

# A note on the convergence theorem of the tridiagonal QR algorithm with Wilkinson's shift

Kensuke Aishima<sup>1</sup>

Received: 2 December 2014 / Revised: 23 March 2015 / Published online: 12 June 2015  
© The JJIAM Publishing Committee and Springer Japan 2015

**Abstract** We discuss the convergence rate of the QR algorithm with Wilkinson's shift for tridiagonal symmetric eigenvalue problems. It is well known that the convergence rate is theoretically at least quadratic, and practically better than cubic for most matrices. In an effort to derive the convergence rate, the limiting patterns of some lower right submatrices have been intensively investigated. In this paper, we first describe a new limiting pattern of the lower right 3-by-3 submatrix with a concrete example, and then prove that the convergence rate of this new pattern is strictly cubic. In addition, we stress that our analysis identifies three classes of the limiting patterns of the tridiagonal QR algorithm with Wilkinson's shift.

**Keywords** Numerical linear algebra · Eigensolver · QR algorithm · Symmetric tridiagonal matrices · Wilkinson's shift · Convergence rate

**Mathematics Subject Classification** 65F15 · 15A18

## 1 Introduction

The standard method for computing the eigenvalues of a real symmetric matrix  $A$  consists of two phases. First,  $A$  is transformed to a tridiagonal matrix  $T$  by an appropriate orthogonal similarity transformation. Then some iterative method is applied to  $T$  to compute its eigenvalues. There are several approaches in the second phase. Among

---

The author is supported by JSPS Grant-in-Aid for Young Scientists (Grant Number 25790096).

---

✉ Kensuke Aishima  
Kensuke\_Aishima@mist.i.u-tokyo.ac.jp

<sup>1</sup> The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo, Japan

them, the historical QR algorithm is still widely used as a reliable tool, particularly with Wilkinson’s shift for accelerating the convergence. In this paper we consider the convergence behavior of this algorithm.

Hence, we write a symmetric irreducible tridiagonal matrix as

$$T = \begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & \beta_{m-1} & \alpha_m \end{pmatrix}. \tag{1}$$

The eigenvalues of  $T$  are all distinct [7], which are denoted as  $\lambda_1 > \dots > \lambda_m$  here. The shifted QR algorithm is described as

$$T^{(n)} - s^{(n)}I = Q^{(n)}R^{(n)}, \tag{2}$$

$$T^{(n+1)} = R^{(n)}Q^{(n)} + s^{(n)}I \tag{3}$$

for  $n = 0, 1, \dots$ , where  $Q^{(n)}$  is orthogonal,  $R^{(n)}$  is upper triangular with nonnegative diagonal elements, and  $s^{(n)}$  is a shift. Similarly to (1), tridiagonal elements of  $T^{(n)}$  are denoted by

$$T^{(n)} = \begin{pmatrix} \alpha_1^{(n)} & \beta_1^{(n)} & & & \\ \beta_1^{(n)} & \alpha_2^{(n)} & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix}. \tag{4}$$

In (2), one way to determine an efficient shift  $s^{(n)}$  is to consider the lower right 2-by-2 submatrix, and pick its eigenvalue closer to  $\alpha_m^{(n)}$ . This is the so called Wilkinson shift.

For this algorithm, there is a long history of convergence analysis. Global convergence (i.e.  $\lim_{n \rightarrow \infty} \beta_{m-1}^{(n)} = 0$  for any initial matrix) was first proved by Wilkinson in [10], and then another elegant proof was given by Hoffmann–Parlett [1]. Regarding the convergence rate, Wilkinson [10] theoretically proved that it is at least quadratic, after which Hoffmann–Parlett [1] showed that for most matrices convergence better than cubic is achieved. Today, the convergence scenarios are classified by the asymptotic behavior of the lower right elements as follows.

- if  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = 0$ , then quadratic convergence only;
- if  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = D \neq 0$ , then cubic convergence at least;

if, in addition,  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = 0$ , then the rate is better than cubic.

This view was not explicitly described in any literature, as far as the authors know. However, in actual fact, this view was first implied by Hoffmann–Parlett [1], and then followed by other researchers [7,8,11] (for more detail, see Sects. 2 and 3). In the present study, we focus on the situation  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = D \neq$

0,  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = C \neq 0$  where the rate is at least cubic. This situation is relatively exceptional. In fact, concrete matrices that belong to this case have not been found in any literature, and most initial matrices result in the case  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = 0$  where the rate is better than cubic. Mathematically strictly speaking, however, it is desired to determine if such an exceptional example actually exists and, if so, derive its exact convergence rate.

In this paper, we first point out an example where  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = D \neq 0$ ,  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = C \neq 0$  are observed. Next, we prove  $\lim_{n \rightarrow \infty} (\alpha_{m-2}^{(n)} - \alpha_m^{(n)}) = -D$  in this case. In other words, we prove that there exists a matrix with the new limiting pattern of the lower right 3-by-3 matrix

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-2}^{(n)} & \beta_{m-2}^{(n)} & 0 \\ \beta_{m-2}^{(n)} & \alpha_{m-1}^{(n)} & \beta_{m-1}^{(n)} \\ 0 & \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_l - D & C & 0 \\ C & \lambda_l + D & 0 \\ 0 & 0 & \lambda_l \end{pmatrix},$$

where  $C \neq 0$  and  $D \neq 0$ . Although several literatures imply that the rate of this case is at least cubic, we theoretically show a stronger result that the rates of all the examples belonging to this limiting pattern are “strictly” cubic. Our analysis is based on the result by Jiang–Zhang [3, Lemma 1]. Also note that our analysis identifies three classes of the limiting patterns of the tridiagonal QR algorithm with Wilkinson’s shift.

This paper is organized as follows. After a brief summary of Wilkinson’s results [10] in Sect. 2, we summarize the limiting matrix patterns owing to Hoffmann–Parlett [1] in Sect. 3. Then in Sect. 4, we present an example where the new limiting pattern above is observed. In Sect. 5, we review a part of the convergence analysis by Jiang–Zhang [3] for our convergence analysis. In Sect. 6, we prove that the convergence rate for the new limiting pattern is strictly cubic, and we give a convergence theorem that covers all the 3-by-3 limiting submatrices. Section 7 is devoted to the conclusions.

*Remark 1* Strictly speaking, in order for our classifications by limiting matrices to work for every initial matrix, it should be also proved that the 3-by-3 submatrix in question always tends to a constant matrix (without exhibiting any oscillatory behaviors). Actually this holds true. This fact might have been noticed by the experts in this research field because its proof is almost the same as those by [2, 6, 9] for the unshifted QR algorithm. However, the present authors do not know any reference where the proof for the shifted algorithm is explicitly stated. For the readers’ convenience, in the present paper a stronger result stating all of the tridiagonal elements in fact tend to constants is shown in Appendix A.

*Remark 2* In this paper, we assume that all the diagonal elements of  $R^{(n)}$  in (2) are nonnegative in the QR algorithm. Even if some diagonal elements of  $R^{(n)}$  are not nonnegative, the absolute values of the subdiagonal elements  $|\beta_1^{(n)}|, \dots, |\beta_{m-1}^{(n)}|$  behave in the same way. We explain it more precisely below. Let  $\hat{T}^{(n)}$  ( $n = 0, 1, \dots$ ) be the tridiagonal matrices computed by the QR algorithm where some diagonal elements of  $R^{(n)}$  are not nonnegative. Then  $\hat{T}^{(n)} = D^{(n)}T^{(n)}D^{(n)}$  for  $n = 0, 1, \dots$  are satisfied where  $D^{(n)}$  are certain diagonal sign matrices. This is proved by induction.

Obviously  $D^{(0)} = I$  is a sign matrix. Note that the Wilkinson’s shift  $s^{(n)}$  for  $\hat{T}^{(n)}$  is the same as  $T^{(n)}$ . Hence, any QR decomposition of  $\hat{T}^{(n)} - s^{(n)}I$  is described as  $(D^{(n)}Q^{(n)}D^{(n+1)})(D^{(n+1)}R^{(n)}D^{(n)})$  for a sign matrix  $D^{(n+1)}$ . Then, we have  $\hat{T}^{(n+1)} = D^{(n+1)}T^{(n+1)}D^{(n+1)}$ . Therefore, the absolute values of the subdiagonal elements  $\hat{T}^{(n+1)}$  are equivalent to those of  $T^{(n+1)}$ . Thus, in some sense, we also cover the convergence analysis of  $\hat{T}^{(n)}$  by replacing  $\beta_1^{(n)}, \dots, \beta_{m-1}^{(n)}$  with  $|\beta_1^{(n)}|, \dots, |\beta_{m-1}^{(n)}|$  in the following sections.

### 2 QR algorithm with Wilkinson’s shift

In this section, Wilkinson’s convergence theorems in [10] are summarized. In the QR algorithm with Wilkinson’s shift, the global convergence (in the sense that  $\lim_{n \rightarrow \infty} \beta_{m-1}^{(n)} = 0$  for any initial matrix) is theoretically guaranteed, and the convergence rate is at least quadratic as the next theorem indicates.

**Theorem 1** (Wilkinson [10]) *Suppose the QR algorithm with Wilkinson’s shift is applied to an irreducible tridiagonal matrix  $T$ . Then we have*

$$\lim_{n \rightarrow \infty} \beta_{m-1}^{(n)} = 0, \quad \left| \beta_{m-1}^{(n+1)} \right| = O \left( \left| \beta_{m-1}^{(n)} \right|^2 \right) \tag{5}$$

for all initial matrices.

Here we like to point out that, although not explicitly written, from the discussions in [10] we can deduce two important facts.

First, Wilkinson’s paper [10] implies  $\lim_{n \rightarrow \infty} \alpha_m^{(n)} = \lambda_l$ . In other words,  $\alpha_m^{(n)}$  converges to a fixed eigenvalue  $\lambda_l$ . It is easy to see that (21) on p. 412 of [10] corresponds to

$$\alpha_m^{(n+1)} = s^{(n)} + r_{mm}^{(n)} \cos \theta^{(n)}, \tag{6}$$

where  $r_{mm}^{(n)}$  is the lower right element of  $R^{(n)}$ , and  $\theta^{(n)}$  is the last angle of the Givens rotation to delete the last subdiagonal element. Obviously,  $r_{mm}^{(\infty)} = 0$ . Hence, we have

$$\lim_{n \rightarrow \infty} \left| \alpha_m^{(n+1)} - \alpha_m^{(n)} \right| \leq \lim_{n \rightarrow \infty} \left| \alpha_m^{(n+1)} - s^{(n)} \right| + \lim_{n \rightarrow \infty} \left| \alpha_m^{(n)} - s^{(n)} \right| = 0. \tag{7}$$

This shows that  $\alpha_m^{(n)}$  converges to a fixed eigenvalue.

Second, Wilkinson’s paper [10] implies the cubic convergence of  $\beta_{m-1}^{(n)}$  in a certain case. The first inequality in (62) on p. 418 of [10] corresponds to

$$\left| \beta_{m-1}^{(n+1)} \right| \leq \frac{\left| \beta_{m-1}^{(n)} \right|^3}{\left| \alpha_{m-1}^{(n)} - s^{(n)} \right| \left( \delta - 3\beta_{m-1}^{(n)} \right)} + \frac{\left| \beta_{m-1}^{(n)} \right|^3}{\left( \delta - 3\beta_{m-1}^{(n)} \right)^2} \tag{8}$$

for all sufficiently large  $n$ , where  $\delta = \min_{i \neq l} |\lambda_i - \lambda_l|$ . It is easy to see that, if  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - s^{(n)}) = D \neq 0$ , then  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-1}^{(n)}|^3)$ . Therefore, the following theorem is established.

**Theorem 2** *Suppose the QR algorithm with Wilkinson’s shift is applied to an irreducible tridiagonal matrix  $T$ . Then  $\lim_{n \rightarrow \infty} \alpha_m^{(n)} = \lambda_l$  holds for some  $l$ . If the lower right 2-by-2 submatrix converges to the limiting matrix*

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-1}^{(n)} & \beta_{m-1}^{(n)} \\ \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_l + D & 0 \\ 0 & \lambda_l \end{pmatrix}, \tag{9}$$

where  $D \neq 0$ , then  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-1}^{(n)}|^3)$ . If the lower right 2-by-2 submatrix converges to the limiting matrix

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-1}^{(n)} & \beta_{m-1}^{(n)} \\ \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_l & 0 \\ 0 & \lambda_l \end{pmatrix}, \tag{10}$$

then  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-1}^{(n)}|^2)$ .

### 3 Convergence rate analysis by Hoffmann–Parlett [1]

Actually, quite often it is observed that the convergence rate is better than cubic. This phenomena is mathematically described in the next theorem, which states that if the second lower right element  $\beta_{m-2}^{(n)}$  converges to 0, then the convergence rate of the lower right element  $\beta_{m-1}^{(n)}$  is better than cubic.

**Theorem 3** (Hoffmann–Parlett [1]) *Suppose the QR algorithm with Wilkinson’s shift is applied to an irreducible tridiagonal matrix  $T$ . If the lower right 3-by-3 submatrix converges to the limiting matrix*

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-2}^{(n)} & \beta_{m-2}^{(n)} & 0 \\ \beta_{m-2}^{(n)} & \alpha_{m-1}^{(n)} & \beta_{m-1}^{(n)} \\ 0 & \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix} = \begin{pmatrix} * & 0 & 0 \\ 0 & \lambda_k & 0 \\ 0 & 0 & \lambda_l \end{pmatrix}, \tag{11}$$

where  $k \neq l$ , then  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3)$ . If the lower right 3-by-3 submatrix converges to the limiting matrix

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-2}^{(n)} & \beta_{m-2}^{(n)} & 0 \\ \beta_{m-2}^{(n)} & \alpha_{m-1}^{(n)} & \beta_{m-1}^{(n)} \\ 0 & \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix} = \begin{pmatrix} * & C & 0 \\ C & \lambda_l & 0 \\ 0 & 0 & \lambda_l \end{pmatrix}, \tag{12}$$

where  $C \neq 0$ , then  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-1}^{(n)}|^2)$ .

More precisely, in (12), Hoffmann–Parlett do not assume  $\lim_{n \rightarrow \infty} \alpha_{m-1}^{(n)} = \lim_{n \rightarrow \infty} \alpha_m^{(n)} = \lambda_l$  but  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - s^{(n)}) = 0$ . However, it is easy to see that  $\lim_{n \rightarrow \infty} \alpha_{m-1}^{(n)} = \lim_{n \rightarrow \infty} \alpha_m^{(n)} = \lambda_l$  is mathematically equivalent to  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - s^{(n)}) = 0$  in the same way as the previous section. Combining the above theorem with Theorem 2, we have the following corollary.

**Corollary 1** *The convergence rate of  $\beta_{m-1}^{(n)}$  is summarized as follows.*

- if  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = 0$ , then quadratic convergence only;
- if  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = D \neq 0$ , then cubic convergence at least; if, in addition,  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = 0$ , then the rate is better than cubic.

In most cases, the limiting matrix satisfies (11), and thus the actual convergence speed is better than cubic [1,7,8,11]. In the case (11), Parlett demonstrates the asymptotic ratio  $|\beta_{m-1}^{(n+1)}|/|\beta_{m-2}^{(n)}|^2|\beta_{m-1}^{(n)}|^3$  approaches a computable limit under the assumption that  $\lim_{n \rightarrow \infty} \beta_{m-3}^{(n)} = 0$  and  $\lim_{n \rightarrow \infty} \alpha_{m-3}^{(n)} = \lambda_j$  for some  $j$  [7, Theorem 8.11.1]. The possibility of the loss of cubic convergence, i.e., the occurrence of (12), was discussed in several studies (see [1,7,8,11]). For example, Parlett [7] and Wang [8] show that, in the case (12), the limiting pattern of the lower right 3-by-3 matrix is

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-2}^{(n)} & \beta_{m-2}^{(n)} & 0 \\ \beta_{m-2}^{(n)} & \alpha_{m-1}^{(n)} & \beta_{m-1}^{(n)} \\ 0 & \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_l & C & 0 \\ C & \lambda_l & 0 \\ 0 & 0 & \lambda_l \end{pmatrix} \tag{13}$$

and  $\lim_{n \rightarrow \infty} \beta_{m-3}^{(n)} = 0$ . But it is still mathematically open whether or not there in fact exists such a matrix that leads to the case (except the case  $m = 3$ , for which a rigorous analysis is given in [4]; see Remark 3 below). However, we notice that the 3-by-3 limiting submatrix is described in a simple way (13) in the case  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = 0$ . Also in the case  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = 0$ , the corresponding limiting pattern is clearly described as (11).

We here focus on the other situation:  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = D \neq 0$  and  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = C \neq 0$ . Does it never happen? If it happens, what is the limiting pattern of the lower right 3-by-3 matrix including  $\lim_{n \rightarrow \infty} \alpha_{m-2}^{(n)}$ ? What is the exact convergence rate?

Below we show an answer to these questions; it turns out that such a situation can actually happen (we show an example), and there the convergence rate is strictly cubic.

### 4 A numerical experiment

Let us apply the QR algorithm with Wilkinson’s shift to a 101-by-101 matrix

$$T^{(0)} = \begin{pmatrix} 0 & 100 & & & \\ 100 & \ddots & \ddots & & \\ & \ddots & 0 & 100 & \\ & & 100 & 0 & 1 \\ & & & 1 & 0 \end{pmatrix}, \tag{14}$$

whose eigenvalues are  $0, \pm c_k$  ( $k = 1, \dots, 50$ ). Also let  $T_2^{(n)}$  be the lower right 3-by-3 submatrix of  $T^{(n)}$ ; for  $n = 0$  we see

$$T_2^{(0)} = \begin{pmatrix} 0 & 100 & \\ 100 & 0 & 1 \\ & 1 & 0 \end{pmatrix}. \tag{15}$$

Then we find the following convergence behavior of  $T_2^{(n)}$  for  $n = 3, 4, 5$ :

$$\begin{aligned} T_2^{(3)} &= \begin{pmatrix} 1.80 & 2.55 & \\ 2.55 & -1.81 & 2.55 \times 10^{-14} \\ & 2.55 \times 10^{-14} & -1.21 \times 10^{-28} \end{pmatrix}, \\ T_2^{(4)} &= \begin{pmatrix} 1.80 & 2.53 & \\ 2.53 & -1.81 & -1.96 \times 10^{-42} \\ & -1.96 \times 10^{-42} & -7.18 \times 10^{-85} \end{pmatrix}, \\ T_2^{(5)} &= \begin{pmatrix} 1.81 & 2.53 & \\ 2.53 & -1.81 & -8.88 \times 10^{-127} \\ & -8.88 \times 10^{-127} & -1.47 \times 10^{-253} \end{pmatrix}. \end{aligned}$$

Thus it tends to

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-2}^{(n)} & \beta_{m-2}^{(n)} & 0 \\ \beta_{m-2}^{(n)} & \alpha_{m-1}^{(n)} & \beta_{m-1}^{(n)} \\ 0 & \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_l - D & C & 0 \\ C & \lambda_l + D & 0 \\ 0 & 0 & \lambda_l \end{pmatrix}, \tag{16}$$

where  $\lambda_l = 0, C \approx 2.5, D \approx -1.8$ . The observed convergence rate of  $\beta_{m-1}^{(n)}$  is cubic:  $|\beta_{m-1}^{(n+1)}|/|\beta_{m-1}^{(n)}|^3 \approx 0.11$ . Despite the long history of the convergence analysis for Wilkinson’s shift, no similar example has been pointed out in the research field of numerical linear algebra, as far as the authors know. In this sense, one of the contribution of this paper is the finding of the matrix (16). In the following sections, we prove that the convergence rate of (16) is strictly cubic.

### 5 Convergence rate analysis by Jiang-Zhang [3]

For the convergence analysis of (16), we review the paper [3] by Jiang and Zhang in 1985. It implies  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3)$  in the case  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = 0$ , though the convergence analysis in [3] focuses on their own shift strategy where both Wilkinson’s shift and the Rayleigh quotient shift are exploited. The convergence rate analysis for the Wilkinson’s shift is described on p. 270 in [3] in the case  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = 0$ .

Let  $T_k^{(n)}$  be the  $k \times k$  leading principal submatrix of  $T^{(n)}$  and

$$d_k^{(n)} := \det \left( T_k^{(n)} - s^{(n)} I \right) \quad (k = 1, \dots, m). \tag{17}$$

The following lemma is crucial [3, Lemma2].

**Lemma 1** ([3]) *For any shift strategy in the tridiagonal QR algorithm,*

$$\left| \beta_{m-1}^{(n+1)} \right| = \frac{\left| \beta_{m-1}^{(n)} d_m^{(n)} \gamma^{(n)} \right|}{\left| \left( d_{m-1}^{(n)} \right)^2 + \left( \beta_{m-1}^{(n)} \gamma^{(n)} \right)^2 \right|} \tag{18}$$

hold for all  $n$ , where

$$\begin{aligned} \left( \gamma^{(n)} \right)^2 &= \left( d_{m-2}^{(n)} \right)^2 + \left( \beta_{m-2}^{(n)} d_{m-3}^{(n)} \right)^2 + \left( \beta_{m-2}^{(n)} \beta_{m-3}^{(n)} d_{m-4}^{(n)} \right)^2 + \dots \\ &+ \left( \beta_{m-2}^{(n)} \beta_{m-3}^{(n)} \dots \beta_2^{(n)} d_1^{(n)} \right)^2 + \left( \beta_{m-2}^{(n)} \beta_{m-3}^{(n)} \dots \beta_1^{(n)} \right)^2 \end{aligned} \tag{19}$$

for all  $n$ .

In order to derive the convergence rate, let us consider the right-hand side of (18). It is easy to see that

$$\lim_{n \rightarrow \infty} d_{m-1}^{(n)} = \prod_{i \neq l} (\lambda_i - \lambda_l) \neq 0 \tag{20}$$

in view of  $\lim_{n \rightarrow \infty} \beta_{m-1}^{(n)} = 0$  and  $s^{(\infty)} = \lambda_l$ . Noting (17) and the definition of the Wilkinson’s shift, we have

$$\begin{aligned} d_m^{(n)} &= \left[ \left( \alpha_m^{(n)} - s^{(n)} \right) \left( \alpha_{m-1}^{(n)} - s^{(n)} \right) - \left( \beta_{m-1}^{(n)} \right)^2 \right] d_{m-2}^{(n)} - \left( \alpha_m^{(n)} - s^{(n)} \right) \left( \beta_{m-2}^{(n)} \right)^2 d_{m-3}^{(n)} \\ &= - \frac{\left( \beta_{m-1}^{(n)} \right)^2 \left( \beta_{m-2}^{(n)} \right)^2 d_{m-3}^{(n)}}{\left( \alpha_{m-1}^{(n)} - s^{(n)} \right)}. \end{aligned} \tag{21}$$



Since we assume here that  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = 0, \lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - s^{(n)}) = \lambda_k - \lambda_l \neq 0$  holds for some  $k, l$ . Therefore, we have

$$|d_m^{(n)}| \sim \frac{|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^2 |d_{m-3}^{(n)}|}{|\lambda_k - \lambda_l|} \tag{22}$$

as  $n \rightarrow \infty$ . Furthermore,  $|\gamma^{(n)}|$  in (18) is bounded because  $\beta_k^{(n)} (1 \leq k \leq m - 1)$  and  $d_k^{(n)} (1 \leq k \leq m)$  in (19) are bounded for all  $n$ . In addition,  $\lim_{n \rightarrow \infty} \beta_{m-1}^{(n)} = 0$  holds. Therefore, we obtain  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3)$  from (18), (20) and (22).

Although the situation:  $\beta_{m-2}^{(\infty)} \neq 0, \alpha_{m-1}^{(\infty)} - s^{(\infty)} \neq 0$  is not explicitly considered in [3], it is easy to see from the above discussion that  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-1}^{(n)}|^3)$  is realized in this case. In the next section, we prove that the convergence rate of  $\beta_{m-1}^{(n)}$  is strictly cubic based on Lemma 1.

### 6 Convergence theorem covering the new limiting pattern

In this section, we prove that all the possible limiting patterns are (11), (13), and (16). Note that (16) is the new pattern. Based on Lemma 1, we prove that the convergence rate of (16) is strictly cubic. As a result, we have a convergence theorem that covers all the possible limiting patterns.

First of all, we note the following lemma. See Appendix B for the proof.

**Lemma 2** *If  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = C \neq 0$ , then  $|\lambda_l - \lambda_{l-1}| = |\lambda_l - \lambda_{l+1}|$  holds. Moreover, the lower right 3-by-3 matrix of  $T^{(n)}$  converges as*

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-2}^{(n)} & \beta_{m-2}^{(n)} & 0 \\ \beta_{m-2}^{(n)} & \alpha_{m-1}^{(n)} & \beta_{m-1}^{(n)} \\ 0 & \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_l - D & C & 0 \\ C & \lambda_l + D & 0 \\ 0 & 0 & \lambda_l \end{pmatrix} \tag{23}$$

for  $D \in \mathbb{R}$  such that  $\sqrt{C^2 + D^2} = |\lambda_l - \lambda_{l-1}| = |\lambda_l - \lambda_{l+1}|$ .

As stated above, [7, 8] showed that, if  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = C \neq 0$  and  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = D = 0$ , then the limiting pattern of the 3-by-3 matrix is (13). Lemma 2 is a generalization of this result to  $D \in \mathbb{R}$ .

Fortunately, in the case  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = D \neq 0$ , we can obtain  $\lim_{n \rightarrow \infty} |\beta_{m-1}^{(n+1)}| / (|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3)$  using (18) in Lemma 1. Recall that  $\lambda_k$  is the closest eigenvalue to  $\lambda_l$ . Let  $\lambda_j$  be the second closest one. We derive the following lemma from (18).

**Lemma 3** *If  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = D \neq 0$ , then*

$$\lim_{n \rightarrow \infty} \frac{|\beta_{m-1}^{(n+1)}|}{|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3} = \frac{|\rho_2|}{|\lambda_k - \lambda_l|^3 |\lambda_j - \lambda_l| |\rho_1|}, \tag{24}$$

where

$$|\rho_1| = |D|/|\lambda_k - \lambda_l|, \tag{25}$$

$$|\rho_2| = \sqrt{1 - |\beta_{m-3}^{(\infty)}|^2 / |\lambda_j - \lambda_l|^2}. \tag{26}$$

In addition,

$$0 < |\rho_1| = |D|/|\lambda_k - \lambda_l| \leq 1 \tag{27}$$

holds. Even if  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{|\beta_{m-1}^{(n+1)}|}{|\beta_{m-1}^{(n)}|^2} \leq \frac{1}{|\lambda_k - \lambda_l|} \tag{28}$$

holds.

See Appendix B for the proof. In fact, (24) is an extension of [7, Theorem 8.11.1]:

$$\lim_{n \rightarrow \infty} \frac{|\beta_{m-1}^{(n+1)}|}{|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3} = \frac{1}{|\lambda_k - \lambda_l|^3 |\lambda_j - \lambda_l|}$$

under the assumption:  $\alpha_{m-2}^{(\infty)} = \lambda_j$ ,  $\alpha_{m-1}^{(\infty)} = \lambda_k$ ,  $\beta_{m-2}^{(\infty)} = 0$ ,  $\beta_{m-3}^{(\infty)} = 0$ . In addition, (24) also covers the answer to [7, Exercise 8.11.3]. In other words, in the case  $\beta_{m-2}^{(\infty)} = 0$ ,  $\alpha_{m-3}^{(\infty)} = \lambda_l (= \alpha_m^{(\infty)})$ ,  $\beta_{m-3}^{(\infty)} \neq 0$ , we have

$$\lim_{n \rightarrow \infty} \frac{|\beta_{m-1}^{(n+1)}|}{|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3} = \frac{\sqrt{1 - |\beta_{m-3}^{(\infty)}|^2 / |\lambda_j - \lambda_l|^2}}{|\lambda_k - \lambda_l|^3 |\lambda_j - \lambda_l|} = 0,$$

where the first equality is due to (24), (25) and (26), the second equality is due to (69) in Appendix B. Also note that (28) is proved by Wilkinson [10]. We prove this based on Lemma 1 in Appendix B.

Using Lemmas 2 and 3, we have the following classification.

**Theorem 4** Suppose the QR algorithm with Wilkinson’s shift is applied to an irreducible tridiagonal matrix  $T$ . Let  $\lambda_l = \alpha_m^{(\infty)}$ , and  $\lambda_k$  be the closest eigenvalue to  $\lambda_l$ . Moreover, let  $\lambda_j$  be the second closest one. All the possible limiting 3-by-3 matrices are divided into three patterns. Based on the limiting patterns, the convergence rates are classified as follows. If the lower right 3-by-3 submatrix converges to the limiting matrix

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-2}^{(n)} & \beta_{m-2}^{(n)} & 0 \\ \beta_{m-2}^{(n)} & \alpha_{m-1}^{(n)} & \beta_{m-1}^{(n)} \\ 0 & \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix} = \begin{pmatrix} * & 0 & 0 \\ 0 & \lambda_k & 0 \\ 0 & 0 & \lambda_l \end{pmatrix}, \tag{29}$$

where  $k \neq l$ , then the convergence rate is better than cubic:

$$\lim_{n \rightarrow \infty} \frac{|\beta_{m-1}^{(n+1)}|}{|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3} = \frac{|\rho_2|}{|\lambda_k - \lambda_l|^3 |\lambda_j - \lambda_l|}, \tag{30}$$

$$|\rho_2| = \sqrt{1 - |\beta_{m-3}^{(\infty)}|^2 / |\lambda_j - \lambda_l|^2} \leq 1. \tag{31}$$

If the lower right 3-by-3 submatrix converges to the limiting matrix

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-2}^{(n)} & \beta_{m-2}^{(n)} & 0 \\ \beta_{m-2}^{(n)} & \alpha_{m-1}^{(n)} & \beta_{m-1}^{(n)} \\ 0 & \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_l - D & C & 0 \\ C & \lambda_l + D & 0 \\ 0 & 0 & \lambda_l \end{pmatrix}, \tag{32}$$

where  $C \neq 0, D \neq 0, \sqrt{C^2 + D^2} = |\lambda_k - \lambda_l| = |\lambda_j - \lambda_l|$ , then the convergence rate is strictly cubic:

$$\lim_{n \rightarrow \infty} \frac{|\beta_{m-1}^{(n+1)}|}{|\beta_{m-1}^{(n)}|^3} = \frac{|C|^2}{|\lambda_k - \lambda_l|^3 |D|}. \tag{33}$$

If the lower right 3-by-3 submatrix converges to the limiting matrix

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-2}^{(n)} & \beta_{m-2}^{(n)} & 0 \\ \beta_{m-2}^{(n)} & \alpha_{m-1}^{(n)} & \beta_{m-1}^{(n)} \\ 0 & \beta_{m-1}^{(n)} & \alpha_m^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_l & C & 0 \\ C & \lambda_l & 0 \\ 0 & 0 & \lambda_l \end{pmatrix}, \tag{34}$$

where  $C \neq 0, |C| = |\lambda_k - \lambda_l| = |\lambda_j - \lambda_l|$ , then the convergence rate is at least quadratic:

$$\limsup_{n \rightarrow \infty} \frac{|\beta_{m-1}^{(n+1)}|}{|\beta_{m-1}^{(n)}|^2} \leq \frac{1}{|\lambda_k - \lambda_l|}. \tag{35}$$

*Proof* If  $\beta_{m-2}^{(\infty)} = 0$ , then  $|D| = |\lambda_k - \lambda_l|$  in Lemma 3. Hence, we have (30). If  $\beta_{m-2}^{(\infty)} \neq 0$ , then  $\beta_{m-3}^{(\infty)} = 0$  because the block size of  $T^{(\infty)}$  is at most 2. See Appendix A for its proof. In addition,  $|\lambda_k - \lambda_l| = |\lambda_j - \lambda_l|$  is proved in Appendix A. Hence, we obtain (33) by Lemma 3. Lemma 3 also shows (35).  $\square$

As noted above, the previous analysis did not explicitly point out the limiting pattern (32), though the cubic convergence of this pattern is guaranteed by Theorem 2. In our experience, this limiting pattern is relatively rare compared to the previously known case (29), but not extremely rare so that it can be left outside our consideration (see also Remark 3 below). We here also like to emphasize that the above classification covers all the possible limiting patterns.

Our analysis in this section is based on the elegant result [3, Lemma2] that is derived by cleverly utilizing the special features of the shifted tridiagonal QR algorithm and Cramer’s rule. Note that Theorem 4 in this section is the stronger result than Corollary 1 that was previously implied by several literatures [1,7,8,11]. By carefully reading these literatures, we immediately find that all the existing proofs of Corollary 1 are also based on some special features of the tridiagonal QR algorithm. However, it is also possible to prove Corollary 1 based on the general gap theorem that is one of the useful perturbation theorems. In Appendix C, we describe another proof of Corollary 1 along this line: the gap theorem and the Jacobi transformation. The alternative proof is readily accessible to many readers in the research field of numerical linear algebra.

*Remark 3* The above result is closely related to the work by Leitte–Saldanha–Tomei [4], which considered the case of 3-by-3 matrices, and proved in the language of dynamical systems theory that there exists a 3-by-3 matrix that converges only quadratically by the QR iteration with Wilkinson’s shift. More precisely, they considered the matrix

$$\tilde{T} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{36}$$

and regarded the QR iterations with Wilkinson’s shift as maps generating a discrete dynamical system in the space of symmetric tridiagonal matrices with the same spectrum as  $\tilde{T}$  (see the original paper for the detail). Then they proved that there exists an open neighborhood of  $\tilde{T}$  such that (i) the iteration maps a point (matrix) back to the set, (ii) the convergence is strictly quadratic if it tends to  $\tilde{T}$  and cubic otherwise, and (iii) the Hausdorff dimension of the set of such initial points that leads to the quadratic convergence is 1.

Although the main topic of [4] is the existence of the quadratic cases, if we view the result from the opposite direction, it is also claiming that there are quite many initial matrices close to  $\tilde{T}$  resulting in cubic convergence (observe (iii) of [4], which states that the quadratic cases are “very thin” [4]). This strongly suggests that cubic convergence can be observed even when  $\beta_{m-2}^{(n)}$  does not tend to 0 (since it should stay around 1). In this way, the work [4] suggests a similar result as above in the case of 3-by-3 matrices, although explicit initial matrix examples are not given. By shrinking the size of the example (14) in the present paper, it is easy to obtain a 3-by-3 example

$$T^{(0)} = \begin{pmatrix} 0 & 100 & \\ 100 & 0 & 1 \\ & 1 & 0 \end{pmatrix}, \tag{37}$$

which actually tends to the form (16) cubically.

### 7 Conclusions and future works

We pointed out a limiting matrix pattern (32) where the convergence rate is strictly cubic as (33). Taking into account the fact that all the elements always converge (see Appendix A), we see that the classification by Theorem 4 covers all the possible scenarios, starting from any initial matrix.

It is still open, however, whether or not there actually exists a strictly quadratic case for 4-by-4 or larger matrices, i.e., if the limiting case (34) actually occurs. Furthermore, even if it is confirmed, still there remains the possibility that the estimate  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-1}^{(n)}|^2)$  is an overestimate, and the actual rate is cubic. These issues are left as future works. (Recall that for 3-by-3 matrices they have been completely settled in [4]; see Remark 3).

**Acknowledgments** The author is grateful to Professor Takayasu Matsuo and Professor Beresford Parlett for their valuable comments and suggestions. The author also thanks the anonymous reviewer for the helpful comments.

### Appendix A: Proof of convergence

We prove here that all the tridiagonal elements of  $T^{(n)}$  converge for any initial matrix  $T^{(0)}$ . To this end, we consider a general shift  $s^{(n)}$  satisfying the following conditions:

- (i) The shift  $s^{(n)}$  converges to a certain eigenvalue  $s^{(\infty)} = \lambda_l$ ;
- (ii)  $|s^{(n)} - \lambda_l| = o(c^n)$  for a positive constant  $c < 1$ .

Note that Wilkinson’s shift satisfies the two conditions: (i) has been proved by [8]; then the convergence rate by the shifts  $s^{(n)} \rightarrow \lambda_l$  is at least quadratic because  $|s^{(n)} - \lambda_l| \leq |s^{(n)} - \alpha_m^{(n)}| + |\alpha_m^{(n)} - \lambda_l| \leq 2|\beta_{m-1}^{(n)}|$  by the definition of Wilkinson’s shift and Gershgorin’s circle theorem, and then we have  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-1}^{(n)}|^2)$  by Wilkinson’s proof, which implies (ii).

In what follows, we prove that all the tridiagonal elements  $T^{(n)}$  always converge for such general shifts. The following convergence proof might have been noticed by the experts in this research field because its proof is almost the same as those by [2, 6, 9] for the unshifted QR algorithm. However, to the best of the authors’ knowledge, the proof for the shifted algorithm is not explicitly stated in any reference. For the readers’ convenience, we prove it as follows.

**Theorem 5** *Suppose that the QR algorithm with a general shift satisfying the above conditions (i) and (ii) is applied to an irreducible tridiagonal matrix  $T$ . Then,  $T^{(n)}$  converges to a block diagonal matrix whose block size is at most 2. Further, the lower*

right diagonal element  $T^{(n)}(m, m)$  converges to  $\lambda_l$  and the lower right off-diagonal element  $T^{(n)}(m, m - 1)$  converges to 0.

*Proof* First of all, we show several important facts for the convergence analysis. Similarly to the discussion in [6] and [9, Chapter8, §28], let  $\tilde{Q}^{(n)}, \tilde{R}^{(n)}$  be

$$\begin{aligned} \tilde{Q}^{(n)} &= Q^{(0)} \dots Q^{(n-1)}, \\ \tilde{R}^{(n)} &= R^{(n-1)} \dots R^{(0)}. \end{aligned}$$

By the orthogonal matrix  $\tilde{Q}^{(n)}, T^{(n)}$  is described as

$$T^{(n)} = \left(\tilde{Q}^{(n)}\right)^T T \tilde{Q}^{(n)}. \tag{38}$$

in view of  $T^{(n+1)} = R^{(n)}Q^{(n)} + s^{(n)}I = (Q^{(n)})^T(Q^{(n)}R^{(n)} + s^{(n)}I)Q^{(n)} = (Q^{(n)})^T T^{(n)} Q^{(n)}$ . Using (38), we see

$$\begin{aligned} \tilde{Q}^{(n)}\tilde{R}^{(n)} &= \tilde{Q}^{(n-1)}\left(T^{(n-1)} - s^{(n-1)}I\right)\tilde{R}^{(n-1)} \\ &= \left(T - s^{(n-1)}I\right)\tilde{Q}^{(n-1)}\tilde{R}^{(n-1)} \\ &= \left(T - s^{(n-1)}I\right)\dots\left(T - s^{(0)}I\right). \end{aligned} \tag{39}$$

Let  $p(l)$  denote a permutation of the indices  $l$  ( $l = 1, \dots, m$ ). Then in view of the condition (i) we can place the shifted eigenvalues in a descending order as

$$\left|\lambda_{p(1)} - s^{(\infty)}\right| \geq \dots \geq \left|\lambda_{p(m-1)} - s^{(\infty)}\right| > \left|\lambda_{p(m)} - s^{(\infty)}\right| = 0. \tag{40}$$

The last inequality follows from (i) (note that all the eigenvalues are distinct since  $T$  is irreducible).

Next, we focus on the eigendecomposition

$$T = X\Lambda X^T, \tag{41}$$

where  $X$  is the orthogonal matrix consisting of the eigenvectors and  $\Lambda$  is the diagonal matrix with the eigenvalues:  $\text{diag}(\lambda_{p(1)}, \dots, \lambda_{p(m)})$ . Then we see

$$\left(T - s^{(0)}I\right)\dots\left(T - s^{(n-1)}I\right) = X\Lambda^{(n)}X^T, \tag{42}$$

where

$$\Lambda^{(n)} = \left(\Lambda - s^{(0)}I\right)\dots\left(\Lambda - s^{(n-1)}I\right). \tag{43}$$

From (38) and (41), we have  $T^{(n)} = (\tilde{Q}^{(n)})^T X\Lambda X^T \tilde{Q}^{(n)}$ . If all the inequalities are strict in (40),  $T^{(n)}$  converges to  $\Lambda$ . Actually,  $T^{(n)} = (\tilde{Q}^{(n)})^T X\Lambda X^T \tilde{Q}^{(n)}$  always converges to a block diagonal matrix whose block size is at most 2. In order to prove it, we

firstly apply the LU factorization  $X^T = LU$ . Note that  $X^T$  constructed by the normalized eigenvectors of an irreducible tridiagonal matrix is always LU factorizable [2,5]. Hence, we see

$$(T - s^{(0)}I) \dots (T - s^{(n-1)}I) = X\Lambda^{(n)}L(\Lambda^{(n)})^{-1}\Lambda^{(n)}U. \tag{44}$$

Combining it with (39) we have

$$\tilde{Q}^{(n)}\tilde{R}^{(n)} = X\Lambda^{(n)}L(\Lambda^{(n)})^{-1}\Lambda^{(n)}U. \tag{45}$$

To reveal the relationship between  $\tilde{Q}^{(n)}$  and  $X$ , we discuss the behavior of the QR factorization of  $\Lambda^{(n)}L(\Lambda^{(n)})^{-1}$  as follows. Let  $D_{\Lambda^{(n)}}$  be an orthogonal matrix

$$\text{diag}\left(\prod_{l=0}^{n-1} \frac{|\lambda_{p(1)} - s^{(l)}|}{(\lambda_{p(1)} - s^{(l)})}, \dots, \prod_{l=0}^{n-1} \frac{|\lambda_{p(m)} - s^{(l)}|}{(\lambda_{p(m)} - s^{(l)})}\right). \tag{46}$$

It is easy to see that

$$D_{\Lambda^{(n)}}\Lambda^{(n)} = \text{diag}\left(\prod_{l=0}^{n-1} |\lambda_{p(1)} - s^{(l)}|, \dots, \prod_{l=0}^{n-1} |\lambda_{p(m)} - s^{(l)}|\right). \tag{47}$$

In the right-hand side of (45), we have

$$\Lambda^{(n)}L(\Lambda^{(n)})^{-1} = D_{\Lambda^{(n)}}^{-1}\left(D_{\Lambda^{(n)}}\Lambda^{(n)}L(D_{\Lambda^{(n)}}\Lambda^{(n)})^{-1}\right)D_{\Lambda^{(n)}}, \tag{48}$$

and by applying the QR factorization we see

$$D_{\Lambda^{(n)}}\Lambda^{(n)}L(D_{\Lambda^{(n)}}\Lambda^{(n)})^{-1} = P^{(n)}\Gamma^{(n)}, \tag{49}$$

where  $P^{(n)}$  is an orthogonal matrix,  $\Gamma^{(n)}$  is an upper triangular matrix whose diagonal elements are positive. Let  $D_U$  be an orthogonal matrix  $D_U = \text{diag}(|u_{11}|/u_{11}, \dots, |u_{mm}|/u_{mm})$ . Then we see

$$\tilde{Q}^{(n)} = XD_{\Lambda^{(n)}}^{-1}P^{(n)}D_U^{-1}, \tag{50}$$

$$\tilde{R}^{(n)} = D_U\Gamma^{(n)}D_{\Lambda^{(n)}}\Lambda^{(n)}U \tag{51}$$

from (45), (48), and (49). Therefore, we have

$$T^{(n)} = D_U(P^{(n)})^T\Lambda P^{(n)}D_U^{-1} \tag{52}$$

from (38) and (41).

Since our aim is to prove the convergence of all the elements of  $T^{(n)}$ , let us discuss the behavior of the orthogonal matrix  $P^{(n)}$  as  $n \rightarrow \infty$ . To this end, we focus on (49). The lower left elements are

$$\left( D_{\Lambda^{(n)}} \Lambda^{(n)} L \left( D_{\Lambda^{(n)}} \Lambda^{(n)} \right)^{-1} \right)_{ij} = l_{ij} \prod_{l=0}^{n-1} \left| \frac{\lambda_{p(j)} - s^{(l)}}{\lambda_{p(i)} - s^{(l)}} \right| \quad (i > j) \tag{53}$$

from (47). Obviously  $\lim_{n \rightarrow \infty} (D_{\Lambda^{(n)}} \Lambda^{(n)} L (D_{\Lambda^{(n)}} \Lambda^{(n)})^{-1})_{ij} = 0$ , when  $|\lambda_{p(i)} - s^{(\infty)}| > |\lambda_{p(j)} - s^{(\infty)}|$ . Otherwise, from (40) and the condition (ii), we have

$$\left| \frac{\lambda_{p(j)} - s^{(l)}}{\lambda_{p(i)} - s^{(l)}} \right| = \left| \frac{\lambda_{p(j)} - \lambda_{p(m)} + \lambda_{p(m)} - s^{(l)}}{\lambda_{p(i)} - \lambda_{p(m)} + \lambda_{p(m)} - s^{(l)}} \right| = 1 + o(c^l). \tag{54}$$

Since a sequence of the size  $o(c^l)$  with  $0 < c < 1$  absolutely converges,  $(D_{\Lambda^{(n)}} \Lambda^{(n)} L (D_{\Lambda^{(n)}} \Lambda^{(n)})^{-1})_{ij}$  represented by the infinite product (53) is convergent:

$$\lim_{n \rightarrow \infty} D_{\Lambda^{(n)}} \Lambda^{(n)} L \left( D_{\Lambda^{(n)}} \Lambda^{(n)} \right)^{-1} = \tilde{L}. \tag{55}$$

The resulting matrix  $\tilde{L}$  is not only unit lower triangular, but also block diagonal with the block sizes at most 2, because if the equality  $|\lambda_{p(k)} - s^{(\infty)}| = |\lambda_{p(k+1)} - s^{(\infty)}|$  in (40) holds, then both inequalities  $|\lambda_{p(k-1)} - s^{(\infty)}| > |\lambda_{p(k)} - s^{(\infty)}|$  and  $|\lambda_{p(k+1)} - s^{(\infty)}| > |\lambda_{p(k+2)} - s^{(\infty)}|$  are satisfied thanks to the fact that the eigenvalues are all distinct. Hence, the orthogonal matrix  $P^{(n)}$  given by the QR factorization of  $D_{\Lambda^{(n)}} \Lambda^{(n)} L (D_{\Lambda^{(n)}} \Lambda^{(n)})^{-1}$  is convergent:

$$\lim_{n \rightarrow \infty} P^{(n)} = \tilde{P}, \tag{56}$$

where  $\tilde{P}$  is a block diagonal matrix whose block size is at most 2. It then follows that

$$\lim_{n \rightarrow \infty} T^{(n)} = D_U \tilde{P}^T \Lambda \tilde{P} D_U^{-1} \tag{57}$$

from (52). Therefore,  $T^{(n)}$  converges to a block diagonal matrix whose block size is at most 2. □

### Appendix B: Proofs of Lemmas 2 and 3

We prove here Lemmas 2 and 3 in turn.

By the discussion in Appendix A,  $\lambda_{p(m-1)}$ ,  $\lambda_{p(m-2)}$  are the eigenvalues of the  $2 \times 2$  submatrix

$$\begin{pmatrix} \alpha_{m-2}^{(\infty)} & \beta_{m-2}^{(\infty)} \\ \beta_{m-2}^{(\infty)} & \alpha_{m-1}^{(\infty)} \end{pmatrix}. \tag{58}$$



In view of  $\beta_{m-2}^{(\infty)} \neq 0$ , we see  $|\lambda_{p(m-2)} - \lambda_{p(m)}| = |\lambda_{p(m-1)} - \lambda_{p(m)}|$ . The eigenvalues are real and distinct, which implies  $(\lambda_{p(m-2)} - \lambda_{p(m)}) + (\lambda_{p(m-1)} - \lambda_{p(m)}) = 0$ . Since the sum of the eigenvalues of the matrix (58) is equal to the trace of that, we have  $(\alpha_{m-2}^{(\infty)} - s^{(\infty)}) + (\alpha_{m-1}^{(\infty)} - s^{(\infty)}) = 0$ . In other words,

$$\lim_{n \rightarrow \infty} \begin{pmatrix} \alpha_{m-2}^{(n)} - s^{(n)} & \beta_{m-2}^{(n)} \\ \beta_{m-2}^{(n)} & \alpha_{m-1}^{(n)} - s^{(n)} \end{pmatrix} = \begin{pmatrix} -D & C \\ C & D \end{pmatrix} \tag{59}$$

holds for constants  $C$  and  $D$ . Obviously, the eigenvalues of the matrix (59) are  $\pm\sqrt{C^2 + D^2}$ . It then follows that

$$\sqrt{C^2 + D^2} = |\lambda_{p(m-2)} - \lambda_{p(m)}| = |\lambda_{p(m-1)} - \lambda_{p(m)}|. \tag{60}$$

This completes the proof of Lemma 2.

Next, we prove Lemma 3 based on Lemma 1. Actually, in Lemma 1,

$$\lim_{n \rightarrow \infty} |\gamma^{(n)}| = \prod_{1 \leq i \leq m-2} |\lambda_{p(i)} - \lambda_{p(m)}| > 0 \tag{61}$$

holds. We prove (61) below. If  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} = 0$ , then  $\lim_{n \rightarrow \infty} \alpha_{m-1}^{(n)} = \lambda_{p(m-1)}$  holds. Therefore, we obtain

$$\lim_{n \rightarrow \infty} |\gamma^{(n)}| = \lim_{n \rightarrow \infty} |d_{m-2}^{(n)}| = \lim_{n \rightarrow \infty} \prod_{1 \leq i \leq m-2} |\lambda_{p(i)} - \lambda_{p(m)}|$$

from (17) and (19). Next, we consider the situation  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} \neq 0$ . It is easy to see that, if  $\lim_{n \rightarrow \infty} \beta_{m-2}^{(n)} \neq 0$ , then  $\lim_{n \rightarrow \infty} \beta_{m-3}^{(n)} = 0$  because the block size of  $T^{(\infty)}$  is at most 2. Combining it with (17), we have

$$\lim_{n \rightarrow \infty} d_{m-3}^{(n)} = \prod_{1 \leq i \leq m-3} (\lambda_{p(i)} - \lambda_{p(m)}) \neq 0. \tag{62}$$

In addition, we see

$$\lim_{n \rightarrow \infty} (\gamma^{(n)})^2 = \lim_{n \rightarrow \infty} \left[ (d_{m-2}^{(n)})^2 + (\beta_{m-2}^{(n)} d_{m-3}^{(n)})^2 \right] \tag{63}$$

in (19). Noting  $\beta_{m-3}^{(\infty)} = 0$  and (17), we see that

$$\begin{aligned} \lim_{n \rightarrow \infty} |\gamma^{(n)}| &= \lim_{n \rightarrow \infty} \sqrt{(\alpha_{m-2}^{(n)} - s^{(n)})^2 + (\beta_{m-2}^{(n)})^2} |d_{m-3}^{(n)}| \\ &= \sqrt{(\alpha_{m-2}^{(\infty)} - \lambda_{p(m)})^2 + (\beta_{m-2}^{(\infty)})^2} \prod_{1 \leq i \leq m-3} |\lambda_{p(i)} - \lambda_{p(m)}| \end{aligned} \tag{64}$$

holds. Therefore, we obtain (61) from (59) and (60).

Obviously,

$$\lim_{n \rightarrow \infty} d_{m-1}^{(n)} = \lim_{n \rightarrow \infty} \prod_{1 \leq i \leq m-1} (\lambda_{p(i)} - s^{(n)}) \neq 0 \tag{65}$$

holds. Moreover, if  $D \neq 0$ , then

$$|d_m^{(n)}| \sim \frac{|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^2 |d_{m-3}^{(n)}|}{|D|} \tag{66}$$

in the same way as (22).

From (18), (61), (65), and (66), we have

$$\lim_{n \rightarrow \infty} \frac{|\beta_{m-1}^{(n+1)}|}{|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3} = \frac{|d_{m-3}^{(\infty)}|}{|D| |\lambda_{p(m-1)} - \lambda_{p(m)}|^2 \prod_{1 \leq i \leq m-2} |\lambda_{p(i)} - \lambda_{p(m)}|}. \tag{67}$$

Noting  $|\rho_1|$  is defined as (25), we have

$$\lim_{n \rightarrow \infty} \frac{|\beta_{m-1}^{(n+1)}|}{|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3} = \frac{|d_{m-3}^{(\infty)}|}{|\rho_1| |\lambda_{p(m-1)} - \lambda_{p(m)}|^3 \prod_{1 \leq i \leq m-2} |\lambda_{p(i)} - \lambda_{p(m)}|} \tag{68}$$

from (67).

We investigate the behavior of  $|d_{m-3}^{(\infty)}|$ . First of all, we assume  $\beta_{m-3}^{(\infty)} \neq 0$ . Then  $\beta_{m-4}^{(\infty)} = 0$  and  $\beta_{m-2}^{(\infty)} = 0$ . Similarly to (59) and (60),

$$|\alpha_{m-3}^{(\infty)} - \lambda_{p(m)}|^2 + |\beta_{m-3}^{(\infty)}|^2 = |\lambda_{p(m-3)} - \lambda_{p(m)}|^2 = |\lambda_{p(m-2)} - \lambda_{p(m)}|^2 \tag{69}$$

holds. Noting  $\beta_{m-4}^{(\infty)} = 0$ , we have

$$\begin{aligned} |d_{m-3}^{(\infty)}| &= |\alpha_{m-3}^{(\infty)} - \lambda_{p(m)}| \prod_{1 \leq i \leq m-4} |\lambda_{p(i)} - \lambda_{p(m)}| \\ &= \sqrt{|\lambda_{p(m-3)} - \lambda_{p(m)}|^2 - |\beta_{m-3}^{(\infty)}|^2} \prod_{1 \leq i \leq m-4} |\lambda_{p(i)} - \lambda_{p(m)}| \end{aligned}$$

in the same way as (59) and (60). Moreover, noting  $|\lambda_{p(m-3)} - \lambda_{p(m)}| = |\lambda_{p(m-2)} - \lambda_{p(m)}|$ , we obtain

$$|d_{m-3}^{(\infty)}| = \sqrt{1 - |\beta_{m-3}^{(\infty)}|^2 / |\lambda_{p(m-2)} - \lambda_{p(m)}|^2} \prod_{1 \leq i \leq m-3} |\lambda_{p(i)} - \lambda_{p(m)}|. \tag{70}$$

Actually, it is easy to see that (70) covers the case  $\beta_{m-3}^{(\infty)} = 0$ . Therefore, we obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{|\beta_{m-1}^{(n+1)}|}{|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3} &= \frac{\sqrt{1 - |\beta_{m-3}^{(\infty)}|^2} / |\lambda_{p(m-2)} - \lambda_{p(m)}|^2}{|\rho_1| |\lambda_{p(m-1)} - \lambda_{p(m)}|^3 |\lambda_{p(m-2)} - \lambda_{p(m)}|} \\ &= \frac{|\rho_2|}{|\rho_1| |\lambda_{p(m-1)} - \lambda_{p(m)}|^3 |\lambda_{p(m-2)} - \lambda_{p(m)}|}, \end{aligned} \tag{71}$$

where the first equality is due to (68) and (70), the second equality is due to the definition of  $|\rho_2|$  in (26). This completes the proof of (24).

The final task is to derive (28) in the case  $\alpha_{m-1}^{(\infty)} = \alpha_m^{(\infty)}$ . In (21),  $|\beta_{m-1}^{(n)}|/|\alpha_{m-1}^{(n)} - s^{(n)}| \leq 1$  in view of the definition of Wilkinson’s shift. Hence,

$$\limsup_{n \rightarrow \infty} \frac{|d_m^{(n)}|}{|\beta_{m-1}^{(n)}|} \leq |\beta_{m-2}^{(\infty)}|^2 |d_{m-3}^{(\infty)}|$$

from (21). Also noting  $|\beta_{m-2}^{(\infty)}| = |\lambda_{p(m-1)} - \lambda_{p(m)}| = |\lambda_{p(m-2)} - \lambda_{p(m)}|, \beta_{m-3}^{(\infty)} = 0$ , we have

$$\limsup_{n \rightarrow \infty} \frac{|d_m^{(n)}|}{|\beta_{m-1}^{(n)}|} \leq \prod_{1 \leq i \leq m-1} |\lambda_{p(i)} - \lambda_{p(m)}|.$$

Therefore, we obtain (28) from (18), (61), and (65).

### Appendix C: Convergence analysis based on perturbation theory

In this section, using perturbation theory we prove Corollary 1. In other words, we prove the following facts:

- if  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = 0$ , then  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-1}^{(n)}|^2)$ ;
- if  $\lim_{n \rightarrow \infty} (\alpha_{m-1}^{(n)} - \alpha_m^{(n)}) = D \neq 0$ , then  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3)$ .

In what follows, we use the so-called gap theorem [7, Theorem 11.7.1].

**Lemma 4** ([7]) *Let  $y$  be a unit vector,  $A$  be a symmetric matrix, and  $\lambda_i$  be the eigenvalue of  $A$  closest to  $y^T Ay$ . Then*

$$|y^T Ay - \lambda_i| \leq \frac{\|Ay - (y^T Ay) y\|^2}{\min_{i \neq l} |\lambda_i - y^T Ay|}$$

holds.



Note that  $s^{(\infty)} = \lambda_{p(m)}$ . Let

$$\delta = \min_{i \neq m} \left| \lambda_{p(i)} - s^{(\infty)} \right| > 0. \tag{77}$$

For any  $\epsilon_2 > 0$ , noting Lemma 4 with  $y := (0, 0, \dots, 1)^T$ , we have

$$\left| \lambda_{p(m)} - s^{(n)} \right| \leq \frac{|z^{(n)}|^2}{\delta - \epsilon_2} \tag{78}$$

for all sufficiently large  $n$ . In addition, it is easy to see that, for any  $\epsilon_3 > 0$ ,

$$\left| \lambda_{p(m-1)} - s^{(n)} \right| \geq \delta - \epsilon_3 \tag{79}$$

for all sufficiently large  $n$ .

Now we consider the case of  $|\alpha_{m-1}^{(n)} - \alpha_m^{(n)}| \rightarrow D \neq 0$ . We obtain

$$\begin{aligned} \left| \beta_{m-1}^{(n+1)} \right| &\leq \frac{\left| \beta_{m-1}^{(n)} \right| |z^{(n)}|^2 |1 + \epsilon_1|}{|\delta - \epsilon_2| |\delta - \epsilon_3|} \\ &= \frac{\left| \beta_{m-1}^{(n)} \right| \left| \beta_{m-2}^{(n)} \sin(\theta^{(n)}) \right|^2 |1 + \epsilon_1|}{|\delta - \epsilon_2| |\delta - \epsilon_3|} \\ &= \frac{\left| \beta_{m-1}^{(n)} \right|^3 \left| \beta_{m-2}^{(n)} \right|^2 |1 + \epsilon_1|}{\left| \alpha_{m-1}^{(n)} - \alpha_m^{(n)} \right|^2 |\delta - \epsilon_2| |\delta - \epsilon_3|} \end{aligned} \tag{80}$$

for all sufficiently large  $n$  by using (74), (79), (78), (76), (75) in turn. We see  $\epsilon_1, \epsilon_2, \epsilon_3 \rightarrow 0$  and  $|\alpha_{m-1}^{(n)} - \alpha_m^{(n)}| \rightarrow |D| > 0$  as  $n \rightarrow \infty$ . Therefore,  $|\beta_{m-1}^{(n+1)}| = O(|\beta_{m-2}^{(n)}|^2 |\beta_{m-1}^{(n)}|^3)$ .

Finally, we prove the quadratic convergence in the case  $|\alpha_{m-1}^{(n)} - \alpha_m^{(n)}| \rightarrow 0$ . Since the estimate (80) by the Jacobi transformation cannot derive the quadratic convergence in the case  $|\alpha_{m-1}^{(n)} - \alpha_m^{(n)}| \rightarrow 0$ , we give another estimate based on Lemma 4. For any  $\epsilon_4 > 0$ , we see

$$\left| \lambda_{p(m)} - \alpha_m^{(n)} \right| \leq \frac{\left| \beta_{m-1}^{(n)} \right|^2}{\delta - \epsilon_4} \tag{81}$$

for all sufficiently large  $n$  from Lemma 4 with  $y := (0, 0, \dots, 1)^T$ ,  $A := T^{(n)}$ . From the definition of Wilkinson’s shift,  $|\alpha_m^{(n)} - s^{(n)}| \leq |\beta_{m-1}^{(n)}|$  holds. Hence,

$$\left| \lambda_{p(m)} - s^{(n)} \right| \leq \left| \lambda_{p(m)} - \alpha_m^{(n)} \right| + \left| \alpha_m^{(n)} - s^{(n)} \right| \leq \frac{\left| \beta_{m-1}^{(n)} \right|^2}{\delta - \epsilon_4} + \left| \beta_{m-1}^{(n)} \right| \tag{82}$$

for all sufficiently large  $n$ . Thus, we have

$$\left| \beta_{m-1}^{(n+1)} \right| \leq \frac{\left| \beta_{m-1}^{(n)} \right|^2 \left( \delta - \epsilon_4 + \beta_{m-1}^{(n)} \right) |1 + \epsilon_1|}{|\delta - \epsilon_3| |\delta - \epsilon_4|}$$

for all sufficiently large  $n$  from (74), (79), (82). We see  $\epsilon_3, \epsilon_4 \rightarrow 0$  and  $\beta_{m-1}^{(n)} \rightarrow 0$  as  $n \rightarrow \infty$ . Therefore,  $\left| \beta_{m-1}^{(n+1)} \right| = O\left(\left| \beta_{m-1}^{(n)} \right|^2\right)$ . This completes the proof.

Although the convergence analysis above is readily accessible to the readers in the research fields of the numerical linear algebra, we also note that the right-hand side of (80) in our proof is an overestimate in the case of  $|\lambda_{p(m-1)} - \lambda_{p(m)}| < |\lambda_{p(m-2)} - \lambda_{p(m)}|$  because we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{\left| \alpha_{m-1}^{(n)} - \alpha_m^{(n)} \right|^2 |\delta|^2} &= \frac{1}{|D|^2 \left| \lambda_{p(m-1)} - \lambda_{p(m)} \right|^2} \\ &> \frac{1}{|D| \left| \lambda_{p(m-2)} - \lambda_{p(m)} \right| \left| \lambda_{p(m-1)} - \lambda_{p(m)} \right|^2} \\ &= \frac{1}{\left| \lambda_{p(m-1)} - \lambda_{p(m)} \right|^3 \left| \lambda_{p(m-2)} - \lambda_{p(m)} \right| |\rho_1|} \\ &\geq \frac{|\rho_2|}{\left| \lambda_{p(m-1)} - \lambda_{p(m)} \right|^3 \left| \lambda_{p(m-2)} - \lambda_{p(m)} \right| |\rho_1|} \\ &= \lim_{n \rightarrow \infty} \frac{\left| \beta_{m-1}^{(n+1)} \right|}{\left| \beta_{m-2}^{(n)} \right|^2 \left| \beta_{m-1}^{(n)} \right|^3}, \end{aligned} \tag{83}$$

where the first equality is due to the assumption  $\alpha_{m-1}^{(\infty)} - \alpha_m^{(\infty)} = D$  and the definition of  $\delta$  in (77), the next inequality is due to (27) and the above condition  $|\lambda_{p(m-1)} - \lambda_{p(m)}| < |\lambda_{p(m-2)} - \lambda_{p(m)}|$ , the next equality is due to (25), the next inequality is due to (26), and the last equality is due to (24). Also note that, in the case of  $\beta_{m-2}^{(\infty)} = C \neq 0$ , the above relation (83) holds because we have the strict inequality in (83) from (60). Therefore, the right-hand side of (80) is an overestimate.

## References

1. Hoffmann, W., Parlett, B.N.: A new proof of global convergence for the tridiagonal QL algorithm. *SIAM J. Numer. Anal.* **15**, 929–937 (1978)
2. Huang, H., Tam, T.: On the QR iterations of real matrices. *Linear Algebra Appl.* **408**, 161–176 (2005)
3. Jiang, E., Zhang, Z.: A new shift of the QL algorithm for irreducible symmetric matrices. *Linear Algebra Appl.* **65**, 261–272 (1985)
4. Leitte, R.S., Saldanha, N.C., Tomei, C.: The asymptotics of Wilkinson’s shift: loss of cubic convergence. *Found. Comput. Math.* **10**, 15–36 (2010)
5. Parlett, B.N.: Canonical decomposition of Hessenberg matrices. *Math. Comput.* **21**, 223–227 (1967)
6. Parlett, B.N.: Global convergence of the basic QR algorithm on Hessenberg matrices. *Math. Comput.* **22**, 803–817 (1968)

7. Parlett, B.N.: The symmetric eigenvalue problem. Prentice-Hall, Englewood Cliffs, New Jersey, 1980; SIAM, Philadelphia (1998)
8. Wang, T.: Convergence of the tridiagonal QR algorithm. *Linear Algebra Appl.* **322**, 1–17 (2001)
9. Wilkinson, J.H.: The algebraic eigenvalue problem. Clarendon Press, Oxford (1965)
10. Wilkinson, J.H.: Global convergence of tridiagonal QR algorithm with origin shifts. *Linear Algebra Appl.* **1**, 409–420 (1968)
11. Zhang, G.: On the convergence rate of the QL algorithm with Wilkinson's shift. *Linear Algebra Appl.* **113**, 131–137 (1989)