DIAGONALIZATION OF MATRIX SEQUENCES

R.J.Kooman

Mathematical Institute, University of Leiden. Niels Bohrweg 1, Leiden, The Netherlands.

Abstract: In recent years many results have been obtained on the behaviour of solutions of the matrix difference equation $M_n\mathbf{x}_n = \mathbf{x}_{n+1}$ where $\{M_n\}_{n=0}^{\infty}$ is a sequence of $k \times k$-matrices with real or complex entries that are close to diagonal matrices. In this paper we study the question of how to transform an arbitrary sequence $\{M_n\}_{n=0}^{\infty}$ where the entries behave not too wildly, into a sequence of almost-diagonal matrices, so that the results for almost-diagonal matrices can be applied to the difference equation with the transformed sequence. In particular, we will try to find explicit matrices B_n such that the matrices $M'_n = B_{n+1}^{-1}M_nB_n$ are close to diagonal matrices. In the case that the M_n are real 2×2-matrices, a fairly general answer is obtained and it is shown how to proceed for a given sequence $\{M_n\}_{n=0}^{\infty}$. Furthermore, we prove a couple of results that are useful for the case of general order k.

Keywords: difference equations, asymptotic behaviour of solutions, eigenvectors, rescaling, Poincaré-Perron theorem, Levinson theorem, Benzaid-Lutz theorem.

AMS Subject Classification: 15A18, 39A06, 39A22, 41A99.

1 INTRODUCTION

We object of this paper is the study of the asymptotic behaviour of solutions $\{\mathbf{x}_n\}$ of matrix difference equations $M_n\mathbf{x}_n = \mathbf{x}_{n+1}$, where M_n are invertible $k \times k$-matrices with real or complex entries, and the solutions $\{\mathbf{x}_n\}_{n=0}^{\infty}$ are sequences of vectors in $\mathbb{R}^k$ or $\mathbb{C}^k$. The starting points of this type of studies is the Poincaré-Perron theorem, which essentially says that if the sequence $\{M_n\}_{n=0}^{\infty}$ converges to a limit matrix M whose eigenvalues a_j have distinct moduli, then there exist matrices F_n, converging to some invertible matrix F (a matrix of eigenvectors for M) such that $M'_n = F_{n+1}^{-1}M_nF_n$ is a diagonal matrix with entries $a_j(n)$ and where $\lim_{n\to\infty} a_j(n+1)/a_j(n) = a_j$. Obviously, there is a simple relationship between the solutions $\{\mathbf{x}_n\}$ of $M_n\mathbf{x}_n = \mathbf{x}_{n+1}$ and the solutions $\{\mathbf{y}_n\}$ of $M'_n\mathbf{y}_n = \mathbf{y}_{n+1}$, namely $F_n\mathbf{y}_n = \mathbf{x}_n$. More recently, many results have been obtained that are generalizations of the Poincaré theorem. Two results that play a central role in this study are: (i) a result that is known as a Levinson type theorem, named after an analogous result in the theory of differential equations. It says that if the matrices M_n are of the form $A_n + P_n$ where A_n is diagonal with non-zero elements $a_j(n)$ and P_n is a perturbation term such that $\sum_{n=0}^{\infty} \|P_n\|/|a_j(n)|$ converges for $j = 1, \ldots, k$, then there exist matrices F_n, converging to the identity matrix, such that $F_{n+1}^{-1}M_nF_n = A_n$ ([1] and, independently but in somewhat different form [2]), provided that the diagonal numbers satisfy some additional condition, which is usually called "possessing an ordinary dichotomy". We will state and discuss this condition below. In Elaydi's book ([3]) this theorem is called the Benzaid-Lutz theorem. (ii) A second, more recent result in this direction is as follows: if the sequence $\{M_n\}_{n=0}^{\infty}$

is of bounded variation, and the limit matrix M has distinct eigenvalues, then there exist matrices F_n, converging to some matrix F of eigenvectors of M, such that $F_{n+1}^{-1}M_nF_n = A_n$, where A_n is a diagonal matrix of eigenvalues $a_j(n)$ of M_n, provided again that the $a_j(n)$ possess an ordinary dichotomy ([4],[5],[6]). The condition that the eigenvalues are distinct cannot be avoided, but, as we will show below, it can be relaxed.

The aim of this paper is study the question of how to transform a given matrix sequence $\{M_n\}_{n=0}^{\infty}$ into a sequence to which one of the above-mentioned results can be applied. More concretely, we try to find explicitly given matrices B_n such that the transformed sequence $\{M_n' = B_{n+1}^{-1}M_nB_n\}_{n=0}^{\infty}$ is of one of the types mentioned above. Since the matrices B_n can be explicitly calculated, we do not lose information if we study the difference equation $M_n'\mathbf{y}_n = \mathbf{y}_{n+1}$ instead of $M_n\mathbf{x}_n = \mathbf{x}_{n+1}$. Certainly it will be necessary to impose at least some conditions on the matrices M_n. The conditions we impose are in fact of bounded variation type. In section 5 we study the case that the matrices M_n are real and of second order, and sufficiently well-behaved; we explicitly show how to find the matrices B_n.

We now give a short overview of the contents of this paper. The paper is, apart from the introduction, divided into four sections. In section 2 we prove a Levinson-type diagonalization result: in fact, we show that if $M_n = A_n + U_nA_n$, where A_n are invertible diagonal matrices (with some weak condition on its entries), and $\sum_{n=0}^{\infty} U_n$ converges, then $M_n = F_{n+1}^{-1}A_nF_n$ and $\{F_n\}_{n=0}^{\infty}$ converges to the identity matrix. This is a somewhat more general version of what is known as the Benzaid-Lutz theorem and we will need this more general form in order to obtain the results of section 3. Moreover, a bound for $\|F_n - I\|$ is given. In the remainder of the paper, we discuss a couple of cases where we can explicitly find a sequence of matrices $\{B_n\}_{n=0}^{\infty}$ such that the result of the second section can be applied to $B_{n+1}^{-1}M_nB_n$ (instead of M_n). In section 3, B_n are matrices of suitably normalized eigenvectors of M_n, and we study the question what conditions the eigenvectors of M_n have to fulfill in order to transform the sequence $\{M_n\}_{n=0}^{\infty}$ into "almost-diagonal form". Not in all cases will there be such a matrix of eigenvectors: in section 4, we treat a simple case of this type and show a way in which this can be dealt with: here, the matrices B_n appear as "rescaling" matrices that separate the eigenvectors of M_n. This idea of rescaling is used in the last section, where we explicitly show how to diagonalize sequences of (sufficiently well-behaved) 2×2-matrices. A couple of examples is given to show how the method works.

We introduce some conventions and give some definitions.

Firstly, we use a matrix norm ([3],[7]): for a $k \times k$-matrix A, its norm is defined as

$$\|A\| = \max_{\mathbf{x}\in\mathbb{C}^k,\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}.$$

In this definition, $\|\mathbf{x}\|$ denotes some vector norm of $\mathbf{x}$. All such matrix norms are equivalent in the sense that for two such norms $\|\ \|$ and $\|\ \|'$ there exist positive constants C_1, C_2 such that $C_1\|A\| \le \|A\|' \le C_2\|A\|$ for all $k\times k$-matrices A. In this paper, we pick as matrix norm the norm based on the vector norm $\|\mathbf{x}\| = \max_{i=1,\dots,k} |x_i|$ (where x_i is the i-th component of $\mathbf{x} \in \mathbb{C}^k$), i.e.

$$\|A\| = \max_{i=1,\ldots,k} \sum_{j=1}^{k} |A_{ij}|,$$

where A_{ij} are the entries of the matrix A. We now call a vector $\mathbf{x}$ *normalized* if $\|\mathbf{x}\| = 1$ and one of the entries of $\mathbf{x}$ is indeed equal to 1.

Secondly, we denote sequences of matrices (or vectors) $\{M_n\}_{n=n_0}^{\infty}$ simply by $\{M_n\}$ or, even shorter, by $\mathbf{M}$. For the asymptotic behaviour the starting value n_0 is irrelevant and we take $n_0 = 0$ unless stated otherwise. Another way to look at this is to consider only equivalence classes of sequences, where to sequences $\{M_n\}$ and $\{M'_n\}$ are equivalent if $M_n = M'_n$ for $n \geq n_0$ for some n_0. The $k \times k$-identity matrix is denoted by I. We call a (matrix or number) sequence $\{M_n\}$ *of bounded variation* if $\sum_{n=0}^{\infty} \|M_n - M_{n+1}\|$ converges. We call a sequence $\{M_n\}$ of $k \times k$-matrices *almost-diagonal* if there exists some matrix sequence $\{F_n\}$ where $\lim_{n\to\infty} F_n = I$ and $F_{n+1}^{-1} M_n F_n$ are diagonal matrices $\mathrm{diag}(a_1(n), \ldots, a_k(n))$. In this case, for the product matrix

$$M_{n-1} \ldots M_1 M_0 = F_n \, \mathrm{diag} \left(\prod_{h=0}^{n-1} a_1(h), \ldots, \prod_{h=0}^{n-1} a_k(h) \right) F$$

where $F = F_0^{-1}$ is fixed for all n. In particular, the matrix recurrence $M_n \mathbf{x}_n = \mathbf{x}_{n+1}$ has a basis of solutions of the form

$$\mathbf{x}_n^{(i)} = \left(\prod_{h=0}^{n-1} a_i(h) \right) (\mathbf{e}_i + \mathcal{O}(\|F_n - I\|)) \quad (i = 1, \ldots, k), \quad n \to \infty$$

where $\mathbf{e}_i$ is the i-th standard basis vector.

For any sequence of invertible matrices $\{M_n\}$ there exist matrices H_n such that $A_n := H_{n+1}^{-1} M_n H_n$ are diagonal. We use the term *diagonalization* of a matrix sequence $\{M_n\}$ in case the matrices A_n are explicitly known and H_n are the products $B_n F_n$ of explicitly known matrices B_n and matrices F_n which are not explicitly known but where $\lim_{n\to\infty} F_n = I$. In order to guarantee the existence of F_n, the result of the first section will be used.

All sequences of diagonal matrices $\{A_n = \mathrm{diag}(a_1(n), \ldots, a_k(n))\}$ will satisfy the following condition on the coefficients: for all pairs i, j there exist positive constants C_1, C_2 such that

$$\text{either } \prod_{n=\ell}^{m} |a_i(n)|/|a_j(n)| < C_1 \text{ or } \prod_{n=\ell}^{m} |a_i(n)|/|a_j(n)| > C_2 \text{ for all } \ell, m \text{ with } 0 \leq \ell \leq m. \tag{1.1}$$

Of course, both conditions may occur simultaneously. Condition (1.1) occurs in similar circumstances elsewhere in the literature and is usually referred to as "possessing an ordinary dichotomy" ([3]). It is interesting to note that condition (1.1) induces a preordering on sequences: if $\mathbf{a} = \{a(n)\}$ and $\mathbf{b} = \{b(n)\}$ are sequences of non-zero complex numbers, then

$\mathbf{a} \preceq \mathbf{b}$ if for all ℓ, m there exists some positive constant C such that $\prod_{n=\ell}^{m} |a(n)|/|b(n)| < C$. If both $\mathbf{a} \preceq \mathbf{b}$ and $\mathbf{b} \preceq \mathbf{a}$, we write $\mathbf{a} \sim \mathbf{b}$. If $\mathbf{a} \preceq \mathbf{b}$ but not $\mathbf{a} \sim \mathbf{b}$, we write $\mathbf{a} \prec \mathbf{b}$. Notice that in the latter case $\prod_{n=\ell}^{\infty} |\mathbf{a}(n)|/|\mathbf{b}(n)| = 0$. So we can restate condition (1.1) by saying that the sequences $\{a_i(n)\}$ are mutually ordered with respect to the preordering $\preceq$.

Lastly, a remark on notation: (i) we use the notation $\sum_{(n)}$ in the following sense: if a_n are non-negative real numbers such that the series $\sum_{h=n_0}^{\infty} a_h$ converges, then $\sum_{(n)} a_h = \sum_{h=n}^{\infty} a_h$. If, on the other hand, the series diverges, then $\sum_{(n)} a_h = \sum_{h=n_0}^{n-1} a_h$. In all of these cases, either n_0 is given or the notation is used in an order estimate where the exact value of n_0 is irrelevant. (ii) If $A_1, \ldots, A_\ell$ are square $(k_1 \times k_1, \ldots, k_\ell \times k_\ell)$-matrices, then $\mathrm{diag}(A_1, \ldots, A_\ell)$ is the $k \times k$ ($k = k_1 + \ldots + k_\ell$) block diagonal matrix with blocks $A_1, \ldots, A_\ell$, in the given order. (iii) For a $k \times k$-matrix B, we denote by $\mathrm{Diag}_{k_1,\ldots,k_\ell}(B)$ the block diagonal matrix whose elements D_{ij} are equal to B_{ij} if $k_1 + \ldots + k_{p-1} < i, j \leq k_1 + \ldots + k_p$ for some p, and $D_{ij} = 0$ otherwise. The matrix $\mathrm{Diag}_{k_1,\ldots,k_\ell}(B)$ is thus a matrix with the same block structure as A. (iv) If $\{u(n)\}$ is some sequence, then $\Delta u(n) = u(n+1) - u(n)$. (v) We use the Landau symbols: $a(n) = \mathcal{O}(b(n))$ (i.e. $|a(n)/b(n)|$ is bounded) and $a(n) = o(b(n))$ (i.e. $a(n)/b(n) \to 0$ as $n \to \infty$). The condition that $n \to \infty$ is always implicit in this paper.

2 ALMOST-DIAGONAL SEQUENCES.

In this section we prove the following result:

Theorem 2.1 *1. Let $\{a_1(n)\}, \ldots, \{a_k(n)\}$ be sequences of non-zero complex numbers possessing an ordinary dichotomy (i.e satisfying condition (1.1)) and let A_n be the matrix* $\mathrm{diag}(a_1(n), \ldots, a_k(n))$. *Furthermore, let $\{U_n\}$ be a sequence of $k \times k$-matrices such that $\sum_{n=0}^{\infty} \|U_n\|$ converges and $A_n + U_n A_n$ is invertible. Then there exists a sequence of invertible matrices $\{F_n\}$ such that*

$$F_{n+1}^{-1}(A_n + U_n A_n)F_n = A_n$$

and

$$\|F_n - I\| \leq C \cdot \max_{i,j=1,\ldots k} \prod_{h=0}^{n-1} \left|\frac{a_i(h)}{a_j(h)}\right| \cdot \sum_{(n)} \|U_\ell\| \prod_{p=0}^{\ell} \left|\frac{a_j(p)}{a_i(p)}\right| \tag{2.1}$$

for some positive constant C. In particular, $\lim_{n\to\infty} F_n = I$.
2. A similar result as (1) holds if we replace $A_n + U_n A_n$ by $A_n + A_n U_n$.

Theorem 2.1 is a corollary of the following somewhat more general result:

Theorem 2.2 *Suppose that $\{a_1(n)\}, \ldots, \{a_k(n)\}$ be sequences of non-zero complex numbers such that $\{a_1(n)\} \preceq \ldots \preceq \{a_k(n)\}$. Let A_n be the diagonal matrix diag$(a_1(n), \ldots, a_k(n))$.*

Suppose that there exists some number L $(0 \le L < k)$ such that the sequences $\displaystyle\sum_{n=0}^{\infty} \frac{|(U_n)_{ij} \cdot a_j(n)|}{|a_\ell(n)|}$ *converge for all i, for $L < \ell \le k$ and for all $j \le \ell$, then a sequence of $k \times k$-matrices $\{F_n\}$ exists such that*

$$F_{n+1}^{-1}(A_n + U_n A_n)F_n = \operatorname{diag}(A_n' + U_n' A_n', a_{L+1}(n), \dots, a_k(n))$$

where A_n' is the $L \times L$-matrix $\operatorname{diag}(a_1(n), \dots, a_L(n))$ and U_n' is some $L \times L$-matrix with elements $(U_n')_{ij} = \mathcal{O}(\max_{p \ge i} |(U_n)_{pj}|)$. Furthermore,

$$\|F_n - I\| \le C \cdot \max_{i,j=1,\dots k} \prod_{h=0}^{n-1} \left| \frac{a_i(h)}{a_j(h)} \right| \cdot \sum_{(n)} V_\ell \prod_{p=0}^{\ell} \left| \frac{a_j(p)}{a_i(p)} \right| \tag{2.2}$$

for some positive constant C, and where $V_\ell = \displaystyle\max_{p=1\dots k; q \le \ell; \ell > L} |(U_n)_{pq} a_q(n)|/|a_\ell(n)|$ and at least one of the numbers i, j is larger than L.

Proof of Theorem 2.1: 1. For some suitable permutation matrix P, the diagonal elements of $P^{-1}A_nP = \operatorname{diag}(a_1'(n), \dots, a_k'(n))$ are ordered such that $\{a_i'(n)\} \preceq \{a_j'(n)\}$ if $i < j$. So we can always replace A_n and U_n by $P^{-1}A_nP$ and $P^{-1}U_nP$ and thus ensure that $\{a_i(n)\} \preceq \{a_j(n)\}$ if $i < j$. Since $\|U_\ell\| \le c \cdot \displaystyle\max_{p=1\dots k; q \le \ell; \ell > 0} |(U_n)_{pq} a_q(n)|/|a_\ell(n)|$, we conclude that Theorem 2.1 follows from the case $L = 0$ of Theorem 2.2.
2. Let $M_n = A_n(I + U_n)$. We define $M_n' = (I + U_{n+1})A_n$. By (1), there exist matrices F_n, converging to I, such that $F_{n+1}^{-1} M_n' F_n = A_n$. Now let $F_n' = (I + U_n)^{-1} F_n$. Then $(F_{n+1}')^{-1} M_n F_n' = A_n$. Moreover, inequality (2.1) holds for $\|F_n' - I\|$ as it does for $\|F_n - I\|$, if we replace n by $n + 1$ on the right-hand side. □

Theorem 2.2 follows from Theorem 1.4 of [4], which we will, for easy reference, state here again:

Theorem 2.3 *Let $\{A_n\} = \{\operatorname{diag}(a_1(n), \dots, a_k(n)\}$ be a sequence of $k \times k$-invertible diagonal matrices with real or complex entries such that $\{a_i(n)\} \preceq \{a_j(n)\}$ for $i \le j$ and let $\{D_n\}$ be a sequence of real or complex matrices such that $A_n + D_n$ is invertible and $\sum_{n=0}^{\infty}(\|D_n\|/|a_j(n)|)$ converges for $j = L+1, \dots k$ (where L is some integer, $0 \le L \le k-1$). Then there exists a sequence $\{G_n\}$ of real or complex $k \times k$-matrices with $\lim_{n\to\infty} G_n = I$ and*

$$G_{n+1}^{-1}(A_n + D_n)G_n = \operatorname{diag}(P_n + Z_n R_n, a_{L+1}(n), \dots, a_k(n))$$

where P_n is the $L \times L$-matrix consisting of the first L rows and columns of $A_n + D_n$, R_n is the $(k - L) \times L$-matrix consisting of the last $k - L$ rows and the first L columns of D_n and Z_n is a $L \times (k - L)$-matrix with

$$\|Z_n\| \le C' \cdot \sum_{h=0}^{n} \frac{\|D_h\|}{|a_{L+1}(h)|} \prod_{p=h+1}^{n} \left| \frac{a_L(p)}{a_{L+1}(p)} \right|$$

for some constant C'. Furthermore,

$$\|G_n - I\| \leq C \cdot \max_{i,j} \prod_{h=0}^{n-1} \left|\frac{a_i(h)}{a_j(h)}\right| \cdot \sum_{(n)} \frac{\|D_\ell\|}{|a_{L+1}(\ell)|} \prod_{p=0}^{\ell} \left|\frac{a_j(p)}{a_i(p)}\right| \tag{2.3}$$

for some constant C and where the maximum is taken over all pairs i, j such that at least one of the i, j is greater than L.

We now proceed to the proof of Theorem 2.2.

Proof: We apply Theorem 2.3 for $L = k-1$ and let $D_n = U_n A_n$. Since $\sum_{n=0}^{\infty} \|U_n\|$ converges if and only if $\sum_{n=0}^{\infty} |(U_n)_{ij}|$ converges for all i, j, we have indeed that $\sum_{n=0}^{\infty} (\|D_n\|/|a_k(n)|)$ converges. Thus there exists a sequence $\{G_n\}$ converging to I such that

$$G_{n+1}^{-1}(A_n + D_n)G_n = \ \mathrm{diag}(P_n + Z_n R_n, a_k(n))$$

where P_n is the $(k-1) \times (k-1)$-matrix that consists of the first $k-1$ rows and columns of $A_n + U_n A_n$, R_n is the row vector of length $k-1$ containing the last row of the matrix $U_n A_n$ except for the element $(U_n A_n)_{kk}$ (i.e. $(R_n)_j = (U_n)_{kj} a_j(n)$), and Z_n is a column vector of length $k-1$ such that

$$\|Z_n\| \leq C' \cdot \sum_{h=0}^{n} \|U_h\| \prod_{p=h+1}^{n} \left|\frac{a_{k-1}(q)}{a_k(q)}\right|$$

and $\|G_n - I\|$ obeys the inequality (2.3) with $\frac{\|D_\ell\|}{|a_{L+1}(\ell)|}$ replaced by $\max_{p,q} |(U_n)_{pq} a_q(n)|/|a_k(n)|$. In particular, Z_n is bounded and since $(R_n)_j = \mathcal{O}(\|U_n\|) \cdot a_j(n)$, we can write the $(k-1) \times (k-1)$-matrix $P_n + Z_n R_n$ in the form $A'_n + U'_n A'_n$ where $A'_n = \mathrm{diag}(a_1(n), \ldots, a_{k-1}(n))$ and $(U'_n)_{ij} = (U_n)_{ij} + (Z_n)_i (U_n)_{kj}$, so in particular, $\sum_{n=0}^{\infty} |(U'_n)_{ij} a_j(n)|/|a_\ell(n)|$ converges for $j \leq \ell$, $\ell > L$ (as it does for $(U_n)_{ij}$ instead of $(U'_n)_{ij}$). Thus, we can apply Theorem 2.3 again to the matrix $A'_n + U'_n A'_n$ and find a sequence of $(k-1) \times (k-1)$-matrices G'_n such that

$$(G'_{n+1})^{-1} \cdot (A'_n + U'_n A'_n) G'_n = \ \mathrm{diag}(P'_n + Z'_n R'_n, a_{k-1}(n))$$

with P'_n the $(k-2) \times (k-2)$-matrix consisting of the first $k-2$ rows and columns of $A'_n + U'_n A'_n$, Z'_n a bounded row vector and R'_n a column vector of length $k-2$ with $(R'_n)_j = (U'_n)_{k-1,j} a_j(n) = \mathcal{O}(\|U_n\|) \cdot a_j(n)$. So, $P'_n + Z'_n R'_n = A''_n + U''_n A''_n$ where $A''_n = \mathrm{diag}(a_1(n), \ldots, a_{k-2}(n))$ and $(U''_n)_{ij} = (U'_n)_{ij} + (Z'_n)_i (U'_n)_{k-1,j}$. Further, if we define $H_n = \mathrm{diag}(G'_n, 1) \cdot G_n$, we have

$$H_{n+1}^{-1}(A_n + D_n)H_n = \ \mathrm{diag}(A''_n + U''_n A''_n, a_{k-1}(n), a_k(n))$$

and $\|H_n - I\|$ obeys the inequality (2.3) with $\frac{\|D_\ell\|}{|a_{L+1}(\ell)|}$ replaced by $\max_{p,q \leq \ell; \ell \geq k-1} |(U_n)_{pq} a_q(n)|/|a_\ell(n)|$ (and where at least one of the i, j is greater than $k-2$). It is clear that we can repeat this procedure for $L = k-3, \ldots 0$ and in this way we obtain the desired result.

Remark 2.1: Theorem 2.1 and 2.2 are generalizations of the Benzaid-Lutz theorem ([1]). In the Benzaid-Lutz theorem, it is required that $\sum_{n=0}^{\infty}(\|U_n A_n\|)/|a_j(n)|$ converges. Although our version seems only a small generalization, it will be seen that we need this improved version for Theorem 3.2.

3 DIAGONALIZATION BY EIGENVECTORS.

In the remainder of the paper we study the problem of how to transform matrix sequences in such a way that Theorem 2.1 can be applied: more explicitly, given a matrix sequence $\{M_n\}$ we try to find an explicit sequence $\{B_n\}$ such that the corollary can be applied directly to the transformed matrix sequence $\{B_{n+1}^{-1} M_n B_n\}$. A simple idea is to use for B_n a suitable matrix of eigenvectors of M_n (or of M_{n-1}). In this section, we explore under what conditions this can be done. Matrix sequences where this procedure does not work will be studied in the next sections.

Suppose that M_n is a invertible and diagonalizable matrix and let E_n be an invertible matrix of eigenvectors for M_n. Then $A_n := E_n^{-1} M_n E_n$ is diagonal and

$$E_{n+1}^{-1} M_n E_n = A_n + U_n A_n$$

where $U_n = E_{n+1}^{-1} E_n - I$. Hence, if $\sum_{n=0}^{\infty} \|E_{n+1}^{-1} E_n - I\|$ converges, then we can apply Theorem 2.1 and find a matrix sequence $\mathbf{F}$ such that $F_n \to I$ and

$$(E_{n+1} F_{n+1})^{-1} M_n E_n F_n = A_n = \text{ diag}(a_1(n), \ldots, a_k(n)),$$

where $a_1(n), \ldots, a_k(n)$ are the eigenvalues of M_n. We have thus diagonalized the matrix sequence $\mathbf{M}$. Usually, the matrices F_n are not explicitly known but $F_n - I$ can be estimated with the aid of Theorem 2.1. We state this result here:

Proposition 3.1 *Let $\{M_n\}$ a sequence of invertible diagonalizable matrices whose eigenvalues $a_1(n), \ldots, a_k(n)$ possess an ordinary dichotomy. Suppose that matrices of eigenvectors E_n of M_n exist such that $\sum_{n=0}^{\infty} \|E_{n+1}^{-1} E_n - I\|$ converges. Then there exists a sequence of matrices F_n such that $\lim_{n\to\infty} F_n = I$ and*

$$(E_{n+1} F_{n+1})^{-1} M_n E_n F_n = A_n = \text{ diag}(a_1(n), \ldots, a_k(n)).$$

In particular, the condition on the eigenvector matrices is satisfied if $\sum_{n=0}^{\infty} \|E_{n+1} - E_n\|$ converges and $E = \lim_{n\to\infty} E_n$ is invertible. Moreover, inequality (2.1) holds for $\|F_n - I\|$, with $\|U_\ell\|$ replaced by $\|E_{\ell+1}^{-1} E_\ell - I\|$ (or $\|E_{\ell+1} - E_\ell\|$, in the latter case).

Proof: Only the last statement needs comment. It follows from the convergence of $\sum_{n=0}^{\infty} \|E_{n+1} - E_n\|$ that E exists and if E is invertible, then $\lim_{n\to\infty} E_n^{-1} = E^{-1}$. Furthermore,

$$\|E_{n+1}^{-1} E_n - I\| \leq \|E_{n+1}^{-1}\| \cdot \|E_n - E_{n+1}\|.$$

This completes the argument. □

We now turn to the question when, for a matrix sequence, a sequence of eigenvector matrices $\{E_n\}$ exists such that the conditions of Proposition 3.1 are fulfilled. Our main result in this direction is as follows:

Theorem 3.2 *1. Let $A_n = \operatorname{diag}(a_1(n), \ldots, a_k(n))$ be diagonal matrices with distinct diagonal elements $a_i(n)$ such that $\{a_i(n)\}$ are bounded variation sequences. Further, let $\{P(n)\}$ be a bounded variation sequences of $k \times k$-matrices with elements $P_{ij}(n)$ such that $\lim_{n\to\infty} P(n) = O$. Furthermore, suppose that, for some number $j \in \{1, \ldots, k\}$ the sequences $\{P_{ij}(n)/(a_j(n) - a_\ell(n))\}$ are of bounded variation for all i, ℓ with $j \neq \ell$ and converge to zero. Then the matrix $M_n := A_n + P(n)$ has a normalized eigenvector $\mathbf{e}_j(n)$ with eigenvalue $\lambda_j(n)$ such that $\lim_{n\to\infty} \mathbf{e}_j(n) = \mathbf{e}_j$ (the j-th standard basis vector in $\mathbb{C}^n$) and $\{\mathbf{e}_j(n)\}$ is a bounded variation sequence. Furthermore*

$$\|\mathbf{e}_j(n) - \mathbf{e}_j\| \leq c \cdot \max_{i,\ell;j\neq\ell} \left| \frac{P_{ij}(n)}{a_j(n) - a_\ell(n)} \right|, \quad \|\mathbf{e}_j(n) - \mathbf{e}_j(n+1)\| \leq c \cdot \max_{i,\ell;j\neq\ell} \left| \Delta \frac{P_{ij}(n)}{a_j(n) - a_\ell(n)} \right|$$

for some number $c > 0$. In addition, the sequences $\left\{ \frac{\lambda_j(n) - a_j(n)}{a_i(n) - a_j(n)} \right\}$ (where $i \neq j$) converge to zero and are of bounded variation, and

$$|\Delta(\lambda_j(n) - a_j(n))|, \quad \left| \Delta \frac{\lambda_j(n) - a_j(n)}{a_i(n) - a_j(n)} \right| = \mathcal{O}(\max_{i,\ell;j\neq\ell} \left| \Delta \frac{P_{ij}(n)}{a_j(n) - a_\ell(n)} \right|).$$

2. If the conditions on $P_{ij}(n)$ hold for all $j = 1, \ldots, k$, then $E_n = (\mathbf{e}_1(n), \ldots, \mathbf{e}_k(n))$ is a matrix of normalized eigenvectors of M_n such that $\lim_{n\to\infty} E_n = I$ and $\{E_n\}$ is of bounded variation. Furthermore,

$$\|E_n - I\| \leq c \cdot \max_{i,j,\ell;j\neq\ell} \left| \frac{P_{ij}(n)}{a_j(n) - a_\ell(n)} \right|, \quad \|E_{n+1} - E_n\| \leq c \cdot \max_{i,j,\ell;j\neq\ell} \left| \Delta \frac{P_{ij}(n)}{a_j(n) - a_\ell(n)} \right|$$

for some number $c > 0$.

In the proof of the theorem we make use of the following lemma that gives an estimate for eigenvalues and eigenvectors of diagonal matrices with a small perturbation.

Lemma 3.3 *Let $a_1, \ldots, a_k$ be distinct complex numbers and let A be the diagonal $k \times k$-matrix $\operatorname{diag}(a_1, \ldots, a_k)$. Let P, Q be $k \times k$-matrices with $\|P\|, \|Q\| \leq \epsilon s$ and $\|P - Q\| = \delta s$, where $s = \min_{i\neq j} |a_i - a_j|/2$, $0 \leq \epsilon < 1$, $0 \leq \delta < 1$. Then $A + P, A + Q$ are diagonalizable with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ and $\mu_1, \ldots, \mu_k$ respectively, and corresponding normalized eigenvectors $\mathbf{f}_1, \ldots, \mathbf{f}_k$ and $\mathbf{g}_1, \ldots \mathbf{g}_k$, respectively. Moreover,*

$$|\lambda_i - a_i| \leq \|P\|, \ |\mu_i - a_i| \leq \|Q\|$$

and

$$\|\mathbf{f}_i - \mathbf{e}_i\| \le \|P\|/s,\ \ \|\mathbf{g}_i - \mathbf{e}_i\| \le \|Q\|/s.$$

Furthermore, if $\epsilon < \frac{-1+\sqrt{5}}{2}$*, then*

$$|\lambda_i - \mu_i| \le \frac{\delta s}{1-\epsilon-\epsilon^2}$$

and

$$\|\mathbf{f}_i - \mathbf{g}_i\| \le \frac{(1+\epsilon)\delta}{1-\epsilon-\epsilon^2}.$$

Proof: Let P_{ij}, Q_{ij} denote the entries of P and Q, and let $\mathbf{f} = (x_1, \dots, x_n)^T$ be some normalized eigenvector of $A + P$ with eigenvector λ. Let $x_j = 1$. It then follows from

$$(\lambda - a_i)x_i = \sum_{\ell=1}^{k} P_{i\ell}x_\ell$$

that

$$|\lambda - a_j| \le \sum_{\ell=1}^{k} |P_{j\ell}| \le \|P\| < s$$

and, for $i \neq j$,

$$|x_i| \le \left(\sum_{\ell=1}^{k} |P_{i\ell}| \right) /|\lambda - a_i| \le \|P\|/s.$$

In particular, for each $j = 1, \dots, k$ there is an eigenvalue that lies in the disk $\{|z - a_j| < s\}$. Since the disks are disjoint, $A + P$ is diagonalizable. A similar statement holds for $A + Q$. Let j again be fixed. We derive inequalities for $|\lambda_j - \mu_j|$ and for $\|\mathbf{f}_j - \mathbf{g}_j\|$: let $\mathbf{f}_j = (x_1, \dots, x_k)^T$ and $\mathbf{g}_j = (y_1, \dots, y_k)^T$ and $x_j = y_j = 1$. Then

$$\lambda_j - \mu_j = \sum_{\ell=1}^{k} P_{j\ell}(x_\ell - y_\ell) + \sum_{\ell=1}^{k} (P_{j\ell} - Q_{j\ell})y_\ell$$

so that, since $\mathbf{f}_j$ and $\mathbf{g}_j$ are normalized vectors,

$$|\lambda_j - \mu_j| \le \|P\| \cdot \|\mathbf{f}_j - \mathbf{g}_j\| + \|P - Q\|.$$

If $i \neq j$, then

$$x_i = \sum_{\ell=1}^{k} \frac{P_{i\ell}x_\ell}{\lambda_j - a_i},\ \ y_i = \sum_{\ell=1}^{k} \frac{Q_{i\ell}y_\ell}{\mu_j - a_i}$$

so that

$$-y_i + x_i = \sum_{\ell=1}^{k} \frac{(P_{i\ell} - Q_{i\ell})y_\ell}{\mu_j - a_i} + \sum_{\ell=1}^{k} \frac{P_{i\ell}(x_\ell - y_\ell)}{\mu_j - a_i} + \sum_{\ell=1}^{k} \frac{P_{i\ell}x_\ell(\mu_j - \lambda_j)}{(\mu_j - a_i)(\lambda_j - a_i)},$$

so that

$$s \cdot \|\mathbf{f}_j - \mathbf{g}_j\| \le \|P - Q\| + \|P\| \cdot \|\mathbf{f}_j - \mathbf{g}_j\| + \|P\| \cdot |\lambda_j - \mu_j|/s.$$

Combining the inequalities for $|\lambda_j - \mu_j|$ and $\|\mathbf{f}_j - \mathbf{g}_j\|$, we obtain the desired result. □

We are now ready to prove Theorem 3.2:

Proof of Theorem 3.2: Let j be fixed. We introduce matrices $P'(n)$ and $\tilde{P}(n)$ where

$$P'_{i\ell}(n) = \frac{P_{i\ell}(n)}{a_j(n) - a_i(n)}, \quad \tilde{P}_{i\ell}(n) = \frac{P_{j\ell}(n)}{a_i(n) - a_j(n)}$$

where $i \ne j$, and $P'_{j\ell} = \tilde{P}_{j\ell} = 0$. The sequences $\{P'(n)\}$ and $\{\tilde{P}(n)\}$ are bounded variation sequences converging to zero. By Lemma 3.3, there is a unique normalized eigenvector $\mathbf{e}_j(n) = (x_1(n), \ldots, x_k(n))^T$ of M_n which is close to $\mathbf{e}_j$ and has eigenvalue $\lambda_j(n)$ such that $|\lambda(n) - a_j(n)| \le \|P(n)\|$. Hence, $x_j(n) = 1$ if n is sufficiently large. Moreover, for n large enough $|P_{i\ell}(n)| < \frac{1}{4}$ and $|\tilde{P}_{i\ell}(n)| < \frac{1}{4}$. By shifting the indices n by some number n_0, we may assume that the above holds for all $n \ge 0$. For the moment we write $\mathbf{e}, x_i, \lambda$ for $\mathbf{e}_j(n), x_i(n), \lambda_j(n)$, thus leaving out the indices j and n. We approximate $\lambda, \mathbf{e}$ by a sequence of numbers $\{\lambda^{(m)}\}$ and a sequence of vectors $\{\mathbf{e}^{(m)}\}$ such that $\lambda^{(0)} = a_j(n)$, $\mathbf{e}^{(0)} = \mathbf{e}_j$ and

$$\lambda^{(m+1)} = a_j(n) + (P(n)\mathbf{e}^{(m)})_j, \qquad (\lambda^{(m+1)} - a_i(n))x_i^{(m+1)} = (P(n)\mathbf{e}^{(m)})_i \tag{3.1}$$

where $m \ge 0$ and $i \ne j$. Hence

$$x_i^{(m+1)} = (P'(n)\mathbf{e}^{(m)})_i \left(1 + \frac{\lambda^{(m+1)} - a_j(n)}{a_j(n) - a_i(n)}\right)^{-1} = (P'(n)\mathbf{e}^{(m)})_i \left(1 - (\tilde{P}(n)\mathbf{e}^{(m)})_i\right)^{-1} \tag{3.2}$$

if $i \ne j$.

We prove that $\{\lambda^{(m)}\}$ and $\{\mathbf{e}^{(m)}\}$ converge to the eigenvalue λ and eigenvector $\mathbf{e}$ of M_n. From (3.1) it follows that, for $i \ne j$,

$$|\lambda^{(m)} - \lambda^{(m+1)}| \le \|P(n)\| \cdot \|\mathbf{e}^{(m)} - \mathbf{e}^{(m-1)}\|$$

and

$$|\lambda^{(m)} - \lambda^{(m+1)}| \le |a_j(n) - a_i(n)| \cdot \|\tilde{P}(n)\| \cdot \|\mathbf{e}^{(m)} - \mathbf{e}^{(m-1)}\|$$

whence it follows from (3.2) that

$$\|\mathbf{e}^{(m)} - \mathbf{e}^{(m+1)}\| \le 2\|P'(n)\| \cdot \|\mathbf{e}^{(m-1)} - \mathbf{e}^{(m)}\| + 4\|P'(n)\| \cdot \|\tilde{P}(n)\| \cdot \|\mathbf{e}^{(m)} - \mathbf{e}^{(m-1)}\| \le$$

$$\le \frac{3}{4}\|\mathbf{e}^{(m)} - \mathbf{e}^{(m-1)}\|.$$

Hence, $\{\mathbf{e}^{(m)}\}$ and $\{\lambda^{(m)}\}$ converge to some vector $\mathbf{e}$ and some number λ and, by (3.1),

$$\lambda = a_j(n) + (P(n)\mathbf{e})_j, \quad (\lambda - a_i(n))x_i = (P(n)\mathbf{e})_i,$$

i.e. $\mathbf{e}$ is a normalized eigenvector of M_n with eigenvalue λ. Moreover,

$$\|\mathbf{e}-\mathbf{e}_j\| = \|\mathbf{e}-\mathbf{e}^{(0)}\| \leq \sum_{m=0}^{\infty} \|\mathbf{e}^{(m)}-\mathbf{e}^{(m+1)}\| \leq 4\cdot\|\mathbf{e}^{(1)}-\mathbf{e}^{(0)}\| \leq 8\cdot\|P'(n)\|.$$

We now show that the sequence $\{\mathbf{e}_j(n)\}$ is of bounded variation. From now on, we write the index n (but not j) explicitly, i.e. we write $\mathbf{e}(n)$ for $\mathbf{e}$ and let the numbers $\lambda^{(m)}=\lambda^{(m)}(n)$, $\mathbf{e}^{(m)}=\mathbf{e}^{(m)}(n)$ be defined by equations (3.1).
By equations (3.1),

$$\Delta(\lambda^{(m+1)}(n)-a_j(n)) = \Delta(P(n)\mathbf{e}^{(m)}(n))_j = (\Delta P(n)\cdot\mathbf{e}^{(m)}(n+1))_j + (P(n)\Delta\mathbf{e}^{(m)}(n))_j$$

so that

$$|\Delta(\lambda^{(m+1)}(n)-a_j(n))| \leq \|\Delta P(n)\| + \|P(n)\|\cdot\|\Delta\mathbf{e}^{(m)}(n)\|$$

and similarly

$$\Delta\left(\frac{\lambda^{(m+1)}(n)-a_j(n)}{a_i(n)-a_j(n)}\right) = (\Delta\tilde{P}(n)\cdot\mathbf{e}^{(m)}(n+1))_j + (\tilde{P}(n)\Delta\mathbf{e}^{(m)}(n))_j$$

so that

$$\left|\Delta\left(\frac{\lambda^{(m+1)}(n)-a_j(n)}{a_i(n)-a_j(n)}\right)\right| \leq \|\Delta\tilde{P}(n)\| + \|\tilde{P}(n)\|\cdot\|\Delta\mathbf{e}^{(m)}(n)\|. \tag{3.3}$$

Furthermore, by (3.2), for $i\neq j$,

$$\|\Delta\mathbf{e}^{(m+1)}(n)\| \leq 2\|\Delta P'(n)\| + 2\|P'(n)\|\left(\|\Delta\mathbf{e}^{(m)}(n)\| + 2\|\Delta\tilde{P}(n)\| + 2\|\tilde{P}(n)\|\cdot\|\Delta\mathbf{e}^{(m)}(n)\|\right)$$

whence it follows, by induction on m, that the sequences $\{\mathbf{e}^m(n)\}$, $\{\lambda^{(m)}(n)\}$ and also $\left\{\dfrac{\lambda^{(m)}(n)-a_j(n)}{a_i(n)-a_j(n)}\right\}$ are of bounded variation for all m. Moreover, let c be some number such that $c\geq 8$ and

$$\|\Delta\mathbf{e}^{(0)}\| \leq c(\|\Delta P'(n)\| + \|\Delta\tilde{P}(n)\|)$$

then we can prove by induction to m that

$$\|\Delta\mathbf{e}^{(m)}\| \leq c(\|\Delta P'(n)\| + \|\Delta\tilde{P}(n)\|)$$

holds for all m. Indeed, since both $\|P'(n)\| < \frac{1}{4}$ and $\|\tilde{P}(n)\| < \frac{1}{4}$, it follows from the above inequalities that

$$\|\Delta\mathbf{e}^{(m+1)}(n)\| \leq 2\|\Delta P'(n)\| + 2\|\Delta\tilde{P}(n)\| + \frac{3}{4}\|\Delta\mathbf{e}^{(m)}\| \leq$$
$$\leq (2+\frac{3}{4}c)(\|\Delta P'(n)\| + \|\Delta\tilde{P}(n)\|) \leq c(\|\Delta P'(n)\| + \|\Delta\tilde{P}(n)\|). \tag{3.4}$$

By dominated convergence, it now follows that the limit sequence $\{\mathbf{e}(n)\}$ is of bounded variation, and inequality (3.4) holds for $\|\Delta\mathbf{e}(n)\|$; from (3.3) it follows that also the sequences $\left\{\dfrac{\lambda(n)-a_j(n)}{a_i(n)-a_j(n)}\right\}$ and $\{\lambda(n)\}$ are of bounded variation and that the order estimates on $\|\mathbf{e}(n)-\mathbf{e}_j\|$, $\|\mathbf{e}(n)-\mathbf{e}(n+1)\|$ and $\left|\Delta\dfrac{\lambda(n)-a_j(n)}{a_i(n)-a_j(n)}\right|$ hold true. This completes the argument.

2. This is an immediate result of (1). □

The case that the diagonal elements $a_j(n)$ converge to distinct numbers $a_1,\dots,a_k$ is worth a special mention; it is in this form that we shall use Theorem 3.2 in section 5 of this paper:

Corollary 3.4 *(i.) Let M be a diagonalizable matrix with distinct eigenvalues and let $\{P_n\}$ be a sequence of $k\times k$-matrices of bounded variation that converges to the zero matrix. Let $a_1(n),\dots,a_k(n)$ be the eigenvalues of $M_n=M+P_n$ and suppose that $\{a_i(n)\}\preceq\{a_j(n)\}$ whenever $i<j$. Furthermore, let $\{Q_n\}$ be a sequence of $k\times k$-matrices such that $\sum_{n=0}^{\infty}\|Q_n\|$ converges. Then there exists some sequence $\{H_n\}$ with $\lim H_n=H$ such that H is a non-singular matrix of eigenvectors of M and*

$$H_{n+1}^{-1}(M_n+M_nQ_n))H_n=\ \mathrm{diag}(a_1(n),\dots,a_k(n))$$

and

$$\|H_n-H\|\le C\cdot\max_{i,j=1,\dots k}\prod_{h=0}^{n-1}\left|\frac{a_i(h)}{a_j(h)}\right|\cdot\sum_{(n)}(\|P_\ell-P_{\ell+1}\|+\|Q_\ell\|)\prod_{p=0}^{\ell}\left|\frac{a_j(p)}{a_i(p)}\right|. \tag{3.5}$$

The sequences $\{a_i(n)\}$ are of bounded variation.

(ii.) Let M,P_n be as in (i) and suppose that $\{b(n)\}$ and $\{c(n)\}$ are bounded variation sequences of real or complex numbers such that the matrices $N_n=b(n)I+c(n)M_n$ are invertible. Then some sequence $\{\tilde H_n\}$ with $\lim\tilde H_n=H$ can be found (where H is, as in (i), a non-singular matrix of eigenvectors of M) and

$$\tilde H_{n+1}^{-1}(N_n+N_nQ_n))\tilde H_n=b(n)I+c(n)\ \mathrm{diag}(a_1(n),\dots,a_k(n))$$

such that inequality (3.5) holds, with $a_q(h)$ replaced by $b(h)+c(h)a_q(h)$ for $q=i,j$, provided that the sequences $\{b(n)+c(n)a_j(n)\}$ possess an ordinary dichotomy.

Proof: (i.) The result follows by applying Theorem 3.2 to $A_n=H^{-1}MH=\mathrm{diag}(a_1,\dots,a_k)$. It follows that a non-singular matrix of normalized eigenvectors E_n of $H^{-1}M_nH$ exists such that $\|E_n-I\|=\mathcal{O}(\|P_n\|)$ and $\|E_n-E_{n+1}\|=\mathcal{O}(\|P_n-P_{n+1}\|)$. By Proposition 3.1, a sequence $\{K_n\}$ of $k\times k$-matrices exists that converges to H and such that

$$K_{n+1}^{-1}(M+P_n)K_n=\ \mathrm{diag}(a_1(n),\dots,a_k(n)).$$

We can now apply Theorem 2.1 to the sequence $\{K_{n+1}^{-1}M_nK_n(I+K_n^{-1}Q_nK_n)\}$. There exist matrices F_n such that $F_n \to I$ and

$$F_{n+1}^{-1} \cdot \text{ diag}(a_1(n), \ldots, a_k(n))(I + K_n^{-1}Q_nK_n)F_n = \text{ diag}(a_1(n), \ldots, a_k(n)).$$

Finally the matrices H_n are defined by $H_n = K_nF_n$.

(ii.) This follows immediately from (i) and Proposition 3.1 since the matrices of eigenvectors E_n of N_n are the same as for M_n. □

Remark 3.1: The case that the eigenvalues of the limit matrix M are distict, as in Corollary 3.4(i), has been studied before. The result is therefore, apart from some slight generalization on the perturbation terms Q_n, not new; it already appears in e.g. [4],[5]. We will make much use of the result in section 5 of this paper.

We conclude this section with two examples of situations where Theorem 3.2 can be applied. We state the examples in the form of propositions. The first example is a direct application and does not need further comment. It involves a diagonal sequence plus a perturbation:

Proposition 3.5 *Let $\{f(n)\}$ be some bounded variation sequence of positive numbers such that $\lim_{n\to\infty} f(n) = 0$, let A_n be a diagonal matrix diag$(a_1(n), \ldots, a_k(n))$ with* distinct *elements $a_j(n) = \alpha_{0j} + \alpha_{1j}f(n) + \ldots + \alpha_{\ell j}f(n)^\ell$. Furthermore, let $\{P_n\}$ be a sequence of $k \times k$-matrices such that $\{P_n/f(n)^\ell\}$ is of bounded variation and $\lim_{n\to\infty}(P_n/f(n)^\ell) = O$. (In particular, this is the case if P_n is a converging power series $\sum_{m=\ell+1}^{\infty} f(n)^m B_m$ where B_m are $k \times k$-matrices). Then $M_n = A_n + P_n$ has a matrix of normalized eigenvectors E_n such that*

$$\|E_n - I\| = \mathcal{O}(\|P_n\|/f(n)^\ell), \quad \|E_n - E_{n+1}\| = \mathcal{O}(\Delta(\|P_n\|/f(n)^\ell)).$$

The next example is somewhat similar to the previous one but now there is a non-diagonal term apart form the perturbation:

Proposition 3.6 *Let A be a diagonal $k \times k$-matrix, B a $k \times k$-matrix and let $\{f(n)\}$ be a bounded variation sequence of positive real numbers such that $\lim_{n\to\infty} f(n) = 0$ and such that $\{f(n+1)/f(n)\}$ is of bounded variation and $\lim_{n\to\infty} \frac{f(n+1)}{f(n)} = 1$. Lastly, let $\{P_n\}$ be a bounded variation sequence of $k \times k$-matrices such that $\lim_{n\to\infty} P_n = O$. Suppose that the eigenvalues $\alpha_i + y\beta_i + o(y)$ of $A + yB$ are distinct up to order y (i.e. if $\alpha_i = \alpha_j$ then $\beta_i \neq \beta_j$ if $i \neq j$). Then the matrices $M_n = A + f(n)B + f(n)P_n$ have eigenvector matrices E_n such that $\{E_n\}$ is of bounded variation and $\lim_{n\to\infty} E_n = E$ where E is some invertible matrix. Moreover,*

$$\|E_n - E\| = \mathcal{O}(\|P_n\|) + \mathcal{O}(f(n)) + \mathcal{O}(\frac{f(n+1)}{f(n)} - 1),$$

$$\|E_n - E_{n+1}\| = \mathcal{O}(\|\Delta P_n\|) + \mathcal{O}(\Delta f(n)) + \mathcal{O}(\Delta(\frac{f(n+1)}{f(n)} - 1)).$$

In the proof, we use the following result:

Lemma 3.7 *Let A, B be $k \times k$-matrices, with A a diagonal matrix of the form $A = diag(a_1 I_{k_1}, \ldots, a_\ell I_{k_\ell})$, where $a_1, \ldots, a_\ell$ are distinct complex numbers. There exists a matrix C such that $B + AC - CA$ has zero elements outside the ℓ blocks that correspond to the block structure of A.*

Proof of the lemma: Let V be the real vector space spanned by the $2k^2$ matrices E_{pq}, iE_{pq}, where E_{pq} is the matrix with elements $(E_{pq})_{rs} = \delta_{pr}\delta_{qs}$. These $2k^2$ matrices form an orthonormal basis with respect to the inner product $\langle C, D\rangle = \mathrm{tr}(C^T D)$ on V. Now let $\phi : V \to V$ be the linear map $\phi(C) = CA - AC$. It is straightforward to check that ϕ is antisymmetric, i.e. $\langle \phi(C), D\rangle = -\langle C, \phi(D)\rangle$. Hence, the kernel and range of ϕ are the orthogonal complement of each other (see e.g. [7]). $\mathrm{Ker}(\phi)$ consists of the matrices that commute with A, i.e. the block diagonal matrices of the form $\mathrm{diag}(C_1, \ldots, C_\ell)$ with C_j a $k_j \times k_j$-matrix. This proves the argument. □

Proof of Proposition 3.6: By permuting the diagonal elements of A, we may suppose that A is of the form as in the lemma. Let

$$A + yB' = \mathrm{diag}(a_1 + y\beta_{11}, \ldots, a_1 + y\beta_{1k_1}, \ldots, a_\ell + y\beta_{\ell k_\ell})$$

be a diagonal matrix with the permuted numbers $\alpha_i + y\beta_i$ on the diagonal. Now, the eigenvalues of $A+yB$ are algebraic functions of y and can be developed in a Puiseux series ([8]) in powers of $y^{1/m}$ for some m. By Lemma 3.7 there exists a matrix C such that

$$(I + yC)^{-1}(A + yB)(I + yC) = A + yB'' + \mathcal{O}(y^2)$$

where $B'' = \mathrm{Diag}_{k_1,\ldots,k_\ell}(B)$ commutes with A: B'' has the same block structure as A. We show that B'' is similar to B'.

Obviously, $A + yB$ is similar to the matrix $A + yB'' + \mathcal{O}(y^2)$. Since the eigenvalues of $A + yB$ are distinct numbers modulo $y^{1+1/m}$, it follows, by applying the principle of the argument from complex analysis ([9]) to the characteristic polynomial of $A+yB''$, that the eigenvalues of $A + yB''$ and $A + yB$ are the same even modulo y^2. But then $A + yB''$ has the same eigenvalues as $A + yB'$ modulo y^2. Since both $A + yB'$ and $A + yB''$ are block diagonal matrices with blocks $a_j I + yB'_j$ and $a_j I + yB''_j$ ($j = 1, \ldots, \ell$), the blocks B'_j and B''_j have the same eigenvalues. Because the eigenvalues of B'_j are distinct, B'_j and B''_j, and hence, also $A + yB''$ and $A + yB'$, are similar. In particular, there exists some invertible matrix E such that $E^{-1}(A + yB'')E = A + yB'$. Hence,

$$E^{-1}(I + yC)^{-1}(A + yB)(I + yC)E = A + yB' + \mathcal{O}(y^2)$$

We can now apply Theorem 3.2 to the matrix sequence $\{M'_n\}$ where

$$M'_n = E^{-1}(I+f(n+1)C)^{-1}M_n(I+f(n)C)E = A+f(n)B''+f(n)E^{-1}P_nE+\mathcal{O}(\Delta f(n))+\mathcal{O}(f(n)^2). \square$$

Remark 3.2: It is not possible to generalize Proposition 3.6 in the obvious way, taking more orders of y than only up to first order. Consider the following example: let A_0, A_1, A_2 be defined by

$$A_0 + yA_1 + y^2A_2 = \begin{pmatrix} 1+y & y \\ 0 & 1+y+y^2 \end{pmatrix}.$$

Hence, $A_0 = I$ is diagonal and $A_0 + yA_1 + y^2A_2$ has eigenvalues $1+y$ and $1+y+y^2$ that are distinct up to order y^2. The matrix of normalized eigenvectors is $E(y) = \begin{pmatrix} 1 & 1 \\ 0 & y \end{pmatrix} = B_0 + yB_1$ and

$$(A_0 + yA_1 + y^2A_2)E(y) = (A_0 + yA_1 + y^2A_2)(B_0 + yB_1) = (B_0 + yB_1)(A_0 + yA_1' + y^2A_2')$$

where $A_1' = \operatorname{diag}(y, y)$ and $A_2' = \operatorname{diag}(0, y^2)$. Here $B_0 = E(0)$ is not invertible: in fact, although A_1 and A_1' have the same eigenvalues, A_1 is non-diagonalizable, whereas A_1' is. Hence, A_1 and A_1' are not similar matrices and the argument in the proof of Proposition 3.6 does not hold.

4 AN EXAMPLE WITH NON-DIAGONIZABLE MATRICES.

In the previous section, the eigenvectors of the matrices M_n are far apart in $\mathbb{C}^n$: M_n have matrices of (normalized) eigenvectors that are close to some non-singular matrix. In this section we study a type of matrix sequence that is in a sense the opposite of the sequences we discussed in the previous chapter and where the matrices are non even diagonalizable. The most simple example is that $\{M_n\} = \{M\}$ is a constant sequence and M is non-diagonalizable. This sequence can be analyzed completely because all the powers M^n can be calculated explicitly (provided that the eigenvalues are known).

As is well known, any matrix is similar to a matrix in Jordan normal form. A Jordan normal form is a block diagonal matrix where the blocks are of the form $aI + J$ where a is some complex number, and $J_{ij} = 1$ if $j - i = 1$ and $J_{ij} = 0$ otherwise. (see e.g. [3],[7]). If $a \neq 0$, such a block matrix is similar to $a(I + J)$. Let us now suppose that $A_n = M = \operatorname{diag}(m_1, \ldots, m_\ell)$ where $m_i = a_i(I + J)$, $a_i \neq 0$ and m_i, I, J are $k_i \times k_i$-matrices. Let L be the maximum of the k_i. If $L > 1$, then M is not diagonalizable. Further, we let $M_n = M + P_n$. We shall show that $\{M_n\}$ can be diagonalized with the aid of Corollary 2.2 if $\sum_{n=1}^{\infty} n^{L-1}\|P_n\|$ converges.

We first assume, for simplicity, that M consists of just one Jordan block, so $M = a(I + J)$ for some $a \neq 0$. Although M itself cannot be diagonalized, it is sufficient to find a matrix sequence $\mathbf{G}$ such that $\Lambda_n := G_{n+1}^{-1}MG_n$ are diagonal. Since M^n can be calculated explicitly, G_n can indeed be found: $G_n = M^nG_0L_n^{-1}$ where $L_n = \prod_{h=0}^{n-1}\Lambda_h$ is a non-singular diagonal matrix and G_0 is an arbitrary non-singular matrix. Set $L_n = a^n \cdot \operatorname{diag}(\ell_1(n), \ldots, \ell_k(n))$. Since the transformed perturbation matrices are $G_{n+1}^{-1}P_nG_n$, we determine G_n by demanding that $\|G_n\| \cdot \|G_n^{-1}\|$ is as small as possible. If we choose $G_0 = I$, then

$$\|G_n\| \sim \max(1/|\ell_1(n)|, n/|\ell_2(n)|, \ldots, n^{k-1}/(k-1)!|\ell_k(n)|)$$

and

$$\|G_n^{-1}\| \sim \max(n^{k-1}|\ell_1(n)|/(k-1)!, \ldots, n|\ell_{k-1}(n)|, |\ell_k(n)|).$$

Neglecting constant factors,

$$\|G_n\| \cdot \|G_n^{-1}\| \sim n^{k-1} \cdot \frac{n^i/|\ell_{i+1}(n)|}{n^j/|\ell_{j+1}(n)|}$$

if $n^i/|\ell_{i+1}(n)|$ is the largest of the numbers $n^p/|\ell_{p+1}(n)|$ and $n^j/|\ell_{j+1}(n)|$ the smallest. Hence, $\|G_n\| \cdot \|G_n^{-1}\|$ is as small as possible if $|\ell_{p+1}(n)|/n^p$ are all approximately equal. Putting $\ell_p(n) = n^{p-1} \cdot \ell(n)$ for some positive numbers $\ell(n)$, we find that $\|G_n\| \cdot \|G_n^{-1}\|$ is of order n^{k-1}. It can be shown that other choices for G_0 do not give better estimates. The value of $\ell(n)$ is not relevant (as long as $\ell(n) \neq 0$) and we choose $\ell(n) = 1$. Thus

$$G_n = (I + J)^n \cdot n^{-B_k}$$

where $B_k = \mathrm{diag}(0, 1, \ldots, k-1)$, and

$$G_{n+1}^{-1} M G_n = a \left(\frac{n+1}{n} \right)^{B_k}$$

and

$$G_{n+1}^{-1} M_n G_n = a \left(\frac{n+1}{n} \right)^{B_k} + U_n$$

where $\sum_{n=0}^{\infty} \|U_n\| = \mathcal{O}(\sum_{n=0}^{\infty} n^{k-1} \|P_n\|)$.
In case M consists of several Jordan blocks $a_i(I + J)$ of order k_i by k_i, we let G_n be the block diagonal matrix whose blocks are $(I + J)^n \cdot n^{-B_{k_i}}$ (where I, J are $k_i \times k_i$-matrices). If L is the maximum of the numbers k_i, then

$$G_{n+1}^{-1} M G_n = a \left(\frac{n+1}{n} \right)^{B}$$

and

$$G_{n+1}^{-1} M_n G_n = a \left(\frac{n+1}{n} \right)^{B} + U_n'$$

where $\sum_{n=0}^{\infty} \|U_n'\| = \mathcal{O}(\sum_{n=0}^{\infty} n^{L-1} \|P_n\|)$ and where $B = \mathrm{diag}(B_{k_1}, \ldots, B_{k_\ell})$ is diagonal. Hence, if $\sum_{n=0}^{\infty} n^{L-1} \|P_n\|$ converges, a matrix sequence $\mathbf{F}$ exists such that $\lim_{n\to\infty} F_n = I$ and

$$F_{n+1}^{-1} G_{n+1}^{-1} M_n G_n F_n = a \left(\frac{n+1}{n} \right)^{B}.$$

Remark 4.1: The above result is not new: a similar result was already obtained by Coffman [10] for linear recurrences, and the matrix version occurs already in [2], if not also elsewhere. The example was given just to show the idea of rescaling, which we shall explore further in the next section.

5 DIAGONALIZATION BY RESCALING SEQUENCES

The matrices G_n in the previous example diagonalize the (constant) matrix sequence $\{M\}$ by separating the eigenvalues of M. In addition, the eigenvectors of the matrices of the transformed sequence do not lie close together in $\mathbb{C}^n$. Even if the eigenvalues of the matrices in a matrix sequence are distinct, the eigenvectors may lie too close together, in which case the matrix sequence cannot be diagonalized by using matrices of normalized eigenvectors (as was done in section 3) because their inverses become too large. Instead, like in section 4, we can try to find matrices G_n such that the transformed sequence $\{G_{n+1}^{-1} M_n G_n\}$ is such that Theorem 2.1 or Theorem 3.2/Corollary 3.4 can be applied. We call such a sequence $\{G_n\}$ a *rescaling sequence*. In this section we will show how to find rescaling sequences for a large class of 2×2-matrices.

We study sequences $\{M_n\}$ of 2×2-matrices of the form

$$M_n = b_0(n)A_0 + \ldots + b_m(n)A_m + P_n =: M_n^{(0)} + P_n \tag{5.1}$$

where $A_0, \ldots, A_m$ are linearly independent fixed matrices with real entries (and where $0 \le m \le 3$), $\{b_i(n)\}$ are sequences of positive real numbers of bounded variation such that $\lim_{n\to\infty} b_0(n) = 1$ and

$$\lim_{n\to\infty} \frac{b_{i+1}(n)}{b_i(n)} = 0$$

and $\{P_n\}$ is a sequence of perturbation matrices of bounded variation where $P_n = o(b_m(n))$. Additional conditions on the coefficients $b_i(n)$ and on P_n are added when necessary. We will try to find explicit matrix sequences $\{H_n\}$ such that $H_{n+1}^{-1} M_n H_n$ is almost-diagonal and such that Theorem 2.1 can be applied. In general, H_n will be of the form $H_n = G_n E_n$ where $\{G_n\}$ is a rescaling sequence and $\{E_n\}$ a sequence of normalized eigenvectors of $G_{n+1}^{-1} M_n G_n$. In some cases, no rescaling will be needed and we can take $G_n = I$.

Definition: In (5.1), we call the matrix A_0 the *leading matrix* if $A_0 \neq aI$ for $a \in \mathbb{C}$. If A_0 is a multiple of the identity matrix, then we call A_1 the leading matrix. The term $b_0 A_0$ - or $b_1 A_1$ - containing the leading matrix will be called the *leading term*.

If $\{p(n)\}$ and $\{q(n)\}$ are sequences of real or complex numbers of bounded variation, the sequences $\{p(n) \pm q(n)\}$ and $\{p(n)q(n)\}$ are also of bounded variation, and if $\lim_{n\to\infty} q(n) = q \neq 0$, the same holds for the sequence $\{p(n)/q(n)\}$. Hence we may assume (by replacing $b_i(n)/b_0(n)$ by $b_i(n)$) that $b_0(n) = 1$ for all n.

We distinguish the following cases:

1. A_0 is diagonalizable with distinct eigenvalues. (type I)
2. $A_0 = aI$ for some $a \neq 0$ and A_1 is diagonalizable. (type II)
3. $A_0 = aI$ for $a \neq 0$ and A_1 is non-diagonalizable. (type IIIa)
4. A_0 is non-diagonalizable and has a non-zero eigenvalue. (type IIIb)

5. A_0 is non-diagonalizable and has eigenvalue zero. (type IV)

We treat the five cases one by one.

1. **Type I.** Corollary 3.4 may be applied if two conditions are met: (i.) The sequences $\{P_n\}$ and $\{b_i(n)\}$ are of bounded variation. (ii.) Condition (1.1) on the eigenvalues $a_1(n), a_2(n)$ of $M_n^{(0)}$ must hold. Since $M_n^{(0)}$ are real matrices, $a_1(n)$, $a_2(n)$ are either both real numbers or $a_2(n)$ is the complex conjugate of $a_1(n)$. In the latter case, $|a_1(n)| = |a_2(n)|$ and (1.1) is satisfied. (1.1) is also satisfied if the eigenvalues a_1, a_2 of A_0 have distinct moduli, so the only case we have to look at more closely, is that $a_1 = -a_2 \neq 0$. In this case, condition (1.1) is certainly satisfied if $|a_1(n)| - |a_2(n)|$ has constant sign (or is zero). But $a_1(n) + a_2(n) = \pm(|a_1(n)| - |a_2(n)|)$ equals the trace of $M_n^{(0)}$ so, at least for n large enough, its sign is indeed constant (it is equal to the sign of the trace of A_i if $\mathrm{tr}(A_j) = 0$ for $j < i$). Hence we can apply Corollary 3.4 and diagonalize the sequence $\{M_n\}$.

2. **Type II.** A non-singular matrix C exists such that $C^{-1}A_1C = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}$ is diagonal and $c_1 \neq c_2$. If we add the requirement that the sequences $\{b_2(n)/b_1(n)\}$, $\{b_3(n)/b_1(n)\}$ and $\{P_n/b_1(n)\}$ are of bounded variation, then
$$C^{-1}M_nC = aI + b_1(n)C^{-1}A_1C + Q_n$$
where $Q_n = o(b_1(n))$ and $\{Q_n/b_1(n)\}$ is of bounded variation. We can now apply Corollary 3.4 to the sequence $\{C^{-1}M_nC\}$ if condition (1.1) on the eigenvalues $a_1(n), a_2(n)$ of $M_n^{(0)}$ is satisfied. If the eigenvalues are not real, one is the complex conjugate of the other and $|a_1(n)| = |a_2(n)|$. If both eigenvalues are real, then condition (1.1) is certainly satisfied if $|a_1(n)| - |a_2(n)|$, and hence, $a_1(n) - a_2(n)$ has constant sign. But by Lemma 3.3, the eigenvalues are $a_1(n) = a + b_1(n)c_1 + o(b_1(n))$ and $a_2(n) = a + b_1(n)c_2 + o(b_1(n))$, so that $a_1(n) - a_2(n) = b_1(n)(c_1 - c_2) + o(b_1(n))$. Hence the conditions of Corollary 3.4 are met and $\{M_n\}$ can be diagonalized.

3. **Type IIIa.** For some non-singular matrix C, $C^{-1}A_1C = \alpha J$, where $J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\alpha \neq 0$. Writing M_n for $C^{-1}M_nC/a$, we have
$$M_n = I + \alpha f(n)J + g(n)\begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} + \beta h(n)J^T + P_n$$
where $f(n) = b_1(n) = o(1)$ and $g(n), h(n)$ are either of the $b_2(n), b_3(n)$ if these occur. We treat this case below.

4. **Type IIIb.** For some non-singular matrix C, $C^{-1}A_0C = a(I + \alpha J)$, where $\alpha \neq 0$. Writing M_n for $C^{-1}M_nC/a$, we have, similar to case 3,
$$M_n = I + \alpha f(n)J + g(n)\begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} + \beta h(n)J^T + P_n$$

where now $\lim_{n\to\infty} f(n) = 1$. We treat types IIIa and IIIb simultaneously.

5. **Type IV.** For some non-singular matrix C, $C^{-1}A_0C = J$. Replacing $C^{-1}M_nC$ by M_n, we may write

$$M_n = J + \beta h(n) J^T + g(n)\begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix} + k(n)\begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} + P_n$$

and where $k(n) = o(g(n))$.

For types III and IV, Corollary 3.4 cannot be applied directly. Instead, we will try and find some rescaling sequence $\{G_n\}$ in such a way that Corollary 3.4 can be applied to the sequence $\{G_{n+1}^{-1}M_nG_n\}$. Since the leading terms $(I+)\alpha f(n)J$ and J have $\mathbf{e}_1$ as the only eigenvector, we may expect that (if the remaining terms are small) the columns of the product matrices $M_{n-1}\ldots M_1M_0$ are close to the direction of $\mathbf{e}_1$, as in the example of the previous section. If we now let G_n be a diagonal matrix $G_n = \begin{pmatrix} g_n^{-1} & 0 \\ 0 & 1 \end{pmatrix}$, where $\lim_{n\to\infty} g_n = 0$, we shorten the $\mathbf{e}_1$-components of the columns with respect to the $\mathbf{e}_2$-components in the orthogonal direction. g_n can be determined by requiring that the eigenvectors of the matrices $G_{n+1}^{-1}M_nG_n$ do not lie close together. We use the following lemma, which we state without proof:

Lemma 5.1 *The eigenvalues of the matrix* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, *where A is not a multiple of the identity matrix, are* $\big((a+d) \pm \sqrt{D}\big)/2$, *where* $D = (a-d)^2 + 4bc$ *and the corresponding eigenvectors are* $\begin{pmatrix} 2b \\ d - a \pm \sqrt{D} \end{pmatrix}$ *or, equivalently, by* $\begin{pmatrix} d - a \mp \sqrt{D} \\ -2c \end{pmatrix}$. *If $D = 0$, A is not diagonalizable.*

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $A \neq aI$, as in the lemma. Suppose $|c| \le |b|$ and $b \neq 0$. We write the eigenvectors as $\begin{pmatrix} b \\ a_1 \end{pmatrix}$ and $\begin{pmatrix} b \\ a_2 \end{pmatrix}$, where $a_1a_2 = -bc$. So at least one of the two numbers $|a_1/b|, |a_2/b|$ is not greater than 1, $|a_1/b| \le 1$ say. Then the directions of the eigenvectors are not too close if $|\sqrt{D}/b| = |(a_1 - a_2)/b|$ is not too small. Similarly, if $|c| \le |b|$, then $|\sqrt{D}/c|$ must not be too small. This leads to the *rescaling conditions* for the rescaling sequence $\{G_n\} = \begin{pmatrix} g_n^{-1} & 0 \\ 0 & 1 \end{pmatrix}$: let $M_n = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix}$. The sequence $\{G_{n+1}M_nG_n = \begin{pmatrix} a_ng_{n+1}/g_n & b_ng_{n+1} \\ c_n/g_n & d_n \end{pmatrix}\}$ is properly rescaled if for some numbers $R > 0$ and $\delta > 0$

- either $|b_ng_ng_{n+1}/c_n| \le R$ and

$$\left| \left(\frac{a_ng_{n+1} - d_ng_n}{c_n} \right)^2 + 4\frac{b_n}{c_n} \cdot g_ng_{n+1} \right| > \delta$$

  for n sufficiently large,

- or $|c_n g_n^{-1} g_{n+1}^{-1}/b_n| \le R$ and

$$\left|\left(\frac{a_n g_n^{-1} - d_n g_{n+1}^{-1}}{b_n}\right)^2 + 4\frac{c_n}{b_n} \cdot g_n^{-1} g_{n+1}^{-1}\right| > \delta$$

  for n sufficiently large.

We now proceed to the treatment of types III and IV.
(i). Type III: $M_n = M_n^{(0)} + P_n$,

$$M_n^{(0)} = I + \alpha f(n) \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + g(n) \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} + \beta h(n) \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{5.2}$$

where $g(n), h(n) = o(f(n))$ and $\{f(n)\}, \{g(n)\}, \{h(n)\}$ are bounded variation sequences and $\lim_{n\to\infty} f(n) = 0$ or 1. We may assume that $\sum_{n=0}^{\infty} f(n) = \infty$, otherwise $\{M_n\}$ can be directly diagonalized with the aid of Theorem 2.1. We define $F(n) = \sum_{\ell=0}^{n} f(\ell)$. Then $\lim_{n\to\infty} f(n)/F(n) = 0$ and $\lim_{n\to\infty} F(n+1)/F(n) = 1$.
After rescaling by $G_n = \mathrm{diag}(g_n^{-1}, 1)$, we obtain

$$G_{n+1}^{-1} M_n G_n = \begin{pmatrix} (1 + c_1 g(n)) g_{n+1}/g_n & \alpha f(n) g_{n+1} \\ \beta h(n)/g_n & 1 + c_2 g(n) \end{pmatrix} + P_n'$$

where $P_n' = G_{n+1}^{-1} P_n G_n = \mathcal{O}(\|P_n\|) \cdot g_n^{-1}$. Larger values of g_n thus give less severe conditions on the perturbation sequence $\mathbf{P}$.
If $h(n)$ is large and $\beta \neq 0$, choosing $g_n = \sqrt{h(n)/f(n)}$ ensures that the term $(c_n/b_n) g_n^{-1} g_{n+1}^{-1}$ is of order unity. On the other hand, if $h(n)$ is sufficiently small, then we must have that

$$\left|g_n^{-1}(1 + c_1 g(n)) - g_{n+1}^{-1}(1 + c_2 g(n))\right| > \delta f(n).$$

If we solve the approximate equality (replacing $> \delta f(n)$ by $\approx f(n)$), we find that

- If $g(n)/f(n) = o(1/F(n))$ then $g_n = 1/F(n-1)$ is a solution.
- If $1/F(n) = o(g(n)/f(n))$, then $g_n = g(n)/f(n)$ is a solution.
- If $\lim_{n\to\infty} g(n)F(n)/f(n) = q \neq 0$, then either $g_n = 1/F(n-1)$ or $g_n = g(n)/f(n)$ is a solution, except if $q(c_1 - c_2) = 1$. In that case, $g_n^{-1} = F(n-1) \cdot \sum_{\ell=0}^{n-1} \frac{f(\ell)}{F(\ell)}$ is a solution.

We collect the results in the form of a proposition:

Proposition 5.2 *Let $M_n^{(0)}$ be a type III sequence as in (5.2) where the sequences $\{f(n)\}$, $\{g(n)\}$ and $\{h(n)\}$ are of bounded variation, $f(n) = o(1)$ or $\lim_{n\to\infty} f(n) = 1$ and $g(n)/f(n)$,*

$h(n)/f(n)$ converge monotonously (at least for n large enough) to zero. Either of the sequences $\{g(n)\}$ or $\{h(n)\}$ may be the zero sequence (in particular, we let $h(n) = 0$ if $\beta = 0$). Furthermore, we assume that

$$\lim_{n\to\infty} \frac{f(n+1)}{f(n)} = \lim_{n\to\infty} \frac{g(n+1)}{g(n)} = \lim_{n\to\infty} \frac{h(n+1)}{h(n)} = 1.$$

(If either of the sequences $\{g(n)\}, \{h(n)\}$ is the zero sequence, the corresponding condition becomes empty.) Define the three sequences $\{p_i(n)\}$ $(i = 1, 2, 3)$ as follows:

$$p_1(n) = 1/F(n-1),\ p_2(n) = g(n)/f(n),\ p_3(n) = \sqrt{h(n)/f(n)}$$

and let $q_i(n) = (1/p_i(n+1) - 1/p_i(n))$ for $i = 1, 2, 3$ (in particular, $q_1(n) = f(n+1)$). Suppose that for all i, j the limits $p_{ij} = \lim_{n\to\infty} q_j(n)/q_i(n)$ exist (p_{ij} is zero or finite) or $p_{ij} = \infty$. Then $p_{ij} = \lim_{n\to\infty} p_i(n)/p_j(n)$. Let g_n be the maximum of the numbers $p_1(n), p_2(n), p_3(n)$. If we choose $G_n = \begin{pmatrix} g_n^{-1} & 0 \\ 0 & 1 \end{pmatrix}$ then the rescaled sequence $\{G_{n+1}^{-1} M_n G_n\}$ is of type II except if $g_n = p_i(n)$ (so that $p_{ji} < \infty$ for all i) and at least one of the numbers p_{ji} with $j \neq i$ is not zero. In the latter case, the rescaled sequence is also of type II except for an exceptional case where the numbers $\alpha, \beta, c_1, c_2, p_{ij}$ satisfy a certain algebraic equation; in this case the rescaled matrix sequence is of type III. In either of the non-exceptional cases, the sequence $\{G_{n+1}^{-1} M_n G_n\}$ satisfies the conditions of Corollary 3.4 and can be diagonalized, provided that the sequences $\{\frac{1}{f(n)}\left(\frac{1}{g_{n+1}} - \frac{1}{g_n}\right)\}$, $\{g(n)/(g_n f(n))\}$ and $\{h(n)/(g_n^2 f(n))\}$ are of bounded variation, as well as $\{G_{n+1}^{-1} P_n G_n / g_n\}$.In the exceptional case, another rescaling is needed.

Remark 5.1: In the proof, the monotony of the sequences $\{g(n)/f(n)\}$ and $\{h(n)/f(n)\}$ is only used to conclude from $p_{ij} = \lim_{n\to\infty} q_j(n)/q_i(n)$ that $p_{ij} = \lim_{n\to\infty} p_i(n)/p_j(n)$. Instead of demanding monotony, it is also possible to impose the somewhat weaker condition that $\{g(n)/f(n)\}$ and $\{h(n)/f(n)\}$ are of bounded variation, but in that case $p_{ij} = \lim_{n\to\infty} p_i(n)/p_j(n)$ must be added as a separate condition.

Remark 5.2: We may take any sequence $\{g_n c_n\}$ instead of $\{g_n\}$ in Proposition 5.2, where $\{c_n\}$ is some bounded variation sequence of positive numbers converging to some number $c > 0$. Indeed, this amounts to replacing a matrix sequence $\{M_n\}$ of type I or II by the sequence $\{C_{n+1}^{-1} M_n C_n\}$ where $\{C_n = \operatorname{diag}(1/c_n, 1)\}$ is a bounded variation sequence converging to some invertible matrix C. The latter sequence is again of type I or II, respectively. Hence, if $g_n = p_i(n)$ for some i, and p_{ij} is finite and not zero, so that $p_i(n)$ and $p_j(n)$ have the same order of magnitude, then we may as well take $g_n = p_j(n)$.

Proof:
(i.) First assume that p_{21} and p_{31} are finite so that we can take $g_n = p_1(n) = 1/F(n-1)$ and $G_n = \operatorname{diag}(F(n-1)^{-1}, 1)$. Then

$$G_{n+1}^{-1} M_n G_n = I + \frac{f(n)}{F(n)} \begin{pmatrix} -1 & \alpha \\ 0 & 0 \end{pmatrix} + g(n) \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} + h(n)F(n-1) \begin{pmatrix} 0 & 0 \\ \beta & 0 \end{pmatrix} + P_n'$$

where $P'_n = G_{n+1}^{-1} P_n G_n$ plus a term of order $g(n)f(n)/F(n)$, which is obviously of lower order than $f(n)/F(n)$. Now if $p_{31} = p_{32} = 0$, the rescaled sequence is of type II since $g(n)$ and $h(n)F(n-1)$ are of lower order than $f(n)/F(n)$. If one or both of the numbers p_{31}, p_{32} are non-zero, the rescaled sequence is of type II unless $(p_{21}(c_1 - c_2) - 1)^2 + 4\alpha\beta p_{31}^2 = 0$ in which case the matrix $\begin{pmatrix} -1 + p_{21}c_1 & \alpha \\ \beta p_{31}^2 & p_{21}c_2 \end{pmatrix}$ is not diagonalizable. In that case, the sequence is of type III and we need to rescale once more. Later we give an example of this situation.

(ii.) Now assume that $p_{12} = 0$ and p_{32} is finite so that we can set $g_n = g(n)/f(n) = p_2(n)$. Then the rescaled sequence has terms

$$G_{n+1}^{-1} M_n G_n = I + g(n) \begin{pmatrix} c_1 & \alpha \\ 0 & c_2 \end{pmatrix} + \left(\frac{p_2(n+1)}{p_2(n)} - 1 \right) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + h(n)f(n)/g(n) \begin{pmatrix} 0 & 0 \\ \beta & 0 \end{pmatrix} + P'_n$$

where $P'_n = G_{n+1}^{-1} P_n G_n$ plus a term of order $g(n)(p_2(n+1)/p_2(n) - 1)$ which is of smaller order than $g(n)$. We show that the term with $g(n)$ is the leading term. Indeed,

$$\left| \frac{p_2(n+1)/p_2(n) - 1}{g(n+1)} \right| = \left(\frac{1}{p_2(n+1)} - \frac{1}{p_2(n)} \right) / f(n+1) = \frac{q_2(n)}{q_1(n)}$$

which converges to zero since $p_{12} = 0$. So we have a type II sequence except when $p_{32} \neq 0$ and $(c_1 - c_2)^2 + 4\alpha\beta p_{32}^2 = 0$ when the matrix $\begin{pmatrix} c_1 & \alpha \\ \beta p_{32}^2 & c_2 \end{pmatrix}$ is not diagonalizable. In that case, the sequence is of type III and we need to rescale once more.

(iii.) Finally, assume that $p_{13} = p_{23} = 0$ so that $g_n = \sqrt{h(n)/f(n)} = p_3(n)$. The rescaled matrices are

$$G_{n+1}^{-1} M_n G_n = I + p_3(n)f(n) \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix} + g(n) \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} + \left(\frac{p_3(n+1)}{p_3(n)} - 1 \right) \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} + P'_n$$

where $P_n = G_{n+1}^{-1} P_n G_n$ plus terms of order $g(n)(\frac{p_3(n+1)}{p_3(n)} - 1)$ and $f(n)p_3(n)(\frac{p_3(n+1)}{p_3(n)} - 1)$ which are of smaller order than the leading term. The factor $p_3(n+1)/p_3(n) - 1$ is also of smaller order than $p_3(n)f(n) = \sqrt{h(n)f(n)}$. This follows since

$$\lim_{n\to\infty} \frac{1}{f(n)} \left(\frac{1}{p_3(n+1)} - \frac{1}{p_3(n)} \right) = \lim_{n\to\infty} \frac{q_3(n)}{q_1(n)} = 0.$$

Hence, the rescaled sequence is of type II.

In all of these three cases where the rescaled matrix sequences are of type II (so excepting the exceptional cases), the assumptions on the bounded variation of the sequences $\{\frac{1}{f(n)}\left(\frac{1}{g_{n+1}} - \frac{1}{g_n}\right)\}$, $\{g(n)/(g_n f(n))\}$, $\{h(n)/(g_n^2 f(n))\}$ and $\{G_{n+1}^{-1} P_n G_n / g_n\}$ allow Corollary 3.4 to be applied to the rescaled sequence $\{G_{n+1}^{-1} M_n G_n\}$. Notice that product sequences of bounded variation are of bounded variation; in particular, the sequence $\{g_{n+1}/g_n\}$ is of bounded variation. This concludes the argument. □

We now give a few examples.

Example 5.1.

$$M_n = \begin{pmatrix} 1+\frac{3}{\sqrt{n}}+\frac{1}{n} & \frac{1}{\sqrt{n}}+\frac{2}{n} \\ -\frac{1}{\sqrt{n}}-\frac{1}{n} & 1+\frac{1}{\sqrt{n}} \end{pmatrix} + P_n := M_n^{(0)} + P_n$$

where $\sum_{n=1}^{\infty} n^{1/4}\|P_n\|$ converges.
We can write this as

$$M_n = I + \frac{1}{\sqrt{n}}\begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} + \frac{1}{n}\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} + P_n.$$

We can take $A_0 = I$ and $A_1 = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$. A_1 is non-diagonalizable: we have a sequence of type IIIa. First write A_1 in Jordan normal form: if $C = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, then $C^{-1}A_1C = 2I + J$ and

$$C^{-1}M_nC = I + \frac{1}{\sqrt{n}}\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} + \frac{1}{n}\begin{pmatrix} -1 & 2 \\ -2 & 2 \end{pmatrix} + C^{-1}P_nC.$$

We can rewrite this in the form

$$C^{-1}M_nC = a(n)I + \tilde{f}(n)\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \tilde{h}(n)\begin{pmatrix} 0 & 0 \\ -2 & 0 \end{pmatrix} + \tilde{g}(n)\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} + C^{-1}P_nC$$

where

$$a(n) = 1 + \frac{2}{\sqrt{n}} + \frac{1}{2n},\ \tilde{f}(n) = \frac{1}{\sqrt{n}} + \frac{2}{n},$$

$$\tilde{g}(n) = \frac{3}{2n},\ \tilde{h}(n) = \frac{1}{n}.$$

Factoring out $a(n)$, this gives

$$C^{-1}M_nC = a(n)\left(I + f(n)\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + h(n)\begin{pmatrix} 0 & 0 \\ -2 & 0 \end{pmatrix} + g(n)\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} + P'_n\right)$$

where

$$f(n) = \frac{1}{\sqrt{n}} + \mathcal{O}(\frac{1}{n\sqrt{n}}),$$

$$g(n) = \frac{3}{2n} + \mathcal{O}(\frac{1}{n\sqrt{n}}),\ h(n) = \frac{1}{n} + \mathcal{O}(\frac{1}{n\sqrt{n}})$$

and $P'_n = C^{-1}P_nC/a(n)$. Now $F(n) = 2\sqrt{n} + \mathcal{O}(1)$.
By Proposition 5.2 and Remark 5.2 we can take $g_n = n^{-1/4}$ and $G_n = \begin{pmatrix} n^{1/4} & 0 \\ 0 & 1 \end{pmatrix}$. Then

$$G_{n+1}^{-1}C^{-1}M_nCG_n = a(n)(I + n^{-3/4}\begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix} + n^{-1}\begin{pmatrix} -3/2 & 0 \\ 0 & -3/2 \end{pmatrix} + D_n + \mathcal{O}(n^{1/4}\|P_n\|))$$

where the coefficients of D_n are power series in $n^{-1/4}$ and $D_n = \mathcal{O}(1/n^{5/4})$. So, since $\begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix}$ is diagonalizable, we have a type II sequence so there exists a sequence of matrices $\{H_n\}$ such that H_n converges to $H = \begin{pmatrix} 1 & 1 \\ -i\sqrt{2} & i\sqrt{2} \end{pmatrix}$ - a matrix of eigenvectors of the leading matrix $\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$, and

$$H_{n+1}^{-1}G_{n+1}^{-1}C^{-1}M_nCG_nH_n = a(n)\cdot\ \mathrm{diag}(\epsilon_1(n), \epsilon_2(n))$$

where $\epsilon_1(n) = 1 - \frac{i\sqrt{2}}{n^{3/4}} + \ldots$ and $\epsilon_2(n) = 1 + \frac{i\sqrt{2}}{n^{3/4}} + \ldots$ are the eigenvalues of $N_n = I + n^{-3/4}\begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix} + n^{-1}\begin{pmatrix} -3/2 & 0 \\ 0 & -3/2 \end{pmatrix} + D_n$ and by applying (3.5) to $n^{3/4}(N_n - I)$ and subsequently, lemma 2.1, to $N_n + \mathcal{O}(n^{1/4}P_n)$, we obtain the estimate

$$\|H_n - H\| \le C' \cdot \max(n^{-1/4}, \sum_{\ell=n}^{\infty} \ell^{1/4}\|P_\ell\|)$$

for some constant C'.

Example 5.2. Let $M_n = \begin{pmatrix} 1 & 1 \\ \frac{c}{n(n+1)} & 1 \end{pmatrix} + P_n$, where $P_n = \begin{pmatrix} p_{11}(n) & p_{12}(n) \\ p_{21}(n) & p_{22}(n) \end{pmatrix}$ such that $\sum_{n=0}^{\infty} |p_{11}(n)|$, $\sum_{n=0}^{\infty} |p_{22}(n)|$, $\sum_{n=0}^{\infty} |p_{12}(n)|/n$ and $\sum_{n=0}^{\infty} n|p_{21}(n)|$ converge.
$\{M_n\}$ is a type IIIb sequence of the form as in Proposition 5.2 where $f(n) = 1$ and $h(n) = c/(n^2+n)$. We have $F(n) = n$ and $\lim_{n\to\infty} h(n)f(n)/F(n)^2 = c$. Letting $g_n = 1/F(n-1) = 1/n$ and $G_n = \begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix}$, we obtain

$$G_{n+1}^{-1}M_nG_n = I + \frac{1}{n+1}\begin{pmatrix} -1 & 1 \\ c & 0 \end{pmatrix} + G_{n+1}^{-1}P_nG_n = I + \frac{1}{n+1}B + G_{n+1}^{-1}P_nG_n.$$

By the assumption on the coefficients of P_n, $\sum_{n=0}^{\infty} \|G_{n+1}^{-1}P_nG_n\|$ converges. If $c \ne -1/4$, the matrix B is diagonalizable, and the rescaled sequence is of type II. If C is a non-singular matrix of eigenvectors of B, then $C^{-1}(I + \frac{1}{n+1}B)C = I + \frac{1}{n+1}\ \mathrm{diag}(b_1, b_2)$, where b_1, b_2 are the distinct eigenvalues of B. Since $\sum_{n=0}^{\infty} \|G_{n+1}^{-1}P_nG_n\|$ converges, there exists, by Theorem 2.1, a matrix sequence $\{F_n\}$ with $\lim_{n\to\infty} F_n = I$, such that

$$F_{n+1}^{-1}C^{-1}G_{n+1}^{-1}M_nG_nCF_n = \ \mathrm{diag}(1 + b_1/(n+1), 1 + b_2/(n+1)).$$

Moreover, we have the following estimate for $F_n - I$:

$$\|F_n - I\| \le C \cdot \max_{i,j=1,2}\left(\sum_{\ell=n}^{\infty} |p_{ij}(\ell)|\ell^{i-j}, n^A \sum_{(n)} |p_{ij}(\ell)|\ell^{i-j-A}, n^{-A} \sum_{(n)} |p_{ij}(\ell)|\ell^{i-j+A}\right)$$

for some constant C and where $A = \text{Re}(b_1 - b_2)$.
In order to see how this works out for a more explicit perturbation matrix P_n, let us suppose that $p_{11}(n)$, $p_{22}(n)$, $p_{12}(n)/n$ and $np_{21}(n)$ are all of order $\mathcal{O}(n^{-3/2})$. Then, by (3.5),

$$\|F_n - I\| \leq C \cdot \max\left(\sum_{\ell=n}^{\infty} \ell^{-3/2}, n^A \sum_{(n)} \ell^{-3/2-A}, n^{-A} \sum_{(n)} \ell^{-3/2+A}\right).$$

Almost always, these terms are of the same order $\mathcal{O}(1/\sqrt{n})$. However, if $A = \pm 1/2$, then one of the three terms is of order $\mathcal{O}(\ln n/\sqrt{n})$.
On the other hand, if $c = -1/4$, B is not diagonalizable. We have type IIIa. Letting $C = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$, we have

$$C^{-1}G_{n+1}^{-1}M_nG_nC = (1 - \frac{1}{2(n+1)})I + \frac{1}{4(n+1)}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + C^{-1}G_{n+1}^{-1}P_nG_nC.$$

We can rescale again with $G'_n = \begin{pmatrix} \ln n & 0 \\ 0 & 1 \end{pmatrix}$ (since $\sum_{\ell=1}^{n} 1/\ell = \ln n + \mathcal{O}(1)$) and find that

$$(G'_{n+1})^{-1}C^{-1}G_{n+1}^{-1}M_nG_nCG'_n = (1-\frac{1}{2(n+1)})\big(I+\frac{1}{n\ln n}\begin{pmatrix} -1 & 1/4 \\ 0 & 0 \end{pmatrix}+\mathcal{O}(\frac{1}{n^2})+\mathcal{O}(\ln n G_{n+1}^{-1}P_nG_n)\big).$$

So if we now assume that a slightly stronger condition on the coefficients of P_n holds, namely that $\sum_{n=0}^{\infty} |p_{11}(n)| \ln n$, $\sum_{n=0}^{\infty} |p_{22}(n)| \ln n$, $\sum_{n=0}^{\infty} |p_{12}(n)| \ln n/n$ and $\sum_{n=0}^{\infty} n \ln n |p_{21}(n)|$ converge, then by Corollary 2.2 a sequence $\{F_n\}$ can be found, converging to I where

$$F_{n+1}^{-1}\Gamma_{n+1}^{-1}M_n\Gamma_nF_n = (1 - \frac{1}{2(n+1)})\left(I + \frac{1}{n\ln n}\begin{pmatrix} -1 & 1/4 \\ 0 & 0 \end{pmatrix}\right),$$

and where $\Gamma_n = G_nCG'_n$ are now the rescaling matrices. In this case there are no *diagonal* rescaling matrices. We have found a rescaling sequence by applying the rescaling procedure twice.

(ii). **Type IV.** $M_n = M_n^{(0)} + P_n$ with

$$M_n^{(0)} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \beta h(n)\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + g(n)\begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix} + k(n)\begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} \tag{5.3}$$

where $k(n) = o(g(n))$ and either $b_1c_2 - b_2c_1 \neq 0$ or $k(n) = 0$.
As for types IIIa and IIIb, we summarize the results in the form of a proposition.

Proposition 5.3 *Let $\{M_n^{(0)}\}$ be a type IV sequence as in (5.3) where the sequences $\{g(n)\}$, $\{\sqrt{h(n)}\}$ and $\{k(n)/g(n)\}$ are bounded variation sequences converging to zero. The sequences $\{g(n)\}$, $\{k(n)\}$ and $\{h(n)\}$ may be zero sequences, in particular is $h(n) = 0$ if $\beta = 0$. Furthermore, we assume that*

$$\lim_{n\to\infty} \frac{g(n+1)}{g(n)} = \lim_{n\to\infty} \frac{h(n+1)}{h(n)} = \lim_{n\to\infty} \frac{k(n+1)}{k(n)} = 1.$$

(If $g(n) \equiv 0$, $k(n) \equiv 0$ or $h(n) \equiv 0$ the corresponding conditions are empty.) We distinguish between two cases.

(i) $b_1 \neq b_2$. Suppose that $\lim_{n\to\infty} g(n)/\sqrt{h(n)} = q$ exists or is infinity. Let g_n be the maximum of the numbers $\sqrt{h(n)}$ and $g(n)$ and define the rescaling matrices by $G_n = \begin{pmatrix} g_n^{-1} & 0 \\ 0 & 1 \end{pmatrix}$. Then the rescaled sequence $\{G_{n+1}^{-1} M_n G_n\}$ is of type I if $q = 0$ or $q = \infty$. If $q \neq 0$ then the rescaled sequence is also of type I except if $q^2(b_1 - b_2)^2 + 4\beta = 0$, in which case the rescaled sequence is of type III. If moreover the sequences $\{g_{n+1}/g_n\}$, $\{g(n)/g_n\}$, $\{k(n)/g_n\}$, $\{h(n)/g_n^2\}$ and $\{G_{n+1}^{-1} P_n G_n/g_n\}$ are of bounded variation, Corollary 3.4 can be applied and the rescaled sequence is diagonalizable.

(ii) $b_1 = b_2 = b$. Let $G(n) = \sum_{\ell=0}^{n} g(\ell)^{-1}$. Define the sequences $\{p_i(n)\}$ ($i = 1, 2, 3$) by

$$p_1(n) = 1/G(n-1),\ p_2(n) = k(n),\ p_3(n) = \sqrt{h(n)}$$

and let $q_i(n) = (1/p_i(n+1) - 1/p_i(n))$ for $i = 1, 2, 3$. Suppose that for all i, j the limits $p_{ij} = \lim_{n\to\infty} q_j(n)/q_i(n)$ exist (p_{ij} is zero or finite) or $p_{ij} = \infty$. Then $p_{ij} = \lim_{n\to\infty} p_i(n)/p_j(n)$. Let g_n be the maximum of the numbers $p_1(n), p_2(n), p_3(n)$ and define the rescaling matrices by $G_n = \begin{pmatrix} g_n^{-1} & 0 \\ 0 & 1 \end{pmatrix}$. Then the rescaled sequence $\{G_{n+1}^{-1} M_n G_n\}$ is of type I or II if $g_n = p_j(n)$ and $p_{ij} = 0$ for $j = 1, 2, 3$ and $j \neq i$ and if moreover $q := \lim_{n\to\infty} \sqrt{h(n)}/g(n) = 0$ or ∞. If at least one of the numbers p_{ij} ($i \neq j$) and q is finite but not zero, the rescaled sequence is again of type I or II except for an exceptional case where the numbers $b_1, c_1, c_2, \beta, p_{ij}, q$ satisfy one of a couple of given algebraic equations, in which case the rescaled sequence is of type III. In the non-exceptional case, the rescaled matrices can be diagonalized with Corollary 3.4 provided that the sequences $\{g_{n+1}/g_n\}$, $\{k(n)/g_n\}$, $\{h(n)/g_n^2\}$ and either $\{g(n)/g_n\}$ (if $\lim_{n\to\infty} g(n)/g_n$ is finite) or $\{g(n)\left(\frac{1}{g_{n+1}} - \frac{1}{g_n}\right)\}$ (if $\lim_{n\to\infty} g(n)/g_n = \infty$) are of bounded variation, as well as $\{G_{n+1}^{-1} P_n G_n/g_n\}$. In the exceptional case, another rescaling is needed.

Proof:

(i.) $b_1 \neq b_2$. First, we assume that $g(n) = o(\sqrt{h(n)})$ so $q = 0$. In particular, $\beta \neq 0$. In this case we let $g_n = \sqrt{h(n)}$ and $G_n = \mathrm{diag}(g_n^{-1}, 1)$. Then the rescaled matrices are

$$G_{n+1}^{-1} M_n G_n = \sqrt{h(n)} \begin{pmatrix} 0 & 1 \\ \beta & 0 \end{pmatrix} + g(n) \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix} + k(n) \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} + P_n' \qquad (5.4)$$

where $P_n' = G_{n+1}^{-1} P_n G_n +$ terms of order $\sqrt{h(n+1)} - \sqrt{h(n)}$, $g(n)\left(\sqrt{h(n+1)}/\sqrt{h(n)} - 1\right)$ and $k(n)\left(\sqrt{h(n+1)}/\sqrt{h(n)} - 1\right)$. The first term (with coefficient $\sqrt{h(n)}$) is the leading term; the leading matrix is diagonalizable and we have a type I sequence $\{G_{n+1}^{-1} M_n G_n/\sqrt{h(n)}\}$. If $q = \lim_{n\to\infty} \sqrt{h(n)}/g(n) \neq 0$, then the leading matrix also contains a contribution from the $g(n)$-term. The leading matrix is then $\begin{pmatrix} b_1 & q \\ q\beta & b_2 \end{pmatrix}$ and we have type I again, unless

$(b_1-b_2)^2+4q^2\beta=0$. Then we have type III, and we need to repeat the procedure. An example is given at the end of this section.

If $\sqrt{h(n)}=o(g(n))$ (in particular, this is the case if $\beta=0$), then let $g_n=g(n)$. The rescaled sequence is

$$G_{n+1}^{-1}M_nG_n=g(n)\begin{pmatrix}b_1&1\\0&b_2\end{pmatrix}+k(n)\begin{pmatrix}c_1&0\\0&c_2\end{pmatrix}+\frac{h(n)}{g(n)}\begin{pmatrix}0&0\\\beta&0\end{pmatrix}+P_n',$$

where $P_n'=G_{n+1}^{-1}P_nG_n+$ terms of order $g(n+1)-g(n)$ and $k(n)\,(g(n+1)/g(n)-1)$. The leading term is the term with coefficient $g(n)$ and is always diagonalizable. The rescaled sequence $\{G_{n+1}^{-1}M_nG_n/g(n)\}$ is of type I.

(ii.) $b_1=b_2=:b\neq 0$. First suppose that p_{21} and p_{31} are finite. We can take $g_n=1/G(n-1)$. Notice that $1/G(n-1)=o(g(n))$. The rescaled matrices are now

$$G_{n+1}^{-1}M_nG_n=bg(n)I+\frac{1}{G(n)}\begin{pmatrix}-b&1\\0&0\end{pmatrix}+k(n)\begin{pmatrix}c_1&0\\0&c_2\end{pmatrix}+h(n)G(n-1)\begin{pmatrix}0&0\\\beta&0\end{pmatrix}+P_n',$$

where $P_n'=G_{n+1}^{-1}P_nG_n+$ a term of order $k(n)/(g(n)G(n))$. If $p_{21}=p_{31}=0$, then the leading term is the term with coefficient $1/G(n)$. The leading matrix is diagonalizable and the type of the sequence $\{G_{n+1}^{-1}M_nG_n/g(n)\}$ is II. If either of the numbers p_{21} or p_{31} is not zero, the leading matrix $\begin{pmatrix}-b+p_{21}c_1&1\\p_{31}^2\beta&p_{21}c_2\end{pmatrix}$ also contains a contribution from other terms. The type is again II unless $(p_{21}(c_1-c_2)-b)^2+4\beta p_{31}^2=0$; then the leading matrix is non-diagonalizable and we have a type III sequence.

Now suppose that $p_{12}=0$ and p_{32} is finite, so we can take $g_n=k(n)$. Notice that $k(n)=o(g(n))$ by definition. The rescaled sequence is now

$$G_{n+1}^{-1}M_nG_n=g(n)\begin{pmatrix}b&0\\0&b\end{pmatrix}+k(n)\begin{pmatrix}c_1&1\\0&c_2\end{pmatrix}+\frac{h(n)}{k(n)}\begin{pmatrix}0&0\\\beta&0\end{pmatrix}+P_n'$$

where $P_n'=G_{n+1}^{-1}P_nG_n+$ terms of order $k(n+1)-k(n)$, $(k(n+1)-k(n))(g(n)/k(n))$. The leading matrix is $\begin{pmatrix}c_1&1\\0&c_2\end{pmatrix}$, which is diagonalizable and we have a type II sequence $\{G_{n+1}^{-1}M_nG_n/g(n)\}$. If $p_{32}\neq 0$, the leading matrix is $\begin{pmatrix}c_1&1\\p_{32}^2\beta&c_2\end{pmatrix}$ which is diagonalizable if $(c_1-c_2)^2+4\beta p_{32}^2\neq 0$. In the latter case we have a type II sequence $\{G_{n+1}^{-1}M_nG_n/k(n)\}$. On the other hand, if $(c_1-c_2)^2+4\beta p_{32}^2=0$, then the sequence is of type III and another rescaling is needed.

Finally, if $p_{13}=p_{23}=0$ we let $g_n=\sqrt{h(n)}$. The rescaled matrices are of the form (5.4). If $\lim_{n\to\infty}\sqrt{h(n)}/g(n)=0$, then the leading matrix $\begin{pmatrix}0&1\\\beta&0\end{pmatrix}$ is diagonalizable and we have a type II sequence $\{G_{n+1}^{-1}M_nG_n/g(n)\}$. If $\lim_{n\to\infty}\sqrt{h(n)}/g(n)=\infty$, then the leading matrix $\begin{pmatrix} 0 & 1 \\ \beta & 0 \end{pmatrix}$ is diagonalizable and we have a type I sequence $\{G_{n+1}^{-1} M_n G_n / \sqrt{h(n)}\}$. If $\lim_{n\to\infty} \sqrt{h(n)}/g(n) = q \neq 0$, then the leading matrix $\begin{pmatrix} b & q \\ q\beta & b \end{pmatrix}$ is diagonalizable and we have a type I sequence $\{G_{n+1}^{-1} M_n G_n / g(n)\}$.

In all of these cases where the rescaled matrix sequences are of type I or type II (so excepting the exceptional cases), the assumptions on the bounded variation of the sequences mentioned in Proposition 5.3 allow Corollary 3.4 to be applied to the rescaled sequences $\{G_{n+1}^{-1} M_n G_n / g(n)\}$ and $\{G_{n+1}^{-1} M_n G_n / \sqrt{h(n)}\}$ respectively. This concludes the argument. □

Finally we give an example of case 5.

Example 5.3: Let $M_n = \begin{pmatrix} \frac{1}{\sqrt{n}} + \mathcal{O}(\frac{1}{n\sqrt{n}}) & 1 - \frac{1}{\sqrt{n}} + \mathcal{O}(\frac{1}{n}) \\ \frac{\beta}{n} + \mathcal{O}(\frac{1}{n^2}) & \frac{2}{\sqrt{n}} + \frac{1}{n} + \mathcal{O}(\frac{1}{n\sqrt{n}}) \end{pmatrix}$ where the coefficients of M_n are power series in $1/\sqrt{n}$ and where $\beta \neq 0$. We split M_n into a main part

$$M_n^{(0)} = \begin{pmatrix} \frac{1}{\sqrt{n}} & 1 - \frac{1}{\sqrt{n}} \\ \frac{\beta}{n} & \frac{2}{\sqrt{n}} + \frac{1}{n} \end{pmatrix} = (1 - \frac{1}{\sqrt{n}}) \left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + \frac{1}{n} \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} + (\frac{1}{n} + \frac{1}{n\sqrt{n}}) \begin{pmatrix} 0 & 0 \\ \beta & 0 \end{pmatrix} \right)$$

and a perturbation term $P_n = \begin{pmatrix} \mathcal{O}(\frac{1}{n\sqrt{n}}) & \mathcal{O}(\frac{1}{n}) \\ \mathcal{O}(\frac{1}{n^2}) & \mathcal{O}(\frac{1}{n\sqrt{n}}) \end{pmatrix}$. The matrix sequence $\{M_n\}$ is of type IV with $g(n) = \frac{1}{\sqrt{n}}$, $k(n) = \frac{1}{n}$, $h(n) = \frac{1}{n} + \frac{1}{n\sqrt{n}}$, $b_1 = 1, b_2 = 2$, and $c_1 = 0, c_2 = 1$. Since $\lim_{n\to\infty} g(n)/\sqrt{h(n)} = 1$, we may take $g_n = \frac{1}{\sqrt{n}}$ and $G_n = \operatorname{diag}(g_n^{-1}, 1)$. The rescaled sequence is

$$G_{n+1}^{-1} M_n G_n = (1 - \frac{1}{\sqrt{n}}) \begin{pmatrix} \frac{1}{\sqrt{n+1}} + \frac{1}{n} + \dots & \frac{1}{\sqrt{n+1}} \\ \frac{\beta}{\sqrt{n}} + \frac{\beta}{n} & \frac{2}{\sqrt{n}} + \frac{3}{n} \end{pmatrix} + G_{n+1}^{-1} P_n G_n =$$

$$= (1 - \frac{1}{\sqrt{n}}) \left(\frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 \\ \beta & 2 \end{pmatrix} + \frac{1}{n} \begin{pmatrix} 1 & 0 \\ \beta & 3 \end{pmatrix} + P_n' \right)$$

where P_n' is a matrix whose terms are power series in $1/\sqrt{n}$ and such that $P_n' = \mathcal{O}(1/n\sqrt{n})$. If $\beta \neq -\frac{1}{4}$, the sequence $\{\sqrt{n} G_{n+1}^{-1} M_n G_n\}$ is of type I and by Corollary 3.4 we can find a sequence $\{H_n\}$ converging to the matrix of eigenvectors $H = \begin{pmatrix} 2 & 2 \\ 1 - \sqrt{1+4\beta} & 1 + \sqrt{1+4\beta} \end{pmatrix}$ of the leading matrix $\begin{pmatrix} 1 & 1 \\ \beta & 2 \end{pmatrix}$ such that

$$H_{n+1}^{-1} G_{n+1}^{-1} M_n G_n H_n = \operatorname{diag}(a_1(n), a_2(n))$$

where $a_1(n) = 1 + (\frac{3}{2} + \frac{1}{2}\sqrt{1+4\beta})\frac{1}{\sqrt{n}} + \ldots$ and $a_2(n) = 1 + (\frac{3}{2} - \frac{1}{2}\sqrt{1+4\beta})\frac{1}{\sqrt{n}} + \ldots$ are the eigenvalues of $G_{n+1}^{-1} M_n G_n$. Furthermore, by (3.5)

$$\|H_n - H\| = \mathcal{O}\left(\max\left(\sum_{\ell=n}^{\infty} \frac{1}{\ell\sqrt{\ell}}, e^{-2A\sqrt{\ell}+\ldots} \sum_{\ell=0}^{n-1} \frac{1}{\ell\sqrt{\ell}} e^{2A\sqrt{\ell}+\ldots}, e^{2A\sqrt{\ell}+\ldots} \sum_{\ell=n}^{\infty} \frac{1}{\ell\sqrt{\ell}} e^{-2A\sqrt{\ell}+\ldots}\right)\right) = \mathcal{O}(\frac{1}{\sqrt{n}})$$

for some number $C_0 > 0$, where $A = \mathrm{Re}\ \sqrt{4\beta+1}$ is the real part of the difference of the eigenvalues of the matrix $\begin{pmatrix} 1 & 1 \\ \beta & 2 \end{pmatrix}$ and the notation $e^{2A\sqrt{\ell}+\ldots}$ is used to denote $\prod_{h=0}^{\ell} \left|\frac{a_1(h)}{a_2(h)}\right| = e^{2A\sqrt{\ell}+\mathcal{O}(\ln \ell)}$.

We next consider the case that $\beta = -1/4$.

Letting $C = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$, we have

$$C^{-1} G_{n+1}^{-1} M_n G_n C = (1 - \frac{1}{\sqrt{n}}) \left(\frac{1}{\sqrt{n}} \begin{pmatrix} \frac{3}{2} & \frac{1}{4} \\ 0 & \frac{3}{2} \end{pmatrix} + \frac{1}{n} \begin{pmatrix} \frac{5}{2} & \frac{1}{4} \\ 3 & \frac{3}{2} \end{pmatrix} + C^{-1} P_n' C \right) =$$

$$= a(n) \left(I + \frac{1}{6} f(n) \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + h(n) \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + g(n) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + P_n'' \right)$$

where

$$a(n) = (1 - \frac{1}{\sqrt{n}})(\frac{3}{2\sqrt{n}} + \frac{2}{n}),\ f(n) = 1 - \frac{1}{3\sqrt{n}} + \mathcal{O}(\frac{1}{n}),$$

$$h(n) = \frac{2}{\sqrt{n}} + \mathcal{O}(\frac{1}{n}),\ g(n) = \frac{1}{3\sqrt{n}} + \mathcal{O}(\frac{1}{n}),\ P_n'' = \mathcal{O}(\frac{1}{n}).$$

The rescaled sequence $\{C^{-1} G_{n+1}^{-1} M_n G_n C / a(n)\}$ is of type IIIb. $F(n) = n - \frac{2}{3}\sqrt{n} + \mathcal{O}(\ln n)$.

$$g_n = \max(1/F(n), g(n)/f(n), \sqrt{h(n)/f(n)}) = \sqrt{h(n)/f(n)} = \sqrt{2} n^{-1/4} + \ldots.$$

According to the same reasoning as in Remark 5.2, we may take $g_n = n^{-1/4}$. So we let the rescaling sequence be $\{G_n' = \begin{pmatrix} n^{1/4} & 0 \\ 0 & 1 \end{pmatrix}\}$. Writing $\Gamma_n = G_n C G_n'$ we obtain a matrix sequence with terms

$$\Gamma_{n+1}^{-1} M_n \Gamma_n = (1 - \frac{1}{\sqrt{n}})(\frac{3}{2\sqrt{n}} + \frac{2}{n}) \left(I + n^{-1/4} \begin{pmatrix} 0 & \frac{1}{6} \\ 2 & 0 \end{pmatrix} + \frac{1}{\sqrt{n}} \begin{pmatrix} 1/3 & 0 \\ 0 & -1/3 \end{pmatrix} + \mathcal{O}(n^{-3/4}) \right).$$

Since all terms are now power series in $n^{-1/4}$, this matrix may be diagonalized by Corollary 3.4, i.e. there exists a sequence $\{H_n\}$, converging to the matrix $H = \begin{pmatrix} \sqrt{3} & -\sqrt{3} \\ 6 & 6 \end{pmatrix}$ of eigenvectors of $\begin{pmatrix} 0 & \frac{1}{6} \\ 2 & 0 \end{pmatrix}$, such that

$$H_{n+1}^{-1} \Gamma_{n+1}^{-1} M_n \Gamma_n H_n = (1 - \frac{1}{\sqrt{n}})(\frac{3}{2\sqrt{n}} + \frac{2}{n}) \begin{pmatrix} 1 - \frac{1}{3}\sqrt{3} n^{-1/4} + \ldots & 0 \\ 0 & 1 + \frac{1}{3}\sqrt{3} n^{-1/4} + \ldots \end{pmatrix}$$

where $1 \pm \frac{1}{3}\sqrt{3}n^{-1/4} + \ldots$ are the eigenvalues of the matrix $\Gamma_{n+1}^{-1} M_n \Gamma_n / a(n)$. Furthermore, by (3.5) it follows that

$$\|H_n - H\| = \mathcal{O}(n^{-1/4}).$$

6 OUTLOOK AND FURTHER RESEARCH.

The obvious next step is to study matrix sequences $\{M_n\}$ where the M_n are of higher order $k \geq 3$. It remains to be seen if again rescaling matrices can be explicitly found for general k or at least if something general can be said about the rescaling procedure. The case that the matrices M_n are complex matrices of order two, can also be studied. There will probably not be any additional difficulties except for the dichotomy condition (1.1), which must be checked separately, whereas in the real case, as we have seen, these conditions are automatically satisfied for the neat matrix sequences of section 5.

References

1 Z.Benzaid and D.A.Lutz, Asymptotic Representation of solutions of perturbed systems of linear difference equations, *Studies Appl.Math.* 77 (1987), 195-221.

2 R.J.Kooman, Decomposition of Matrix Sequences, *Indag. Mathem., N.S.* 5(1) (1994), 61-79.

3 S.N.Elaydi, "An introduction to Difference Equations", Springer Verlag, 2005 (3rd ed.).

4 R.J.Kooman, Asymptotic Behaviour of Solutions of Linear Recurrences and Sequences of Möbius-Transformations, *J.Approx.Th.* 93 (1998), 1-58.

5 J.Janas and M.Moszynski, Spectral properties of Jacobi matrices by asymptotic analysis, *J.Approx.Th.* 120 (2003), 309-336.

6 S.N.Elaydi, Asymptotics for Linear Difference Equations II: Application *New trends in difference equations: Proceedings of the Fifth International Conference on Difference Equations, Temuco, Chile, Jan.2000*, Taylor and Francis 2002, 111-133.

7 P.D.Lax, "Linear Algebra", John Wiley and Sons, Inc. 1997.

8 H.M.Farkas, I.Kra, "Riemann Surfaces", Springer Verlag, 1992.

9 S.Lang, "Complex Analysis", Springer Verlag, 1998.

10 C.V.Coffman, Asymptotic behaviour of solutions of ordinary difference equations, *Trans.Amer.Mat.Soc.* 110 (1964), 22-51.