrence equation, whose highest and lowest order coefficients are, up to the sign, the highest and the lowest coefficient of the linear Eq.\((8)\), respectively. Hence the coefficients of such a recurrence are independent of the particular basis of solutions for \( (8) \). A bit more generally, the coefficients \( (-1)^{m-r} \alpha_n^{(r)} \), for \( r = 0, \ldots, m \), are easily seen to be proportional to

\[
\det \left[ x_n^{(j)} \right]_{i=0, \ldots, r; j=1, \ldots, m} \quad r = 0, \ldots, m,
\]

see [46, §285]. We also recall that, regardless of the Eq.\((8)\), \( r \) sequences \( z_n^{(1)}, \ldots, z_n^{(s)} \) are linearly independent if and only if

\[
\det \left[ z_n^{(j)} \right]_{i=1, \ldots, s; j=1, \ldots, s} \neq 0 \quad \text{for infinitely many } n,
\]

see [8, §7], [46, §279].

The purpose of this section is to study the asymptotic behavior of the \( l \times l \) minors of \( x_n \). We outline a possible strategy to achieve this goal, which is to find a difference equation satisfied by those minors; see [1, Problem 2.16.23] for the case \( l = m - 1 \) with contiguous rows, while
(8) obviously copes with the case \( l = 1 \). Then we explain how to circumvent the difficulties that arise from that method. Before starting with, we briefly recall the most important results by Poincaré, Perron and Pituk about the asymptotic behavior of solutions of (8) satisfying (7).

Minors of the Casoratian (or Wronskian) matrix are key tools in the theory of difference (or differential) equations, with regard to disconjugacy, factorization, discrete Rolle theorem, and several important results on the same vain: see the milestone paper [26]; we refer to [1, Chapter 10] for a wide and (relatively) updated literature, and to [13] for a nice introduction and some perspectives on older results.

### 3.2 A short account on Poincaré-Perron-Pituk’s theorems

Concerning solutions of (8) with the property (7), Poincaré [50] [1, Theorem 2.14.1] proved the following: if the moduli of the roots \( \lambda_1, \ldots, \lambda_m \) of the characteristic polynomial

\[
d'(m) \lambda^m + d'(m-1) \lambda^{m-1} + \cdots + d(0)
\]

are distinct, then either \( x_n = 0 \) for any sufficiently large \( n \), or

\[
\lim_{n \to \infty} \frac{x_{n+1}}{x_n} = \lambda_j \quad \text{for some } j = 1, \ldots, m.
\]

Later on, Perron [43] [1, Theorem 2.14.2] obtained a more precise result: there exists a basis \( x_n^{(1)}, \ldots, x_n^{(m)} \) of solutions of (8) such that

\[
\lim_{n \to \infty} \frac{x_n^{(j)}}{x_n^{(j)}} = \lambda_j \quad \text{for all } j = 1, \ldots, m.
\]

Also, in the more general situation where the moduli \( |\lambda_j| \), and even the roots \( \lambda_j \) themselves, may coincide, Perron [43, 44] proved that there exists a basis of solutions (8) such that

\[
\limsup_{n \to \infty} \sqrt[n]{|x_n^{(j)}|} = |\lambda_j| \quad \text{for all } j = 1, \ldots, m.
\]

In the early 2000’s, Pituk [48] obtained a new limit relation for all nonzero solutions of (8), without any assumption on the moduli \( |\lambda_j| \), and even without assuming anything on \( d_n^{(0)} \) or \( d(0) \):

\[
\lim_{n \to \infty} \sqrt[n]{|x_n| + |x_{n+1}| + \cdots + |x_{n+m-1}|} = |\lambda_j| \quad \text{for some } j = 1, \ldots, m.
\]

The last result is sufficient for certain applications, see [36]. Thus, a rough guess about how to manage with asymptotic behaviors of minors of \( x_n \) is to obtain a difference equation for them. It is also worth mentioning Buslaev’s strengthening of Poincaré’s theorem (see [7]), where again the roots \( \lambda_j \) of the characteristic polynomial need not to be distinct in modulus: for any nonzero solution \( x_n \) of (8)

\[
\limsup_{n \to \infty} \sqrt[n]{|x_n|} = |\lambda_j| \quad \text{for some } j = 1, \ldots, m,
\]

and \( x_n \) satisfies a linear recurrence equation similar to (8), whose monic characteristic polynomial divides the monic characteristic polynomial of (8), and whose characteristic roots are all equal in modulus. As it was remarked by Zudilin [65], this implies that if \( x_n \) is a nonzero solution of (8) such that
Vectors of type II Hermite–Padé approximations and a new linear...

\[ \limsup_{n \to \infty} \sqrt[n]{|x_n|} = |\lambda_1|, \]

and if \(|\lambda_j| \neq |\lambda_1|\) for \(j \neq 1\), then

\[ \lim_{n \to \infty} \frac{x_{n+1}}{x_n} = \lambda_1. \]

### 3.3 The difference equations for the minors

The criteria in Sect. 2 are designed to deal with two different situations: either we have \(l\) linearly independent solutions of (8), or we have a basis of solutions, and select \(l\) solutions within this basis. Here we unify the exposition: if \(x_n^{(1)}, \ldots, x_n^{(m)}\) are \(m\) solutions of (8), for now not necessarily linearly independent, we select the first \(l\) of them and denote

\[ x_n^{(-,l)} = \begin{bmatrix} x_n^{(j)} \\ x_{n+i-1}^{(j)} \end{bmatrix} i = 1, \ldots, m \cdot j = 1, \ldots, l \]

If we take only \(l\) solutions of (8) from the very beginning, the discussion that follows does not change.

For any \(1 \leq \mu_1 < \cdots < \mu_l \leq m\), we pick the square sub-matrix of \(x_n^{(-,l)}\) with the corresponding rows:

\[ x_n^{(\mu,l)} = \begin{bmatrix} x_n^{(j)} \\ x_{n+\mu_i-1}^{(j)} \end{bmatrix} i = 1, \ldots, l \cdot j = 1, \ldots, l \]

We wish to find a linear recurrence equation satisfied by

\[ \det x_n^{(\mu,l)} \cdot \]

To this end, we write

\[ x_{n+1} = \Psi_n x_n, \]

where

\[ \Psi_n = \begin{bmatrix} \Psi_n^{(i,j)} \\ \Psi_n^{(j,i)} \end{bmatrix} i = 1, \ldots, m \cdot j = 1, \ldots, m \]

is the companion matrix of (8):

\[ \Psi_n^{(i,j+1)} = 1 \ (i = 1, \ldots, m - 1), \quad \Psi_n^{(i,j)} = -\frac{\alpha_n^{(j-1)}}{\alpha_n^{(m)}} \ (j = 1, \ldots, m), \quad \Psi_n^{(i,j)} = 0 \text{ otherwise.} \]

It is worth noticing that

\[ \Psi_n^{-1} = \begin{bmatrix} \Psi_n^{-1}^{(i,j)} \\ \Psi_n^{-1}^{(j,i)} \end{bmatrix} i = 1, \ldots, m \cdot j = 1, \ldots, m \]
where
\[
\Psi^{(1,i)}_n = \frac{\alpha^{(j)}_n}{\alpha^{(0)}_n} (j = 1, \ldots, m), \quad \Psi^{(i+1,i)}_n = 1 (i = 1, \ldots, m - 1), \quad \Psi^{(i,i)}_n = 0 \text{ otherwise},
\]
and that if \(\lambda^{(1)}_n, \ldots, \lambda^{(m)}_n\) are the eigenvalues of \(\Psi_n\), then \(w_{n+1} = \Lambda_n z_n\), where
\[
\Lambda^{(i,i)}_n = \lambda^{(i)}_n (j = 1, \ldots, m), \quad \Lambda^{(i,i+1)}_n = 1 (i = 1, \ldots, m - 1), \quad \Lambda^{(i,i)}_n = 0 \text{ otherwise},
\]
and \(z_n\) and \(w_{n+1}\) are defined by \(z^{(i)}_n = x^{(i)}_n\), \(x^{(i)}_n = x^{(i)}_n - \lambda^{(i)}_n x^{(i+1)}_{n+1}\) for \(i = 1, \ldots, m - 1\), and similarly \(w^{(i)}_{n+1} = x^{(i)}_{n+1}\), \(w^{(i)}_{n+1} = x^{(i)}_{n+1} - \lambda^{(i)}_{n+1} x^{(i)}_{n+2}\) for \(i = 1, \ldots, m - 1\). So far, we are not assuming that the Casoratian matrix \(x_n\) is non-singular. Incidentally,
\[
\det x_{n+1} = \det \Psi_n \det x_n,
\]
with
\[
\det \Psi_n = (-1)^n \frac{\alpha^{(j)}_n}{\alpha^{(m)}_n},
\]
gives us the recurrence equation for \(\det x_n\) displayed above, i.e. it settles the case \(l = m\), while (8) obviously copes with the case \(l = 1\). Plainly, \(\det x_n = \det z_n = \det w_n\) and \(\det \Psi_n = \det \Lambda_n = \lambda^{(1)}_n \ldots \lambda^{(m)}_n\).

By induction on \(k\),
\[
x_{n+k} = \Psi_{n+k-1} \ldots \Psi_n x_n,
\]
hence
\[
x^{(\mu,i)}_{n+k} = \left[ \Psi_{n+k-1} \ldots \Psi_n \right]^{(\mu,-)} x^{(-i)}_n.
\]
Here and hereafter, the empty product of matrices is the identity matrix. By the Binet–Cauchy formula,
\[
\det x^{(\mu,i)}_{n+k} = \sum_{\nu} \det \left[ \Psi_{n+k-1} \ldots \Psi_n \right]^{(\mu,\nu)} \det x^{(\nu,i)}_n,
\]
where the sum is over all \(\nu = (\nu_1, \ldots, \nu_l)\) such that \(1 \leq \nu_1 < \cdots < \nu_l \leq m\).

For a fixed \(\mu\), and for each \(k\), we consider the coordinates of \(\det x^{(\mu,i)}_{n+k}\) with respect to \(\det x^{(\nu,i)}_n\), where \(\nu\) varies: these are precisely \(\det \left[ \Psi_{n+k-1} \ldots \Psi_n \right]^{(\mu,\nu)}\). Thus, by letting \(k\) vary from 0 to \(\binom{m}{l}\), it is straightforward to obtain a linear difference equation of order \(\binom{m}{l}\) satisfied by \(y_n = \det x^{(\mu,i)}_n\):
\[
\sum_{k=0}^{\left(\begin{array}{l} m \\ l \end{array}\right)} (-1)^k \det \left[ \Psi_{n+j-1} \ldots \Psi_n \right]^{(\mu,\nu)} \left(\begin{array}{l} m \\ l \end{array}\right) y_{n+k} = 0,
\]
where \(\hat{k}\) means that the index \(k\) is omitted in the range for \(j\). Note that any minor \(y_n = \det x^{(\mu,\nu)}_n\) is a solution of (9), because the coefficients in (9) do not depend on the
choice of the columns \( v_1, \ldots, v_l \), while they do depend on the choice of the rows \( \mu_1, \ldots, \mu_l \).

Also, in (9) and in similar formulas below, unless otherwise stated, we can take any ordering in the set for \( \nu \) (of course, the same each time, in the same formula).

Again keeping \( \mu \) fixed, the \( \left( \begin{array}{c} m \\ l \end{array} \right) \) solutions \( y_n = \det x_n^{(\mu, \nu)} \) of (9) found above, with \( 1 \leq v_1 < \cdots < v_l \leq m \), are linearly independent if and only if the Casoratian matrix

\[
\begin{bmatrix}
\det x_{n+j-1}^{(\mu, \nu)} \\
j = 1, \ldots, \left( \begin{array}{c} m \\ l \end{array} \right) \\
1 \leq v_1 < \cdots < v_l \leq m
\end{bmatrix}
\]

is non-singular. Just as above, we have

\[
\det \left[ \det x_{n+j-1}^{(\mu, \nu)} \right] \\
j = 1, \ldots, \left( \begin{array}{c} m \\ l \end{array} \right) \\
1 \leq v_1 < \cdots < v_l \leq m
\]

\[
= \det \left[ \sum_\nu \det \left[ \Psi_{n+j-2}^{(\mu, \nu)} \cdots \Psi_n^{(\mu, \nu)} \right] \det x_n^{(\nu, \nu)} \right] \\
j = 1, \ldots, \left( \begin{array}{c} m \\ l \end{array} \right) \\
1 \leq v_1 < \cdots < v_l \leq m
\]

\[
= \det \left[ \det \left[ \Psi_{n+j-2}^{(\mu, \nu)} \cdots \Psi_n^{(\mu, \nu)} \right] \right] \\
j = 1, \ldots, \left( \begin{array}{c} m \\ l \end{array} \right) \\
1 \leq v_1 < \cdots < v_l \leq m
\]

\[
\det \left[ \det x_n^{(\nu, \nu)} \right] \\
1 \leq v_1 < \cdots < v_l \leq m
\]

where the Binet formula \( \det AB = \det A \det B \) was used. Here, one more time, the ordering in the set for \( \nu \) (resp. for \( \nu \)) must be the same at each occurrence, while it needs not to be identical for \( \nu \) and \( \nu \), though it would be more consistent; on the other hand, there is no way to choose the same ordering for \( j \) and \( \nu \) (that would just be non-sense). By (19) below,

\[
\det \left[ \det x_{n+j-1}^{(\mu, \nu)} \right] \\
j = 1, \ldots, \left( \begin{array}{c} m \\ l \end{array} \right) \\
1 \leq v_1 < \cdots < v_l \leq m
\]

\[
= \det \left[ \det \left[ \Psi_{n+j-2}^{(\mu, \nu)} \cdots \Psi_n^{(\mu, \nu)} \right] \right] \det x_n^{(\nu, \nu)} \left( \begin{array}{c} m-1 \\ l-1 \end{array} \right)
\]

\[
1 \leq v_1 < \cdots < v_l \leq m
\]

We remark that the highest and the lowest order coefficients in (9), respectively, are
\[ (-1)^j \det \left[ \det [\Psi_{n+j-2} \cdots \Psi_n]^{(\mu, \nu)} \right] \]

\[ j = 1, \ldots, \begin{pmatrix} m \\ l \end{pmatrix} \]

\[ 1 \leq v_1 < \cdots < v_l \leq m \]

and, using the Binet formula and (19) again,

\[ \det \left[ \det [\Psi_{n+j-1} \cdots \Psi_n]^{(\mu, \nu)} \right] \]

\[ j = 1, \ldots, \begin{pmatrix} m \\ l \end{pmatrix} \]

\[ 1 \leq v_1 < \cdots < v_l \leq m \]

\[ = \det \left[ \det [\Psi_{n+j-1} \cdots \Psi_{n+1}]^{(\mu, \nu)} \right] \]

\[ j = 1, \ldots, \begin{pmatrix} m \\ l \end{pmatrix} \]

\[ (\det \Psi_n) \begin{pmatrix} m - 1 \\ l - 1 \end{pmatrix}, \]

in accordance with the discrete Abel formulas for (8) and for (9).

Our conclusion, for this subsection, reads as follows: if the quantity in (10) is nonzero for some \( n \) (thus is so for any \( n \)), and if \( \lambda_n^{(1)}, \ldots, \lambda_n^{(m)} \) is a basis of solutions of (8), then \( \{ \det \lambda_n^{(\mu, \nu)} : 1 \leq v_1 < \cdots < v_l \leq m \} \) is a basis of solutions of (9).

**Remark 3.1** The coefficients of the recurrence Eq. (9) only depend on the coefficients of the recurrence (8), and do not depend on a basis of solutions, nor on a choice for the columns of the minor. To be more precise, since the coefficients of the matrix \( \Psi_n \) are either 0 or, up to the sign, an elementary symmetric function in the eigenvalues \( \lambda_n^{(1)}, \ldots, \lambda_n^{(m)} \),

\[ \text{Sym}^r(\lambda_n^{(1)}, \ldots, \lambda_n^{(m)}) = \sum_{1 \leq v_1 < \cdots < v_l \leq m} \lambda_n^{(v_1)} \cdots \lambda_n^{(v_l)}, \quad r = 0, \ldots, m, \]

where \( \text{Sym}^0(\lambda_n^{(1)}, \ldots, \lambda_n^{(m)}) = 1 \), for all \( \nu \) with \( 1 \leq v_1 < \cdots < v_l \leq m \) we have universal polynomials in \( z_{i,r} \), with \( i = 1, \ldots, \begin{pmatrix} m \\ l \end{pmatrix} - 1 \), \( r = 1, \ldots, m \), with integer coefficients and partial degree not exceeding 1 in each of \( z_{i,r} \), such that their values at \( z_{i,r} = \text{Sym}^r(\lambda_n^{(1)}, \ldots, \lambda_n^{(m)}) \) are the coefficients of the Eq. (9).

**Remark 3.2** We stress that the lowest and highest order coefficients in (8) and in (9) are related by (10), and that the lowest order coefficient for a given \( n \) is, up to a nonzero constant, the highest order coefficient for \( n + 1 \). For this reason, it is sufficient to check the non-vanishing of one of the two (say: the highest order coefficient) for any \( n \), in order to apply the described method.

**Remark 3.3** There is an equivalent way to get the recurrence (9), that we outline here. Let \( \lambda_n^{(1)}, \ldots, \lambda_n^{(m)} \) be the (nonzero) roots of

\[ \alpha_n^{(m)} \lambda^m + \alpha_n^{(m-1)} \lambda^{m-1} + \cdots + \alpha_n^{(0)} = 0, \]

which, essentially, may be supposed to be distinct, as we are going to see. Then the rows of the \( (m + 1) \times (m + 1) \) matrix

\[ \text{Sym}^r(\lambda_n^{(1)}, \ldots, \lambda_n^{(m)}) = \sum_{1 \leq v_1 < \cdots < v_l \leq m} \lambda_n^{(v_1)} \cdots \lambda_n^{(v_l)}, \quad r = 0, \ldots, m, \]
are linearly dependent, so that its determinant vanishes. Here, \( \nu_1, \ldots, \nu_m \) are arbitrarily chosen indices with \( 1 \leq \nu_1 < \cdots < \nu_m \leq m \), so that we have \( \binom{m}{l-1} \) such vanishing determinants, for each \( n \). Each determinant can be expanded with the help of Laplace formula along the first \( l \) columns, to obtain

\[
\sum_{0 \leq \mu_1 < \cdots < \mu_l \leq m} (-1)^{m-\mu_l} \det x_{n+i}^{(j)} = \mu_1, \ldots, \mu_l \det \lambda_n^{(\nu)} = 0, \quad j = 1, \ldots, l
\]

(11)

where \( |\mu| = \mu_1 + \cdots + \mu_l \), and \( \hat{\mu}_1, \ldots, \hat{\mu}_m \) are the complementary indices of \( \mu_1, \ldots, \mu_l \) in \( 0, \ldots, m \). Thus, each sum contains \( \binom{m+1}{l} = \binom{m}{l} + \binom{m}{l-1} \) terms, note, however, that only \( 2 \binom{m}{l} - \binom{m-1}{l} = \binom{m}{l} + \binom{m-1}{l-1} \) of them have a minor of the Casoratian matrix as a factor. By considering consecutive values for \( n \), we have only \( \binom{m}{l-1} \) new terms, where \( \text{new} \) refers to their \( x \)-determinant factor, for each new value of \( n \), and the same number of new equations that correspond to different choices of \( \nu \). Thus, for a fixed \( \mu \) with \( 1 \leq \mu_1 < \cdots < \mu_l \leq m \), taking a linear combination of (11) for \( n, n+1, \ldots, n+\binom{m}{l}-1 \), we get a vanishing linear combination of terms of the type

\[
\det x_{n+i-1}^{(j)} = \mu_1, \ldots, \mu_l \quad j = 1, \ldots, l
\]

only, for \( n, n+1, \ldots, n+\binom{m}{l} \). Finally, we observe that each Eq. (11) can be divided by Vandermonde \( (\lambda_{\nu_1}, \ldots, \lambda_{\nu_m}) \), and after this operation the \( \lambda \)-determinant factors in (11) are replaced with polynomials in \( \text{Sym}^r(\lambda_n^{(1)}, \ldots, \lambda_n^{(m)}) \), and we do not need to assume that \( \lambda_n^{(1)}, \ldots, \lambda_n^{(m)} \) are distinct.

**Remark 3.4** Seemingly, yet another way to obtain the recurrence Eq. (9) is by induction on \( m - l \), using the condensation formula [31, (2.16)].

### 3.4 Recurrences with constant coefficients

Let us consider the special case when the coefficients of Eq. (8) are independent of \( n \):

\[
\alpha^{(m)} x_{n+m} + \alpha^{(m-1)} x_{n+m-1} + \cdots + \alpha^{(0)} x_n = 0,
\]

(12)

and suppose that \( \alpha^{(0)} \alpha^{(m)} \neq 0 \). If the roots \( \lambda_1, \ldots, \lambda_m \) of the polynomial

\[
\alpha^{(m)} \lambda^m + \alpha^{(m-1)} \lambda^{m-1} + \cdots + \alpha^{(0)}
\]

(13)

are distinct, then \( x_n^{(j)} = \lambda_j^n (j = 1, \ldots, m) \) is a basis of solutions of (12), because
\[ \det \begin{bmatrix} x_{n+i-1}^{(j)} \\ \vdots \end{bmatrix} \bigg|_{i=1, \ldots, m} = \det \begin{bmatrix} \lambda_j^{n+i-1} \\ \vdots \end{bmatrix} \bigg|_{j=1, \ldots, m} = (\lambda_1 \cdots \lambda_m)^n \prod_{1 \leq i < j \leq m} (\lambda_j - \lambda_i) \neq 0. \]

(14)

The columns of the matrix in (14) are the eigenvectors of the companion matrix

\[ \Psi = \begin{bmatrix} \Psi^{(ij)} \\ \vdots \end{bmatrix} \bigg|_{i=1, \ldots, m} \bigg|_{j=1, \ldots, m} \]

(15)

of the recurrence Eq. (12), defined by

\[ \Psi^{(i+1)} = 1 \ (i = 1, \ldots, m-1); \quad \Psi^{(m,j)} = -\frac{\alpha^{(j-1)}}{\alpha^{(m)}} \ (j = 1, \ldots, m); \quad \Psi^{(ij)} = 0 \text{ otherwise,} \]

so that

\[ \Psi x_n = x_n \Delta, \]

where \( \Delta = \text{diag} (\lambda_1, \ldots, \lambda_m) \).

This holds in particular when \( \lambda_1, \ldots, \lambda_m \) additionally satisfy

\[ \lambda_{i+1} - \lambda_i = \varepsilon \quad \text{for any} \ i \neq k_1, k_1 + k_2, \ldots, k_1 + \cdots + k_r - 1, \]

where \( k_1 + \cdots + k_r = m \). After a few elementary manipulations on the columns of \( x_n \), dividing by a suitable power of \( \varepsilon \) and making \( \varepsilon \to 0 \) (keeping the \( r \) numbers \( \lambda_{k_1}, \lambda_{k_1+k_2}, \ldots, \lambda_{k_1+\cdots+k_r} \) fixed) in (14), we obtain, by changing the notation, a basis of solutions of (12), which also is a basis of eigenvectors of \( \Psi \), when \( \lambda_1, \ldots, \lambda_r \) are the distinct roots of (13) with multiplicities \( k_1, \ldots, k_r \):

\[ x_n^{(k)} = \binom{n}{k-1} \lambda_1^{n-k+1} \ (k = 1, \ldots, k_1), \]

\[ x_n^{(k_1+k)} = \binom{n}{k-1} \lambda_2^{n-k+1} \ (k = 1, \ldots, k_2), \]

\[ \ldots, x_n^{(k_1+\cdots+k_r)} = \binom{n}{k-1} \lambda_r^{n-k+1} \ (k = 1, \ldots, k_r), \]

because now

\[ \det \begin{bmatrix} x_{n+i-1}^{(j)} \\ \vdots \end{bmatrix} \bigg|_{i=1, \ldots, m} = (\lambda_1^k \cdots \lambda_r^k)^n \prod_{1 \leq i < j \leq r} (\lambda_j - \lambda_i)^{k_j} \neq 0, \]

see [38, pp.174–176], or, for a modern and well informed source, [31, Theorem 20].

Let us suppose, to avoid complications, that \( \lambda_1, \ldots, \lambda_m \) are distinct. We may apply the arguments in Sect. 3.3, and find a difference equation for the minors

\[ y_n = \det x_n^{\mu, \nu} = \det \begin{bmatrix} \lambda_{\nu_1}^{\mu_1-1} \\ \vdots \end{bmatrix} \bigg|_{i=1, \ldots, l} = (\lambda_{\nu_1} \cdots \lambda_{\nu_l})^n \det \begin{bmatrix} \lambda_{\nu_j}^{\mu_j-1} \\ \vdots \end{bmatrix} \bigg|_{j=1, \ldots, l}, \]

namely
If
\[
\begin{bmatrix}
  m \\
  l
\end{bmatrix}
\begin{bmatrix}
  \cdots
  \lambda_{v_1} \cdot \lambda_{v_l}
\end{bmatrix},
\]
\[
\begin{bmatrix}
  m \\
  l
\end{bmatrix}
\begin{bmatrix}
  \cdots
  \lambda_{v_1} \cdot \lambda_{v_l}
\end{bmatrix} \neq 0
\]
for all \( v \), then by (16) the \( \left( \begin{array}{c}
  m \\
  l
\end{array} \right) \) products \( \lambda_0 = \lambda_{v_1} \cdots \lambda_{v_l} \), for \( 1 \leq v_1 < \cdots < v_l \leq m \), are roots of the polynomial

\[
\sum_{k=0}^{m \choose l} (-1)^k \det \left[ \begin{bmatrix}
  \mu \\
  v
\end{bmatrix} \right] \begin{bmatrix}
  \begin{array}{c}
    \cdots
    \lambda_{v_1} \cdot \lambda_{v_l}
  \end{array}
\end{bmatrix},
\]
\[
\begin{bmatrix}
  m \\
  l
\end{bmatrix}
\begin{bmatrix}
  \cdots
  \lambda_{v_1} \cdot \lambda_{v_l}
\end{bmatrix} = 0.
\]

In addition, if \( \lambda_0 \) are all distinct, then the highest and lowest (see remark 3.2) coefficients of (17) are nonzero, because

\[
\det \left[ \begin{bmatrix}
  \mu \\
  v
\end{bmatrix} \right] \begin{bmatrix}
  \begin{array}{c}
    \cdots
    \lambda_{v_1} \cdot \lambda_{v_l}
  \end{array}
\end{bmatrix},
\]
\[
\begin{bmatrix}
  m \\
  l
\end{bmatrix}
\begin{bmatrix}
  \cdots
  \lambda_{v_1} \cdot \lambda_{v_l}
\end{bmatrix} = 1
\]

and

\[
\prod_{1 \leq v_1 < \cdots < v_l \leq m} \det \lambda_{v_1} \cdot \det \lambda_{v_l} \neq 0
\]

\[
\begin{bmatrix}
  m \\
  l
\end{bmatrix}
\begin{bmatrix}
  \cdots
  \lambda_{v_1} \cdot \lambda_{v_l}
\end{bmatrix} = 1
\]

\[
\begin{bmatrix}
  m \\
  l
\end{bmatrix}
\begin{bmatrix}
  \cdots
  \lambda_{v_1} \cdot \lambda_{v_l}
\end{bmatrix} = 0.
\]
Putting this in a different way, by performing the previous trick for all the coefficients of the polynomial \((17)\), we see that, under the non-vanishing assumption 
\[
\det x_0^{\mu,\nu} \neq 0 \quad \text{for all } \nu,
\]
the polynomial \((17)\) is a multiple (by a nonzero coefficient) of
\[
\sum_{k=0}^{m-l} (-1)^k \det [\lambda_j^k]_{j=0, \ldots, \kappa, \ldots, (m/l)}_{1 \leq v_1 < \cdots < v_l \leq m}
\]
where
\[
\det x_0^{\mu,\nu} \neq 0.
\]
Moreover, the Casoratian matrix
\[
\left[ \det x_{n+j-1}^{\mu,\nu} \right]_{j=0, \ldots, (m/l)}_{1 \leq v_1 < \cdots < v_l \leq m}
\]
is non-singular, because
\[
\det \left[ \det x_{n+j-1}^{\mu,\nu} \right]_{j=0, \ldots, (m/l)}_{1 \leq v_1 < \cdots < v_l \leq m} = \prod_{1 \leq v_1 < \cdots < v_l \leq m} \det x_0^{\mu,\nu} \cdot \det [\lambda_j^k]_{j=0, \ldots, (m/l)}_{1 \leq v_1 < \cdots < v_l \leq m} \neq 0.
\]
Clearly, if \(\lambda\) are distinct, then \textit{a fortiori} \(\lambda_j\) are distinct. Note, however, that \(y_n\) satisfies the difference equation
\[
y_{n+q} + \beta^{(q-1)} y_{n+q-1} + \cdots + \beta^{(0)} y_n = 0,
\]
for \(q = \binom{m}{l}\) and \(\beta^{(0)}, \ldots, \beta^{(q-1)}\) defined by
\[
\prod_\nu (\lambda - \lambda) = \lambda^q + \beta^{(q-1)} \lambda^{q-1} + \cdots + \beta^{(0)},
\]
regardless to whether \( \lambda_v \) are distinct or not. If they are not distinct, the minors \( \det J^{\mu,0}_n \) are no longer a basis of solutions of the recurrence (16).

It is fairly possible that in concrete applications of the outlined method in the environment of our criteria in Sect. 2, the assumption that \( \lambda_v \) are distinct is fulfilled. In this case, one can deal with the requirement that (10) does not vanish, by combining the above discussion with (7), and recalling that \( a^{(0)} a^{(m)} \neq 0 \). However, we seek for more generality, specially because the result that we present looks like much more ready-to-use than the recurrence (9). On the other hand, in some cases one may wish to apply Buslaev’s theorem and Zudilin’s corollary described above to (9), which therefore is of some interest by itself.

### 3.5 The Sylvester–Franke theorem

The following fundamental result in the theory of determinants made its appearance in Sect. 3.3, and is crucial in rest of this section.

**Theorem 3.1** (Sylvester–Franke’s theorem [24, 56]) Let \( \rho \) be a \( m \times m \) matrix with entries in \( \mathbb{C} \), and let \( \lambda_1, \ldots, \lambda_m \) be the eigenvalues of \( \rho \), repeated with their algebraic multiplicity. Then for all \( l = 1, \ldots, m \) the eigenvalues of the matrix whose \( \binom{m}{l} \) rows and columns are arranged with the same (say: the lexicographical) ordering, are \( \lambda_{\mu_v} = \lambda_{\mu_1} \cdots \lambda_{\mu_l} \) for \( 1 \leq \mu_1 < \cdots < \mu_l \leq m \), again repeated with their algebraic multiplicity.

In particular

\[
\det \left[ \det \rho^{(\mu, \nu)} \right]_{\mu_1 < \cdots < \mu_l} = (\det \rho)_{\frac{m - 1}{l - 1}}.
\]

**Proof** Let us prove (19), up to the sign, and under the assumption \( \det \rho \neq 0 \). Up to reordering the columns (or the rows) in \( \rho \) we may suppose that all principal minors \( \det \rho^{(k, k)} \), for \( k = 1, \ldots, m \), are nonzero. In this setting one could even determine all the eigenvalues of (18), and, as a result, obtain (18); note, however, that the eigenvalues may change because of the permutation of the rows (or of the columns) in \( \rho \). By assumptions, there exist an upper triangular matrix \( U \), with 1’s on its diagonal, a lower triangular matrix \( L \) with 1’s on its diagonal, and a diagonal matrix \( \Delta \) with the \( \lambda_j \)'s on its diagonal (possibly up to a permutation), such that

\[
\rho = LU = L(\lambda)U = L(\lambda_1) \cdots L(\lambda_m).
\]

By the Binet–Cauchy formula applied twice,

\[
\det \left[ \det \rho^{(\mu, \nu)} \right]_{\mu_1 < \cdots < \mu_l} = \det L^{(\mu, \tau)}_{\mu_1 < \cdots < \mu_l} \det \Delta^{(\tau, \nu)}_{\tau_1 < \cdots < \tau_l} \det U^{(\nu)}_{\nu_1 < \cdots < \nu_l},
\]
where the ranges for $\mu$, $\tau$, $\nu$ and $\nu$ are the same, and the lexicographical ordering is chosen at any occurrence of each multi-index. The last formula displays a product of three matrices, namely: a lower triangular matrix with 1’s on the diagonal, a diagonal matrix with the products $\lambda_\nu$ on its diagonal, and an upper triangular matrix with 1’s on the diagonal. Therefore

$$\det[\det \rho^{(\mu,\nu)}]_\mu = \det[\det \Delta^{(r,\omega)}]_\tau = (\det \Delta^{(r,\omega)})^{(m-1)} l^{-1} = (\det \rho^{(\mu,\nu)})^{(m-1)} l^{-1},$$

as we claimed. In particular, if $\rho$ is non-singular, then its $l$th compound matrix (18) is also non-singular. It could be seen that the products $\lambda_\nu$, which are the eigenvalues of the $l$th compound matrix of $\Delta$, are also the eigenvalues of (18), which would imply our claim in this special case, but we are about to prove it in general.

We now prove that $\lambda_\nu$ are the eigenvalues of (18) without assuming neither $\det \rho \neq 0$, nor the non-vanishing of the principal minors of $\rho$. In $\mathbb{C}$ a passage to the limit would suffice, but we prefer an algebraic proof. Also, the abstract argument in [23] seemingly requires using the axiom of choice, which we do not require here.

Let $\sigma$ be a non-singular square matrix such that

$$\rho \sigma = \sigma J,$$

where $J$ is Jordan’s canonical form of $\rho$, so that $J = \Lambda + \Omega$, where $\Lambda$ is a diagonal matrix with $\lambda_1, \ldots, \lambda_m$ on its diagonal (up to the order), and $\Omega$ is a nilpotent, strictly upper (according to some authors, lower), triangular matrix. Just as above, we have

$$[\det \rho^{(\mu,r)}]_\mu [\det \sigma^{(r,\nu)}]_\nu = [\det \sigma^{(\mu,\nu)}]_\mu [\det \rho^{(\mu,\nu)}]_\nu$$

where

$$[\det \rho^{(\mu,\nu)}]_\nu 1 \leq \nu_1 < \cdots < \nu_l \leq m$$

is an upper triangular matrix with $\lambda_\nu$ on its diagonal, and

$$[\det \sigma^{(r,\nu)}]_\nu 1 \leq \nu_1 < \cdots < \nu_l \leq m$$

is a non-singular matrix by the previous argument. Thus, the eigenvalues of (18) are the same as the eigenvalues of (20), which plainly are the products $\lambda_\nu$, and the theorem is proved.

**Remark 3.5** The matrix (18) is called the $l$th **compound matrix**, or the $l$th **adjugate**, of $\rho$. A proof of the Sylvester-Franke theorem by induction, and several interesting historical notes with a rich bibliography, can be found in [51]. Further proofs are in [60] and [23]. The second part of our proof has intersection with [23] when $\rho$ is diagonalizable, i.e. when $J$ is diagonal. The natural environment of the compound matrices is the exterior algebra $\Lambda^l \mathbb{C}^m$. 

 Springer
Remark 3.6 Rather interestingly, the $L\Delta U$ factorization of the Hessian matrix is a cornerstone of the $\mathbb{C}^N$-saddle point method in [47].

Remark 3.7 If we wish to find the eigenvectors of the compound matrix (18), assuming that we already know the eigenvectors of $\rho$, which are (some of) the columns of $\sigma$ in the above proof, then we are confronted with the entirely combinatorial problem of finding the Jordan normal form of the compound matrix of $J$, which is Jordan’s normal form of $\rho$. The solution of this problem is detailed in [2] and [33].

3.6 Asymptotic behavior of the minors

Pituk [48] considered Poincaré-Perron type difference systems

$$p_{n+1} = [A + B_n]p_n,$$

where $p_n \in \mathbb{C}^m$, the matrix $A \in \mathbb{C}^{m \times m}$ is independent of $n$, and the sequence of matrices $B_n \in \mathbb{C}^{m \times m}$ satisfies

$$\lim_{n \to \infty} \|B_n\| = 0.$$

Here, $\| \cdot \|$ can be any norm on $\mathbb{C}^{m \times m}$.

Putting two theorems together, we have

**Theorem 3.2** (Pituk [48, 49]) Let $\lambda_1, \ldots, \lambda_m$ be the eigenvalues of $A$. If $p_n$ is a solution of (21), then either $p_n = 0$ for any sufficiently large $n$, or

$$\lim_{n \to \infty} \sqrt[n]{\|p_n\|} = |\lambda_j|$$

for some $j = 1, \ldots, m$. (22)

Furthermore, if in (22) we have $p_n \in \mathbb{R}^m_{\geq 0}$, then $|\lambda_j| = \lambda_j \in \mathbb{R}_{\geq 0}$ in (22), and there exists an eigenvector $q$ of $A$ such that $A q = \lambda_j q$ (with the same eigenvalue as in (22)) and $q \in \mathbb{R}^m_{\geq 0}$.

Again, in (22) we can take any norm on $\mathbb{C}^m$. The limit equation we reported about in Sect. 3.2 above was obtained by choosing the $\ell_1$-norm in $\mathbb{C}^m$, $A = \Psi$ and $B_n = \Psi n - \Psi$, where $\Psi_n$ and $\Psi$ are the companion matrices in Sects. 3.2–3.3.

The quoted theorems by Pituk require a very weak assumption on the sequence $B_n$ and essentially no assumption on the matrix $A$, which is very remarkable in comparison with previous results by Perron, Máté and Nevai, Coffman, Li, Trench, Pituk himself and other authors. The second part of the theorem above does not require that $A$ is a nonnegative matrix, but implies a weak form of the Perron-Frobenius theorem about the spectral radius of a nonnegative matrix, see [49, pp. 492–493].

Combining (22) and (19), we get the following

**Theorem 3.3** Let $x_{n}^{(1)}, \ldots, x_{n}^{(l)}$ be linearly independent solutions of (8). Suppose that (7) holds, and let $\lambda_1, \ldots, \lambda_m$ be the eigenvalues of the matrix $\Psi$ in (15). Let

$$x_n = \left[ x_{n+i-1}^{(j)} \right] i = 1, \ldots, m, j = 1, \ldots, l$$

Then
\[
\lim_{n \to \infty} \left\| \left[ \det x_n^{(\mu,-)} \right]_{\mu} \right\| = |\lambda_{v_1} \cdots \lambda_{v_l}| \quad \text{for some } 1 \leq v_1 < \cdots < v_l \leq m.
\]

**Proof** By Sect. 3.1, our first assumption on \(x_n^{(1)}, \ldots, x_n^{(l)}\) implies that

\[
det \left[ x_{n+i-1}^{(l)} \right]_{i = 1, \ldots, l} \neq 0 \quad \text{for infinitely many } n.
\]

We may apply (22) to the system

\[
\left[ \det x_{n+1}^{(\mu,-)} \right]_{\mu} = \left[ \det \Psi_n^{(\mu,v)} \right]_{\mu} \left[ \det x_n^{(\mu,-)} \right]_{\mu},
\]

because

\[
\left[ \det \Psi_n^{(\mu,v)} \right]_{\mu} = \left[ \det \Psi_n^{(\mu,v)} \right]_{\mu} + \left[ \det \Psi_n^{(\mu,v)} - \det \Psi_n^{(\mu,v)} \right]_{\mu}
\]

and

\[
\lim_{n \to \infty} \left\| \left[ \det \Psi_n^{(\mu,v)} - \det \Psi_n^{(\mu,v)} \right]_{\mu} \right\| = 0.
\]

By the Sylvester-Franke theorem, the eigenvalues of

\[
\left[ \det \Psi_n^{(\mu,v)} \right]_{\mu}
\]

are precisely the products \(\lambda_{v_1} \cdots \lambda_{v_l}\), for \(1 \leq v_1 < \cdots < v_l \leq m\). \(\square\)

**Remark 3.8** The above result can be made more precise, using the Jordan normal form of the compound matrix of \(\Psi\), and we refer the reader to our previous remark 3.7.

### 4 Some applications of our criterion

In this section we outline a concrete application of our criterion on two examples. The exposition that follows is a bit sketchy, for two reasons. The first one is that we want to keep the focus of the paper on the criterion itself, and the examples below are merely illustrative. The second reason is that we do not try here to optimize the parameters in the first example, see the end of Sect. 4.1, nor we put special care in the general upper bound for the linear forms, see below. Thus, the experimental results we present here are very likely improvable; in the first example, with the help of the refined criterion, see Theorem 2.4 above, combined with the so-called permutation-group method; in the second example, with a clever estimate of the linear forms.

Let \(\alpha_1, \ldots, \alpha_m \in \mathbb{C}\) be distinct, and let \(k, m, n \in \mathbb{N}\) with \(k \geq 1\). The \(k\)th polylogarithm of \(z\) is defined for \(z \in \mathbb{C}\) with \(|z| < 1\), by

\[
\text{polylog}_k(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^k}.
\]
Vectors of type II Hermite–Padé approximations and a new linear…

\[ \text{Li}_k(z) = \sum_{l=0}^{\infty} \frac{z^l}{k^l}. \]

We put, recursively,

\[ U_0(z) = \prod_{i=1}^{m} (z + \alpha_i)^{kn}, \quad U_j(z) = \frac{1}{n!} \frac{d^n}{dz^n} (z^n U_{j-1}(z)) \hspace{1cm} (j = 1, \ldots, k). \]

The polynomials \( U_j(z) \) have degree \( kmn \). Let \( V_0(z) = z^{kmn} U_0(1/z) \). There exist \( km \) polynomials \( W_{ij}(z) \) with \( i = 1, \ldots, m \) and \( j = 1, \ldots, k \) having degree not exceeding \( kmn \) such that

\[ V_0(z) \text{Li}_j(-\alpha_i z) - W_{ij}(z) = O(z^{kmn+n+1}) \quad (z \to 0). \]  

(23)

In other words, \( V_0, W_{1,1}, \ldots, W_{1,k}, \ldots, W_{m,1}, \ldots, W_{m,k} \) is a system of \((n, \ldots, n)\) type II Padé approximations to \( 1, \text{Li}_1(-\alpha_1 z), \ldots, \text{Li}_k(-\alpha_k z), \ldots, \text{Li}_1(-\alpha_{m-1} z), \ldots, \text{Li}_k(-\alpha_{m-1} z) \) at \( z = 0 \). This is a special case of a more general analytic construction introduced in [15]. The case \( k = 1 \) was introduced in [41] and [52], and the case \( m = 1 \) was introduced, in the more general context of the Lerch functions, in [27]. In particular, the above statement is a special case of [15, Theorem 3.6]. Moreover, the construction can be slightly twisted to obtain a \((km+1) \times (km+1)\) square matrix of polynomials whose determinant does not vanishes for any \( n \), and is independent of \( z \); see [15, Proposition 5.1].

We shall need the following more explicit expressions for the polynomials \( V_0 \) and \( W_{ij} \):

\begin{align*}
V_0(z) &= \sum_{l_1=0}^{kn} \cdots \sum_{l_m=0}^{kn} \left( \frac{n + l_1 + \cdots + l_m}{n} \right)^k \left( \frac{kn}{l_1} \right)^{kn-l_1} \cdots \left( \frac{kn}{l_m} \right)^{kn-l_m} z^{kn-n-l_1-\cdots-l_m} \\
W_{ij}(z) &= \sum_{l_1=0}^{kn} \cdots \sum_{l_m=0}^{kn} \left( \frac{n + l_1 + \cdots + l_m}{n} \right)^k \left( \frac{kn}{l_1} \right)^{kn-l_1} \cdots \left( \frac{kn}{l_m} \right)^{kn-l_m} \\
&\quad \times \frac{1}{s!} \sum_{s=1}^{l_1+\cdots+l_m} (-1)^s \alpha_i^s z^{kn-n-l_1-\cdots-l_m} \hspace{1cm} (i = 1, \ldots, m; j = 1, \ldots, k). 
\end{align*}

Taking into account (23), if \( |\alpha_i| < 1 \) for \( i = 1, \ldots, m \), then

\[ |V_0(1) \text{Li}_j(-\alpha_i) - W_{ij}(1)| \leq (kmn + 1) |\text{Li}_j(1)| |\alpha_i|^{n + \max_{s=1,\ldots,m} \max_{l=0,\ldots,\infty} \left( \frac{n + ml}{n} \right)^k \left( \frac{kn}{l} \right)^{kn-n-l}}. \]

as is easily seen using the absolute convergence of the approximation series.

For a generalization of the main result of [35] to values of the polylogarithm at algebraic points outside the unit disc, we refer the reader to the recent paper [21]. That requires much deeper arguments, which we do not tackle here.

From

\[ \left( \frac{n + ml}{n} \right)^k \left( \frac{kn}{l} \right)^{kn} \leq \left( \frac{n + kmn}{n} \right)^k \left( \frac{kn}{\left[ \frac{kn}{2} \right]} \right)^{kn} \]

we infer that

\[ \text{Li}_k(z) = \sum_{l=0}^{\infty} \frac{z^l}{k^l}. \]
\[ \log |V_0(1)\text{Li}_1(-\pi) - W_{i\pi}(1)| \]
\[ \leq (km + 1)n \log \max_j |\alpha_j| + ((km + 1) \log(km + 1) - km \log(km) + m \log 2)kn + O(\log n), \]
for \( n \to \infty \). This will be used in the second example.

### 4.1 First example

In the first example we focus on the linear independence of
\[
1, \text{Li}_1(1/q), \text{Li}_2(1/q), \text{Li}_1(2/q), \text{Li}_2(2/q)
\]
over \( \mathbb{Q} \) for all sufficiently large positive integers \( q \). Let \( \alpha = -1/q \) and \( \beta = -2/q \), and let
\[
u_1(z; n) = \frac{1}{n!} \frac{d^n}{dz^n} \left( \frac{z^n}{n!} \frac{d^n}{dz^n} \left( \frac{z^n}{n!} \frac{d^n}{dz^n} (z + \alpha)^2(z + \beta)^2 \right) \right),
\]
\[
u_2(z; n) = a \int_0^1 \frac{u(z; n) - u(-at; n)}{z + at} \log t \, dt,
\]
\[
u_2(z; n) = \frac{1}{n!} \frac{d^n}{dz^n} \left( \frac{z^n}{n!} \frac{d^n}{dz^n} \left( \frac{z^n}{n!} \frac{d^n}{dz^n} (z + \alpha)^2(z + \beta)^2 \right) \right),
\]
\[
u_2(z; n) = a \int_0^1 \frac{u(z; n) - u(-at; n)}{z + at} \log t \, dt,
\]
\[
u_2(z; n) = a \int_0^1 \frac{u(z; n) - u(-at; n)}{z + at} \log t \, dt,
\]
\[
u_2(z; n) = a \int_0^1 \frac{u(z; n) - u(-at; n)}{z + at} \log t \, dt,
\]
see [52, p.285] and [29, p.375] for similar formulas. Roughly, when \( q \) is large the linear forms
\[
u_1(z; 1), \text{Li}_1(1/q), \text{Li}_2(1/q), \text{Li}_1(2/q), \text{Li}_2(2/q)
\]
are small. However, the coefficients \( u(1; n), v_1(1; n), v_2(1; n), w_1(1; n), w_2(1; n) \) are rational. It is sufficient to multiply each of them by
\[
q^{2n}d_n^2,
\]
where \( d_n \) is the least common multiple of the integers \( 1, \ldots, n \), to obtain approximations with integer coefficients. More precisely, our refined criterion, i.e. Theorem 2.4 above, is less demanding, but in this first attempt we prefer to skip these nuances. Computations show that using just the same sequences of approximations we experimentally find out that the five numbers \( 1, \text{Li}_1(1/q), \text{Li}_2(1/q), \text{Li}_1(2/q), \text{Li}_2(2/q) \) are linearly independent over \( \mathbb{Q} \) for all \( q \geq 1323 \), and that four out of the same five numbers are linearly independent over \( \mathbb{Q} \) for all \( q \geq 1287 \). The improvement on the range for \( q \) in the second case depends on the asymptotic behavior of the \( 2 \times 2 \) determinants of linear forms, as showed in Theorem 3.3 above. For comparison, we recall that \( 1, \text{Li}_1(1/q), \text{Li}_2(1/q) \) are known to be linear independent over \( \mathbb{Q} \) for \( q \geq 6 \), and \( 1, \text{Li}_1(2/q), \text{Li}_2(2/q) \) are linear independent over \( \mathbb{Q} \) for \( q \geq 51 \), see [54, p.94].

To achieve our plan rigorously, one should compute a linear recurrence relation satisfied by all the coefficients \( u(1; n), v_1(1; n), v_2(1; n), w_1(1; n), w_2(1; n) \). According to the experimental style of this section, we proceed differently. The coefficient \( u(1; n) \) can be easily written as a double Cauchy integral, so that
\[
\lim_{n \to \infty} \frac{1}{n} \log |u(1;n)|
\]
(26)
depends on the critical values of the function
\[
f(s, t) = \frac{s(s - 1/q)^2 (s - 2/q)^2 t}{(s - t)(t - 1)}.
\]
Solving
\[
\frac{\partial}{\partial t} f(s, t) = 0,
\]
we obtain \( s = r^2 \). If we were using the \( C^2 \) saddle point method [30], we would look for the critical values of \( f(r^2, t) = (g(t))^2 \), where
\[
g(t) = \frac{t(t^2 - 1/q)(t^2 - 2/q)}{t - 1}.
\]
If we solve \( g'(t) = 0 \), i.e. we find the roots \( t_1, \ldots, t_5 \) of
\[
4r^5 - 5r^4 - \frac{6}{q} r^3 + \frac{9}{q} r^2 - \frac{2}{q^2} = 0,
\]
and compute
\[
\log |t_i| + \log |t_i^2 - 1/q| + \log |t_i^2 - 2/q| - \log |t_i - 1| + 2 \log q + 4,
\]
for \( i = 1, \ldots, 5 \), the maximum of those values, doubled, is an upper bound for (26), and the second maximum, doubled, is an upper bound for the linear forms
\[
\begin{align*}
q^{4n} d_{4n}^2 u(1;n) Li_1(1/q) &- q^{4n} d_{4n}^2 v_1(1;n), \\
q^{4n} d_{4n}^2 u(1;n) Li_2(1/q) &- q^{4n} d_{4n}^2 v_2(1;n), \\
q^{4n} d_{4n}^2 u(1;n) Li_2(2/q) &- q^{4n} d_{4n}^2 w_2(1;n).
\end{align*}
\]
Suppose that the roots \( t_1, \ldots, t_5 \) are ordered in such a way that
\[
|g(t_1)| > |g(t_2)| > |g(t_3)| > |g(t_4)| > |g(t_5)|.
\]
Note that \( |g(t_2)| = O(|q|^{-\frac{1}{2}}) \) for \( q \to \infty \) by (23). The five numbers
\[
1, Li_1(1/q), Li_2(1/q), Li_1(2/q), Li_2(2/q)
\]
are linearly independent over \( \mathbb{Q} \) if
\[
\log |g(t_2)| + 2 \log q + 4 < 0,
\]
and at least four among the above five numbers are linearly independent over \( \mathbb{Q} \) if
\[
\log |g(t_2)| + \log |g(t_5)| + 4 \log q + 8 < 0.
\]
This explains the difference in the ranges for \( q \) in the two cases.

It is not very difficult to turn the above heuristic argument into a fully rigorous one. Let us outline how to do this. First of all, we need a characterization of the polynomials \( u(z; n) \) in terms of certain orthogonality conditions. This means that
\[
\int_0^1 t^l u(-\gamma t; n)(\log t)^j \, dt = 0 \quad \gamma = \alpha, \beta; \ j = 0, 1; \ l = 0, \ldots, n - 1,
\]
and that any polynomial \( U(z) \) in \( z \) of degree not exceeding \( 4n - 1 \) and satisfying
\[
\int_0^1 t^l U(-\gamma t)(\log t)^j \, dt = 0 \quad \gamma = \alpha, \beta; \ j = 0, 1; \ l = 0, \ldots, n - 1,
\]
must be identically zero. Secondly, we can find five polynomials \( A_0(z; n), \ldots, A_4(z; n) \) in \( z \) and \( n \), not all zero, such that
\[
\deg_z A_0(z; n) \leq 7, \ \deg_z A_1(z; n) \leq 11, \ \deg_z A_2(z; n) \leq 10, \ \deg_z A_3(z; n) \leq 9, \ \deg_z A_4(z; n) \leq 8,
\]
and that
\[
\deg_z \left( \sum_{j=0}^4 A_j(z; n)u(z; n - j) \right) < 4(n - 12).
\]
By the orthogonality conditions above, we get
\[
\sum_{j=0}^4 A_j(z; n)u(z; n - j) = 0.
\]
After dividing by a suitable power of \( n \), this is a Poincaré-Perron-Pituk-type recurrence, and we may apply our results in Sects. 2 and 3. Using the orthogonality conditions again, it is easy to see that the polynomials \( v_1(z; n), v_2(z; n), w_1(z; n) \) and \( w_2(z; n) \) satisfy the same recurrence relation as \( u(z; n) \). Moreover, instead of actually computing the recurrence, maybe with the help of the algorithm in [64] implemented in some computer algebra system, one can also use the explicit form
\[
u(z; n) = \sum_{p=0}^{2n} \sum_{q=0}^{2n} \binom{2n}{p} \binom{2n}{q} \binom{5n - p - q}{n}^2 \alpha^p \beta^q z^{4n-p-q}
\]
to obtain explicitly the limit equation, i.e. (13), of the recurrence, like, e.g., in [36, Theorem 5.1]. We understand that this is just a sketch, but that was all what we promised.

We do not consider the sets of numbers \( 1, L_i(\alpha), L_i(\beta), L_i(\gamma), L_i(\delta) \) when \( (\alpha, \beta) \) is one of
\[
(-1/q, -2/q), \quad (1/q, -2/q), \quad (-1/q, 2/q),
\]
because we experimentally found that in each of these cases \( g(t_2) \) and \( g(t_3) \) happen to be complex conjugate solutions of a polynomial of degree 5, therefore \( |g(t_2)| = |g(t_3)| \), so that our criterion would not have an interesting application. On the other hand, a way to circumvent this difficulty could be the use of approximations more general than those considered above, obtained, e.g. changing \( u(z; n) \) into
\[
\frac{z^{-q_2n}}{(p_2 - q_2)n!} \frac{d^{(p_2 - q_2)n}}{dz^{(p_2 - q_2)n}} z^{(p_2 - q_1)} \frac{d^{(p_1 - q_1)n}}{dz^{(p_1 - q_1)n}} (z^{(p_1)n}(z + \alpha)^m(z + \beta)^m).
\]
This will be the subject of some future paper. We also remark that there is another strategy that is totally independent of consideration of linear recurrence sequences. The linear forms (25) can be written as double integrals, as to the two on the left of (25), or sums of double integrals, as to the two on the right of (25), by using the orthogonality properties of \( u(z; n) \) like in the papers [52] and [29], and then applying the \( \mathbb{C}^2 \)-saddle method. As to the upper bound, in the \( 2 \times 2 \) determinant one can get rid of one of the four double integrals involved, thus carrying the value \( |g(t)| \) into play. For the non-vanishing assumption, it could be handled through a lower bound of the determinant, again using the \( \mathbb{C}^2 \)-saddle method.

4.2 Second example

This example aims to illustrate the refined criterion, i.e. Theorem 2.4. To keep the exposition as simple as possible, we now disregard the contribution that comes from the application of Theorem 3.3, and just use the trivial estimation of the determinants, i.e Hadamard’s inequality. Let us consider the \( km + 1 \) numbers

\[
1, \{ \text{Li}_1(1/lq), \ldots, \text{Li}_k(1/lq) \mid l = 1, \ldots, m \},
\]

where \( q \) is a sufficiently large integer, positive or negative. With the notation at the beginning of the section, we set \( \alpha_i = 1/lq \). A special case of [15, Theorem 2.1] is the following: the \( km + 1 \) numbers above are linearly independent over \( \mathbb{Q} \) whenever

\[
\log |q| > km(k + \log d_m + k \log(5/2)) + k \log 3.
\]  

(28)

We have

\[
d_m^{kn} V_0(1) \in \mathbb{Z}[1/q], \quad d_m^{knn} d_m^l W_{ij}(1) \in \mathbb{Z}[1/q] \quad (j = 1, \ldots, k; i = 1, \ldots, m),
\]

and

\[
\lim_{n \to \infty} \frac{1}{n} \log(d_m^{knn} d_m^{kl}) = km \log d_m + k^2 m.
\]

Combining

\[
q^{knn} d_m^k V_0(1) \in \mathbb{Z}, \quad q^{knn} d_m^{knn} d_m^l W_{ij}(1) \in \mathbb{Z} \quad (j = 1, \ldots, k; i = 1, \ldots, m)
\]

with the above lower bounds for the linear forms, and using the non-vanishing result in [15, Proposition 5.1], we see that the \( km + 1 \) numbers (27) are linear independent over \( \mathbb{Q} \) if

\[
\log |q| > k^2 m + km \log(2d_m) + k(km + 1) \log(km + 1) - k^2 m \log(km),
\]

and, a fortiori, if

\[
\log |q| > k^2 m + km \log(2d_m) + k \log(km + 1) + k,
\]

which improves upon (28). Numerically, for \( k = m = 10 \) we require \( \log |q| \geq 1909 \), and for \( k = m = 11 \) we need \( \log |q| \geq 2717 \); for comparison, in [15, Example 6.1] it is proved, when \( k = m = 10 \) in our notation, that the \( 10^2 + 1 \) numbers in (27) are linearly independent over \( \mathbb{Q} \) whenever \( \log |q| \geq 2715 \). We would like to underline that the approximations we are using here are just the same as in [15, Example 6.1], and the tiny improvement
only comes from a slightly more precise analytic estimate of the linear forms, which could be improved a bit further, e.g., by a refinement of the upper bound in (24). However, our purpose here is not to make previous results more precise, but rather to modulate past and future results in our setting, which allows the dimension of the vector space to be less than its maximal value, in the context of Hermite–Padé approximations of type II. In this section, we can do that by exploiting certain fine arithmetic properties of the coefficients.

Using our Theorem 2.4, we can prove a lower bound for the dimension of the vector space spanned over \( \mathbb{Q} \) by the above numbers, say \( \delta(k, m) \), for a wider range for \( q \). Looking more closely at the definition of the polynomials \( W_{ij}(z) \), we have

\[
d_m^{(k-1)mn}W_{ij}(1) \in \mathbb{Z}[1/q] \text{ if } j \leq k - 1, \quad d_m^{(m-1)mn}W_{ij}(1) \in \mathbb{Z}[1/q] \text{ if } i \leq m - 1.
\]

Thus

\[
d_m^{(k-1)mn}d_{m-1}^{mn}d_{kmn}^{k}W_{ij}(1) \in \mathbb{Z}[1/q] \text{ if } (i, j) \neq (m, k).
\]

Similarly, we have:

- \( \delta(k, m) \geq km \) if \( k, m \geq 2 \) and
  
  \[
  2 \log |q| > 2k^2m + (2k - 1)m \log d_m + m \log d_{m-1} + 2km \log 2 + 2k \log(km + 1) + 2k;
  \]

- \( \delta(k, m) \geq km - 1 \) if \( k, m \geq 3 \) and
  
  \[
  3 \log |q| > 3k^2m + (3k - 2)m \log d_m + 2m \log d_{m-1} + 3km \log 2 + 3k \log(km + 1) + 3k.
  \]

Now let \( k, m \geq 4 \). Refining again the above denominator estimate, we remark that

\[
d_m^{(k-2)mn}d_{m-1}^{mn}d_{m-2}^{mn}d_{kmn}^{k}W_{ij}(1) \in \mathbb{Z}[1/q] \text{ if } (i, j) \neq (m, k), (m - 1, k), (m, k - 1).
\]

Therefore

- \( \delta(k, m) \geq km - 2 \) if \( k, m \geq 4 \) and
  
  \[
  4 \log |q| > 4k^2m + (4k - 3)m \log d_m + 2m \log d_{m-1} + m \log d_{m-2} + 4km \log 2 + 4k \log(km + 1) + 4k;
  \]

- \( \delta(k, m) \geq km - 3 \) if \( k, m \geq 5 \) and
  
  \[
  5 \log |q| > 5k^2m + (5k - 4)m \log d_m + 2m \log d_{m-1} + 2m \log d_{m-2} + 5km \log 2 + 5k \log(km + 1) + 5k;
  \]

- \( \delta(k, m) \geq km - 4 \) if \( k, m \geq 6 \) and
  
  \[
  6 \log |q| > 6k^2m + (6k - 5)m \log d_m + 2m \log d_{m-1} + 3m \log d_{m-2} + 6km \log 2 + 6k \log(km + 1) + 6k.
  \]

The argument can be continued as far as \( k \) and \( m \) are large enough, and each time we can cut a larger triangular corner from the range for \( i \) and \( j \), and get a better, i.e. smaller, denominator outside that corner.

Numerically, \( \delta(11, 11) \geq 1 + 121 - (1 + 2 + 3 + 4) = 112 \) if
\[ 10 \log |q| > 1011^2 + (110 - 9)10 \log(d_{11}) + 55 \log(d_{10}) \\
+ 44(\log d_8) + 1210 \log 2 + 110 \log(122) + 110, \]

i.e. \(|q| > e^{2585} \).

**Remark 4.1** Further examples can be obtained, e.g., with the set of numbers

\[ 1, \{Li_1(1/p^{l-1}q), \ldots, Li_k(1/p^{l-1}q) \mid l = 1, \ldots, m \}, \]

where \(p\) is a positive integer.

**Remark 4.2** It would be interesting to obtain an application of our criterion, or of a suitable quantitative version of it, to the linear independence of values of \(G\)-functions at several points, in the spirit of [11, Theorem 1], whose proof uses Hermite–Padé approximations of type II.

**Acknowledgements** The core of the present note, a short less-than-seven-pages draft without applications, was written several years ago partly during, partly after a stay at the Centre International de Rencontres Mathématiques de Luminy, France. However, my attention to the topic was recently refreshed by the papers [15] and [62], and specially by the proof of [62, Lemma 5.3]. Along the time, I had the pleasure of chatting on this topic with F.Amoroso, M.Laurent and W.Zudilin; a special thank to them, and to whom else made all this possible, in a way or another. I owe a debt of gratitude to an anonymous referee for a very careful reading. Some computations in Sect. 4 were made with the help of the free software Pari/GP [42].

**References**

1. Agarwal, R.P.: Difference equations and inequalities: theory, methods, and applications. 2nd, revised and expanded ed., Pure and Applied Mathematics, Marcel Dekker. 228. New York, xiii+971 (2000)
2. Aitken, A.C.: The normal form of compound and induced matrices. Proc. Lond. Math. Soc. 2(38), 354–376 (1934)
3. Amoroso, F.: Indépendance linéaire décalée, manuscript (2004)
4. Bedulev, E.V.: On the linear independence of numbers over number fields. Mat. Zametki [Math. Notes] 64(44), 506–517 (1998). (440–449)
5. Bombieri, E., Vaaler, J.: On Siegel’s lemma. Invent. Math. 73(1983), 11–32 (1983); addendum: ibid. 75 (1984), 377
6. Bundschuh, P., Töpfer, T.: Über lineare Unabhängigkeit. Monatsh. Math. 117(1–2), 17–32 (1994)
7. Buslaev, V.I.: Relations for the coefficients and singular points of a function. Mat. Sb. 131(173), (1986), 357–384 (Russian); English: Math. USSR Sb. 59(2), 349–377 (1988)
8. Casorati, F.: Il calcolo delle differenze finite interpretato ed accresciuto di nuovi teoremi a sussidio principalmente delle odierne ricerche basate sulla variabilità complessa. Ann. Mat. Pura Appl. (2) 10, 10–43 (1880)
9. Chantanasiri, A.: On the criteria for linear independence of Nesterenko, Fischler and Zudilin. Chandchuri J. Math. 2(1), 31–46 (2010)
10. Chantanasiri, A.: Généralisation des critères pour l’indépendance linéaire de Nesterenko, Amoroso, Colmez, Fischler et Zudilin. Ann. Math. Blaise Pascal 19(1), 75–105 (2012)
11. Chudnovsky, D.V., Chudnovsky, G.V.: Applications of Padé approximations to diophantine inequalities in values of G-functions. In: Chudnovsky D.V., Chudnovsky G.V., Cohn H., Nathanson M.B. (eds.), Number Theory. Lecture Notes in Mathematics 1135, Springer, Berlin, pp. 9–51 (1985)
12. Colmez, P.: Arithmétique de la fonction zêta, in: La fonction zêta. Berline, N., Sabbah, C. (ed). Palais: Les Éditions de l’École Polytechnique, pp. 37–164 (2003)
13. Coppel, W.A.: Disconjugacy, Lecture Notes in Mathematics. 220. Springer, Berlin, p. 147 (1971)
14. Dauguet, S.: Généralisations quantitatives du critère d’indépendance linéaire de Nesterenko. J. Théor. Nombres Bordeaux 27(2), 483–498 (2015)
15. David, S., Hirata-Kohno, N., Kawashima, M.: Can polylogarithms at algebraic points be linearly independent? arXiv: 1912.03811 [math.NT], Mosc. J. Comb. Number Theory 9:4 (2020), 389–406
16. Dubickas, A.: On the approximation of \(\pi/\sqrt{3}\) by rational fractions Vestn. Mosk. Univ., Ser. I 1987, No. 6 (1987), 73–76 (Russian); English: Mosc. Univ. Math. Bull. 42(6), 76–79 (1987)
17. Fischler, S.: Nesterenko’s criterion when the small linear forms oscillate. Arch. Math. 98(2), 143–151 (2012)
18. Fischler, S.: Nesterenko’s linear independence criterion for vectors. Monatsh. Math. 177(3), 397–419 (2015)
19. Fischler, S., Hussain, M., Kristensen, S., Levesley, J.: A converse to linear independence criteria, valid almost everywhere. Ramanujan J. 38(3), 513–528 (2015)
20. Fischler, S., Rivoal, T.: Multiple zeta values, Padé approximation and Vasilyev’s conjecture. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5) 15(15), 24–1 (2016)
21. Fischler, S., Rivoal, T.: Linear independence of values of G-functions, II. Outside the disk of convergence, arXiv: 1811.08758 [math.NT], HAL: hal-01927576, Ann. Math. Qué. (2020)
22. Fischler, S., Zudilin, W.: A refinement of Nesterenko’s linear independence criterion with applications to zeta values. Math. Ann. 347(4), 739–763 (2010)
23. Flanders, H.: A note on the Sylvester-Franke theorem. Am. Math. Monthly 60(8), 543–545 (1953)
24. Franke, E.: Ueber Determinanten aus Unterdeterminanten. J. Reine Angew. Math. 61, 350–355 (1863)
25. Goncharov, A.B.: Multiple polylogarithm, cyclotomy and modular complexes. Math. Res. Lett. 5, 497–516 (1998)
26. Hartman, P.: Difference equations: disconjugacy, principal solutions, Green’s functions, complete monotonicity. Trans. Am. Math. Soc. 246, 1–30 (1978)
27. Hata, M.: On the linear independence of the values of polylogarithmic functions. J. Math. Pures Appl. (9) 69(2), 133–173 (1990)
28. Hata, M.: Rational approximations to \(\pi\) and some other numbers. Acta Arith. 63(4), 335–349 (1993)
29. Hata, M.: Rational approximations to the dilogarithm. Trans. Am. Math. Soc. 336(1), 363–387 (1993)
30. Hata, M.: \(C^2\)-saddle method and Beukers’ integral, ibid. 352 (2000), 4557–4583
31. Krattenthaler, C.: Advanced determinant calculus. Sém. Lothar. Combin. 42, B42q, 67pp (1999)
32. Laurent, M., Roy, D.: Criteria of algebraic independence with multiplicities and approximation by surfaces. J. Reine Angew. Math. 536, 65–114 (2001)
33. Littlewood, D.E.: On induced and compound matrices. Proc. Lond. Math. Soc. 2(40), 370–381 (1935)
34. Mahler, K.: Perfect systems. Compos. Math. 19, 95–166 (1968)
35. Marcovecchio, R.: Linear independence of linear forms in polylogarithms. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5) 5, 1–11 (2006)
36. Marcovecchio, R.: Multiple Legendre polynomials in diophantine approximation. Int. J. Number Theory 10, 1829–1855 (2014)
37. Marcovecchio, R.: Linear independence of polylogarithms at algebraic points. Mosc. J. Comb. Number Theory 6(2–3), 208–232 (2016)
38. Meray, C.: Extension aux équations simultanées des formules de Newton pour le calcul des sommes de puissances semblables des racines des équations entières. Ann. Sci. École Norm. Sup. 1(4), 159–193 (1867)
39. Nesterenko, Yu.V.: On the linear independence of numbers, Vestn. Mosk. Univ. Ser. I, No. 1 [Mosc. Univ. Math. Bull. 40:1] (1985), 46–49 [69–74]
40. Nesterenko, Yu.V.: On a criterion of linear independence of \(p\)-adic numbers. Manuscripta Math. 139(3–4), 405–414 (2012)
41. Nikishin, E.M.: On logarithms of natural numbers, Izv. Akad. Nauk SSSR Ser. Mat. 43:6 (1979), 1319–1327; correction: ibid. 44:4 (1980), 972 [Math. USSR-Izv. 15:3 (1980), 523–530]
42. The PARI Group, PARI/GP version 2.11.0, Université Bordeaux, 2017, http://pari.math.u-bordeaux.fr/
43. Perron, O.: Über die Poincaré’sche lineare Differenzengleichung. J. Reine Angew. Math. 137, 6–64 (1909)
44. Perron, O.: Über Summengleichungen und Poincaré’sche Differenzengleichungen. Math. Ann. 84, 1–15 (1921)
45. Philippon, P.: Critères pour l’indépendance algébrique. Publ. Math. Inst. Hautes Études Sci. 64, 5–52 (1986)
46. Pincherle, S., Amaldi, U.: Le operazioni distributive e le loro applicazioni all’analisi, Bologna, Zanichelli, xii+490 (1901), reprinted at the occasion of the XIX Congress of the U.M.I, Bologna, September 12–17, 2011
47. Pinna, F., Viola, C.: The saddle-point method in $\mathbb{C}^N$ and the generalized Airy functions. Bull. Soc. Math. France 147(2), 221–257 (2019)
48. Pituk, M.: More on Poincaré’s and Perron’s theorems for difference equations. J. Diff. Equ. Appl. 8(3), 201–216 (2002)
49. Pituk, M.: A link between the Perron-Frobenius theorem and Perron’s theorem for difference equations. Linear Algebra Appl. 434, 490–500 (2011)
50. Poincaré, H.: Sur les équations lineaires aux différentielles ordinaires et aux différences finies. Am. J. Math. 7, 203–258 (1885)
51. Price, G.B.: Some identities in the theory of determinants. Am. Math. Monthly 54(2), 75–90 (1947)
52. Rhin, G., Tofflin, P.: Approximants de Padé simultanés de logarithmes. J. Number Theory 24, 284–297 (1986)
53. Rhin, G., Viola, C.: On a permutation group related to $\zeta(2)$. Acta Arith. 77(1), 23–56 (1996)
54. Rhin, G., Viola, C.: Linear independence of $1$, $\operatorname{Li}_1$ and $\operatorname{Li}_2$. Mosc. J. Comb. Number Theory 8(1), 81–96 (2019)
55. Rukhadze, E.A.: A lower bound for the approximations of $\ln 2$ by rational numbers. Vestn. Mosk. Univ. Ser. I Mat. Mekh. 6, (1987), 25–29 (Russian); English: Mosc. Univ. Math. Bull. 42(66), 30–35 (1987)
56. Sylvester, J.J.: On the relations between the minor determinants of linearly equivalent quadratic functions. Phil. Mag. (4) 1(4), 395–405 (1851)
57. Töpfer, T.: Über lineare Unabhängigkeit in algebraischen Zahlkörpern. Result. Math. 25(1–2), 139–152 (1994)
58. Töpfer, T.: An axiomatization of Nesterenko’s method and applications on Mahler functions. J. Number Theory 49(1), 1–26 (1994)
59. Töpfer, T.: An axiomatization of Nesterenko’s method and applications on Mahler functions. II. Compos. Math. 95(3), 323–342 (1995)
60. Tornheim, L.: The Sylvester-Franke Theorem. Am. Math. Monthly 59(6), 389–391 (1952)
61. Viola, C.: On Siegel’s method in diophantine approximation to transcendental numbers. Rend. Semin. Mat. Univ. Politec. Torino 53(4), 455–469 (1995)
62. Viola, C., Zudilin, W.: Linear independence of dilogarithmic values. J. Reine Angew. Math. 736, 193–223 (2018)
63. Waldschmidt, M.: Introduction to Diophantine methods: irrationality and transcendence, notes of the course, Ho Chi Minh University of Natural Sciences, September 12–October 4, 2007. Mission effectuée dans le cadre du PICS Formath Vietnam, 94 pp. (https://webusers.imj-prg.fr/~michel.waldschmidt/coursHCMUNS2007.html, last update: 03/04/2019)
64. Wilf, H.S., Zeilberger, D.: An algorithmic proof theory for hypergeometric (ordinary and “$q$”-) multisum/integral. Invent. Math. 108(3), 575–633 (1992)
65. Zudilin, V.V.: Difference equations and the irrationality measure of numbers. Tr. Mat. Inst. Steklova, Anal. Teor. Chisel i Prilozh. [Proc. Steklov Inst. Math.] 218, 165–178 (1997). ([160–174])
66. Zudilin, W.: Two hypergeometric tales and a new irrationality measure of $\zeta(2)$. Ann. Math. Qué. 38, 101–117 (2014)
67. Zudilin, W.: A determinantal approach to irrationality. Constr. Approx. 45(2), 301–310 (2017)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.