IMA Journal of Numerical Analysis (2009) Page 1 of 13
doi:10.1093/imanum/drnxxx

From qd to LR, or, How were the qd and LR algorithms discovered?

MARTIN H. GUTKNECHT†,
Seminar for Applied Mathematics, ETH Zurich, CH-8092 Zurich, Switzerland

AND

BERESFORD N. PARLETT‡
Department of Mathematics, University of California, Evans Hall, Berkeley, CA 94720-3840, USA

[Received on 17 August 2009; revised on ...]

Perhaps the most astonishing idea in eigenvalue computation is Rutishauser's idea of applying the LR transform to a matrix for generating a sequence of similar matrices that become more and more triangular. The same idea is the foundation of the ubiquitous QR algorithm. It is well known that this idea originated in Rutishauser's qd algorithm, which precedes the LR algorithm and can be understood as applying LR to a tridiagonal matrix. But how did Rutishauser discover qd, and when did he find the qd–LR connection? We checked some of the early sources and come up with an explanation.

Keywords: matrix eigenvalues, qd algorithm, LR algorithm, QR algorithm, Rutishauser.

1. Introduction

In the year 2000 the QR algorithm was placed on the list of the top ten algorithms of the 20th century in the journal "Computers in Science and Engineering" (see Parlett, 2000). The honor was well merited; the QR algorithm is the ubiquitous tool for computing eigenvalues of dense matrices. Its predecessor, the LR algorithm, is now largely forgotten and rarely taught to students. What we wish to say here is that, from an intellectual viewpoint, it was the LR algorithm that made the seminal contribution; QR was merely a stable version of LR. But how could anyone come up with the apparently ridiculous idea of factoring a square matrix $\mathbf{A} = \mathbf{A}_1$ into two triangular matrices, $\mathbf{A}_1 = \mathbf{L}_1\mathbf{R}_1$, and then forming a new matrix $\mathbf{A}_2 = \mathbf{R}_1\mathbf{L}_1$. This LR transform requires a lot of arithmetic and creates no zero entries in the matrix. But since $\mathbf{L}_1$, by convention, is lower triangular with ones on the diagonal, it is invertible and $\mathbf{A}_2$ has the same spectrum as $\mathbf{A}_1$. Yet it is not at all obvious that, if one is rich enough and keeps on computing $\mathbf{A}_3, \mathbf{A}_4, \mathbf{A}_5, \ldots$ by iterating the LR transform, then slowly but (nearly) surely the iterates become upper triangular and the diagonal entries converge to eigenvalues. Who would have thought of such a bizarre process?

It was Heinz Rutishauser who discovered this LR algorithm as a byproduct of his qd algorithm. In "The Algebraic Eigenvalue Problem" the eminent numerical analyst J.H. Wilkinson called Rutishauser an "algorithmic genius" (see Wilkinson, 1965, page vii), and regarding the LR algorithm he wrote in the same book (see Wilkinson, 1965, page 485): "In my opinion its development is the most significant advance which has been made in connexion with the eigenvalue problem since the advent of automatic

†Corresponding author. Email: mhg@math.ethz.ch
‡Email: parlett@math.berkeley.edu

computers." Surely the invention of these two algorithms is evidence of genius. That still leaves open our question: How did he do it? What follows is our guess of how it happened.

Our main conclusion is that he found the qd algorithm by studying previous work of Hadamard (1892), Aitken (1926, 1931), and Lanczos (1950a, Ch. VI) , and by improving on it. The insight that was truly impressive was to see that a step of the progressive qd algorithm (see below) can be interpreted as the LR transform on a tridiagonal matrix. After that it is not hard to see that tridiagonality of the matrix is not essential and so the LR transform may be applied to any matrix that permits triangular factorization.

2. Stiefel's "assignment" for Rutishauser

When founding the Institute of Applied Mathematics at ETH Zurich in 1948, Eduard Stiefel hired Heinz Rutishauser, who had just finished his dissertation in complex analysis, as a research assistant. Rutishauser was hired to help to construct a digital electronic computer and to explore and develop numerical methods for using it. In 1952, he finished his *Habilitation* thesis, in which he developed a compiler, and became a *Privatdozent*. After that, around 1953, on Stiefel's suggestion, Rutishauser approached the key problem of determining the poles of a rational function given by a power series in z^{-1},

$$f(z) := \sum_{\nu=0}^{\infty} \frac{s_\nu}{z^{\nu+1}} . \tag{2.1}$$

The application he had in mind was the following: Assume $\mathbf{A}$ is an $N \times N$ matrix and $\mathbf{x}_0$, $\mathbf{y}_0$ are two N-vectors. Then, for $s_\nu := \mathbf{y}_0^\mathsf{T} \mathbf{A}^\nu \mathbf{x}_0$, the series in (2.1) is the Taylor expansion at ∞ of

$$f(z) := \left\langle \mathbf{y}_0, (z\mathbf{I} - \mathbf{A})^{-1} \mathbf{x}_0 \right\rangle = \left\langle \mathbf{y}_0, \tfrac{1}{z} (\mathbf{I} - \tfrac{1}{z}\mathbf{A})^{-1} \mathbf{x}_0 \right\rangle , \tag{2.2}$$

which is a proper rational fraction of degree $n \leqslant N$ whose poles are eigenvalues of $\mathbf{A}$. This is seen from the representation

$$f(z) = \frac{\mathbf{y}_0^\mathsf{T} \,\mathsf{adj}\,(z\mathbf{I} - \mathbf{A})\, \mathbf{x}_0}{\mathsf{det}(z\mathbf{I} - \mathbf{A})}, \tag{2.3}$$

which also reveals that only the numerator depends on $\mathbf{x}_0$ and $\mathbf{y}_0$ unless some zeros and poles cancel. This application to the matrix eigenvalue problem was the starting point and the target of Rutishauser's investigation. He called the coefficients s_ν *Schwarz constants*, but today they are referred to as *moments* in numerical linear algebra and as *Markov parameters* in systems and control theory, where the sequence of moments is the *impulse response* of the linear time-invariant discrete-time single-input single-output (SISO) control system given by the state matrix $\mathbf{A}$ and the vector $\mathbf{x}_0$ and $\mathbf{y}_0$. So, Stiefel's proposal for Rutishauser was to determine the eigenvalues of $\mathbf{A}$ given the sequence of moments. Rutishauser (1957b) wrote in his foreword: "Following this suggestion the author developed an algorithm that solves the posed problem."

We next describe what was known before about this problem and how Rutishauser came up with his new solution and new insight. In view of the two dominant quantities involved he called his algorithm *quotient-difference algorithm* or, briefly, *QD algorithm*. Nowadays, the abbreviation in lower case letters, *qd algorithm*, is widely used to emphasize that q and d are not matrices (in contrast, say, to the LR and QR algorithms).

We know now that Stiefel's proposal was actually a bad one: the problem of determining the eigenvalues from the moments is typically extremely ill-conditioned. It is well-known that the elements of a symmetric tridiagonal matrix (and the nodes and weights of the corresponding Gauss-Christoffel quadrature formula) are badly determined by its moments (see, e.g., Gautschi, 1968, 1982). Rutishauser became aware of this ill-conditioning and of a better solution of the matrix eigenvalue problem (see Rutishauser, 1955a), namely using the Lanczos algorithm (see Lanczos, 1950a) for reducing the matrix to tridiagonal form and then applying the *progressive form* of his qd algorithm or, what amounts to the same, his *LR algorithm*. This approach may also run into stability problems or even break down, but this occurs only rarely, while ill-conditioning is nearly inescapable when using moments.

3. Finding the poles of f from the moments: Hadamard and Aitken

If f is a proper rational function q/p of degree n with explicitly known denominator

$$p(z) = \pi_0 z^n + \pi_1 z^{n-1} + \cdots + \pi_n ,$$

it follows from the expansion (2.1) that the moments satisfy the difference equation

$$\pi_0 s_{k+n} + \pi_1 s_{k+n-1} + \cdots + \pi_n s_k = 0 \qquad (k \geqslant 0). \tag{3.1}$$

The initial values $s_0, \ldots, s_{n-1}$ are determined by the numerator polynomial q, while the recursion (3.1) only depends on p. It was known to Daniel Bernoulli (1700-1782) that if p has a unique zero z_1 of maximum modulus (and hence, the series (2.1) converges for $|z| > |z_1|$), then the solution $\{s_\nu\}$ of the difference equation (3.1) satisfies

$$\lim_{\nu\to\infty} \frac{s_{\nu+1}}{s_\nu} = z_1 . \tag{3.2}$$

This is Bernoulli's method for finding such a greatest root (see Bernoulli, 1732; Aitken, 1926) .

König (1884) established more than 150 years later that the analogous result holds for any power series of an analytic function with a single simple pole on the boundary of the disk of convergence. Soon after that the French mathematician Jacques Hadamard (1865–1963), in his thesis (see Hadamard, 1892) published in 1892, solved the problem of finding *all* the poles of f from the moments by a beautiful procedure that is very ill-suited to computer implementation, however. Now the function just had to be meromorphic in a disk around the origin and analytic at the origin, where its Taylor series was given. Here we formulate the results assuming f is analytic at ∞ and given by the series (2.1). For simplicity, we further assume that f is a proper rational function of order n.

Hadamard considered the double sequence of *Hankel determinants*

$$H_k^{(\nu)} := \begin{vmatrix} s_\nu & s_{\nu+1} & \cdots & s_{\nu+k-1} \\ s_{\nu+1} & s_{\nu+2} & \cdots & s_{\nu+k} \\ \vdots & \vdots & \ddots & \vdots \\ s_{\nu+k-1} & s_{\nu+k} & \cdots & s_{\nu+2k-2} \end{vmatrix} \qquad (k = 1, 2, \ldots;\ \nu = 0, 1, \ldots) \tag{3.3}$$

and, adapted to our situation, established the following main result.

THEOREM 3.1 [Hadamard (1892)] Assume the series (2.1) represents a rational function whose n poles, counted including multiplicity, are ordered such that

$$|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_{n-1}| \geqslant |\lambda_n| . \tag{3.4}$$

If $1 \leqslant k < n$ and $|\lambda_{k+1}| < \Lambda < |\lambda_k|$, or if $k = n$ and $\Lambda < |\lambda_n|$, then

$$H_k^{(\nu)} = \mathsf{const} \cdot (\lambda_1 \cdots \lambda_k)^\nu \left[1 + \mathcal{O}\left(\frac{\Lambda}{|\lambda_k|} \right)^\nu \right] \qquad \text{as} \quad \nu \to \infty . \tag{3.5}$$

Assuming simple poles Henrici (1958) gave a simpler proof of this result. Multiple poles can be treated with a technique used by Golomb (1943). New proofs of Hadamard's theorem have also been a topic in the subsequent qd literature (see Gragg, 1973; Gragg & Householder, 1966; Housholder, 1974).

Here are some obvious conclusions.

COROLLARY 3.1 Under the assumptions of Theorem 3.1, $H_{n+1}^{(\nu)} = 0$ $(\forall \nu)$. Moreover, if f has n simple poles, then

1. $H_k^{(\nu)} \neq 0$ $(k = 1, \ldots, n)$ for large enough ν.
2. If $|\lambda_k| > |\lambda_{k+1}|$, then

$$\frac{H_k^{(\nu+1)}}{H_k^{(\nu)}} \to \lambda_1 \lambda_2 \cdots \lambda_k \qquad \text{as} \quad \nu \to \infty. \tag{3.6}$$

3. If $|\lambda_{k-1}| > |\lambda_k| > |\lambda_{k+1}|$, then

$$q_k^{(\nu)} := \frac{H_k^{(\nu+1)}}{H_k^{(\nu)}} \cdot \frac{H_{k-1}^{(\nu)}}{H_{k-1}^{(\nu+1)}} \to \lambda_k \qquad \text{as} \quad \nu \to \infty. \tag{3.7}$$

Statement (3.7) persists for $k = 1$ if we let $H_0^{(\nu)} := 1$ $(\forall \nu)$. In view of $H_1^{(\nu)} = s_\nu$, it reduces then to Bernoulli's result (3.2).

We observe that the LU decomposition of the Hankel matrix in (3.3), if it is possible without pivoting, yields for fixed ν the Hankel determinants $H_1^{(\nu)}, \ldots, H_N^{(\nu)}$ in one factorization. Nevertheless, computing this LU decomposition for a sufficient number of indices ν is an unacceptably big effort. Obviously, Hadamard was only interested in theorems, not in practical computation. The motivation for developing an efficient algorithm must have been missing, though, in fact, Hadamard had the key in his hands: the striking nonlinear relation

$$\left(H_k^{(\nu)}\right)^2 = H_k^{(\nu-1)} H_k^{(\nu+1)} - H_{k+1}^{(\nu-1)} H_{k-1}^{(\nu+1)} \tag{3.8}$$

among neighboring Hankel determinants. It is now often called *Jacobi's identity* for Hankel determinants Henrici (1974). Note that it expresses the square of $H_k^{(\nu)}$ as a difference of products of next neighbors.

It was the New Zealander Alexander Craig Aitken (1895–1967), in Scotland, who, in 1926 came up with the now obvious algorithmic conclusion (see Aitken, 1926, 1931). He was unaware of Hadamard's work, but rediscovered Theorem 3.1 and knew Jacobi's identity (3.8), which he considered as a special case of a "theorem of compound determinants". He realized that it can be used to build up — from the left or from the top — the triangular table

$$\begin{array}{cccccc}
1 & & & & & \\
1 & H_1^{(0)} & & & & \\
1 & H_1^{(1)} & H_2^{(0)} & & & \\
1 & H_1^{(2)} & H_2^{(1)} & H_3^{(0)} & & \\
1 & H_1^{(3)} & H_2^{(2)} & H_3^{(1)} & H_4^{(0)} & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array} \tag{3.9}$$

Unfortunately, when one of the determinants vanishes, both the horizontal and the vertical recursions break down. But similar types of breakdown can occur in many related recursive schemes.

Another constructive tool that was available when Rutishauser solved Stiefel's problem was the *Chebyshev algorithm* (see Chebyshev, 1859), that allows us to compute the recurrence coefficients of a set of orthogonal polynomials if the sequence of moments of the underlying weight function is known. Specifically, we need the first $2m$ moments to construct recursively the orthogonal polynomials up to degree m. Neither Stiefel nor Rutishauser seem to ever mention this tool, despite the fact that, in its initial phase, Rutishauser qd algorithm serves the same purpose, as we will see. The Chebyshev algorithm was later revived, analyzed, and modified by Sack & Donovan (1972), Wheeler (1974), and, in a series of papers by Gautschi (the first of these being Gautschi, 1970) , who also came up with the name *modified Chebyshev algorithm* for the more stable version using *modified moments*.

4. Rutishauser's qd algorithm

Rutishauser (1954b) was aware of the work of Hadamard (1892), Aitken (1926, 1931), and Lanczos (1950b) when he worked on Stiefel's problem. It seems that in the second half of 1952 or early in 1953 he took Aitken's work, improved it in a significant way, and made the connections to a number of related topics and applications. The key result was his *qd algorithm*, on which he published three papers (Rutishauser, 1954b,a,c) in 1954, the first and most fundamental of which was received by ZAMP on Aug. 5, 1953. The following year he had yet another seminal article (see Rutishauser, 1955a) on the application of qd to the eigenvalue problem. In partly revised form this early work on qd is collected in (Rutishauser, 1957b), which covers also some additional material, in particular a short appendix on the LR algorithm. Another appendix contains a shortened version of (Rutishauser, 1954c), which is also related to cg and qd. The volume was issued on January 15, 1958, but it seems to have been compiled long before. In fact, the preface is dated June 26, 1956. Moreover, Rutishauser (1957a) cites preprints dated 1956 and 1955, respectively, of his and Stiefel's contributions to the volume.

Rutishauser does nowhere clearly state how he found the qd algorithm; he only gives an indication, not a complete derivation. Henrici (1958) writes that the qd algorithm "by a simple but ingenious modification of Aitken's method, entirely bypasses the computation of Hankel determinants". And that "It it is remarkable that in the computation of the q_k^n, the determinants H_k^n do not have to be used if a set of auxiliary quantities is introduced."

The details have been hinted to by Henrici (1958) and worked out by Parlett (1996): first, in view of Hadamard's Theorem 3.1, in particular conclusion (3.7), the target of the computation are the quotients $q_k^{(\nu)}$. By multiplying Jacobi's identity (3.8) centered at $H_{k-1}^{(\nu+1)}$,

$$\left(H_{k-1}^{(\nu+1)}\right)^2 = H_{k-1}^{(\nu)} H_{k-1}^{(\nu+2)} - H_k^{(\nu)} H_{k-2}^{(\nu+2)} ,$$

with

$$\frac{H_k^{(\nu+1)}}{H_{k-1}^{(\nu+1)} H_k^{(\nu)} H_{k-1}^{(\nu+2)}}$$

we can turn the first term on the right-hand side into $q_k^{(\nu)}$:

$$\frac{H_{k-1}^{(\nu+1)} H_k^{(\nu+1)}}{H_k^{(\nu)} H_{k-1}^{(\nu+2)}} = q_k^{(\nu)} - \frac{H_{k-2}^{(\nu+2)} H_k^{(\nu+1)}}{H_{k-1}^{(\nu+1)} H_{k-1}^{(\nu+2)}} . \tag{4.1}$$

Likewise, we write down Jacobi's identity centered at $H_k^{(\nu+1)}$,

$$\left(H_k^{(\nu+1)}\right)^2 = H_k^{(\nu)} H_k^{(\nu+2)} - H_{k+1}^{(\nu)} H_{k-1}^{(\nu+2)},$$

and multiply it with

$$\frac{H_{k-1}^{(\nu+1)}}{H_k^{(\nu)} H_{k-1}^{(\nu+2)} H_k^{(\nu+1)}}.$$

This turns the first term on the right-hand side into $q_k^{(\nu+1)}$:

$$\frac{H_{k-1}^{(\nu+1)} H_k^{(\nu+1)}}{H_k^{(\nu)} H_{k-1}^{(\nu+2)}} = q_k^{(\nu+1)} - \frac{H_{k-1}^{(\nu+1)} H_{k+1}^{(\nu)}}{H_k^{(\nu)} H_k^{(\nu+1)}}. \tag{4.2}$$

Clearly the left-hand sides of (4.1) and (4.2) are identical and the second terms on the right-hand sides have the same structure. So, after introducing the auxiliary quantity

$$e_k^{(\nu)} := \frac{H_{k-1}^{(\nu+1)} H_{k+1}^{(\nu)}}{H_k^{(\nu)} H_k^{(\nu+1)}}, \tag{4.3}$$

we can conclude from (4.1) and (4.2) that

$$q_k^{(\nu)} + e_k^{(\nu)} = q_k^{(\nu+1)} + e_{k-1}^{(\nu+1)}. \tag{4.4}$$

This relation can be seen to also hold for $k = 1$ if we define $e_0^{(\nu)} := 0$ for all ν. Moreover, under our assumption of $f(z)$ being a proper rational fraction of degree n, one can show that $e_n^{(\nu)} = 0$ for all ν. In addition, from the definitions (3.7) of $q_k^{(\nu)}$ and (4.3) of $e_k^{(\nu)}$ it is readily verified that

$$q_{k+1}^{(\nu)} e_k^{(\nu)} = q_k^{(\nu+1)} e_k^{(\nu+1)}. \tag{4.5}$$

The relations (4.4) and (4.5) are the *rhombus rules* defining the qd algorithm.[1] Rutishauser (1954b) suggested to write down the quantities $e_k^{(\nu)}$ and $q_k^{(\nu)}$ in a triangular scheme called *qd scheme* (also known as *qd table*); see Fig. 1.[2] Recall that at this time, the simple computations needed to build up this scheme were normally done on a desk calculator, so a suitable scheme to write down the numbers obtained was most useful.

With the "initial values" $e_0^{(\nu)} = 0$ and $q_1^{(\nu)} = s_{\nu+1}/s_\nu$ the rhombus rules allow us to build up the qd scheme from the first column. Alternatively, they can be used to build up the scheme from its top diagonal. The latter application is called the *progressive qd algorithm*. It is the version that is still of importance. Rutishauser (1953b, 1958) knew that the top diagonal of the qd scheme can be computed with the Lanczos algorithm (see Lanczos, 1950b; Rutishauser, 1953a) or with the conjugate gradient method (see Hestenes & Stiefel, 1952).

[1] According to footnotes of Henrici (1958) and Rutishauser (1957b) the name "rhombus rules" was coined by Stiefel (see Stiefel, 1955, page 42) or (Stiefel, 1958, page 18).

[2] The first example of a qd scheme in (Rutishauser, 1954b) listed, for each ν, not only the two columns with $e_k^{(\nu)}$ and $q_k^{(\nu)}$, but additional two columns for $s_k^{(\nu)}$ and $d_k^{(\nu)} := q_k^{(\nu+1)} - q_k^{(\nu)}$. The rhombus rules made these two columns obsolete.

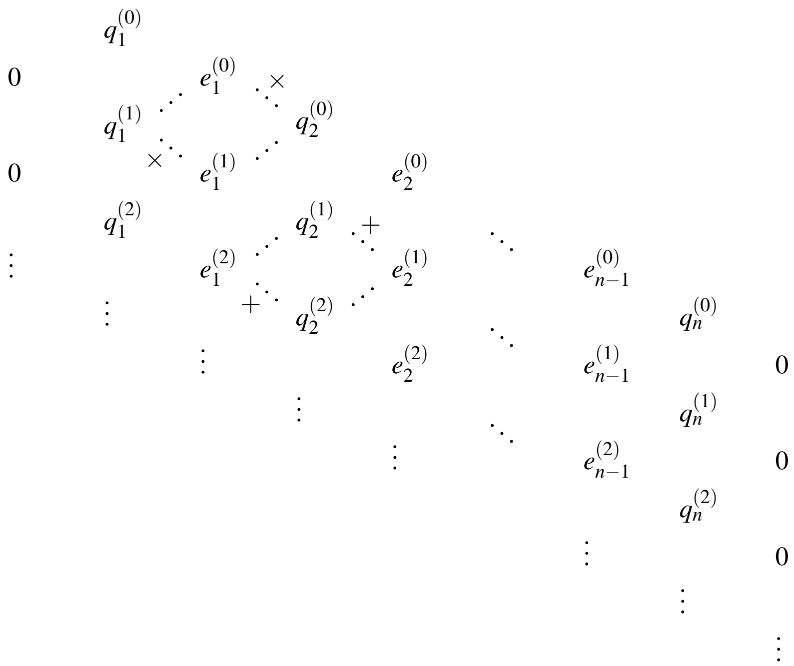

FIG. 1. The qd table of a rational function f of degree n and manifestations of the two rhombus rules.

In the case of a proper rational fraction f of exact degree n, as in (2.2), $e_n^{(\nu)} = 0$ holds for all ν, and thus the table is not defined beyond the nth e–column. Assuming that all the poles of f have different moduli, Rutishauser (1954b) could readily conclude from Aitken's work (see Aitken, 1926) that

$$\left.\begin{array}{lcl} \lim_{\nu\to\infty} q_k^{(\nu)} & = & \lambda_k \\ \lim_{\nu\to\infty} e_k^{(\nu)} & = & 0 \end{array}\right\} \qquad (k = 1, 2, \ldots, n). \tag{4.6}$$

Rutishauser (1954b) also gave the generalization of this to meromorphic functions, referring for its proof to Hadamard ("The proof of the proposition can be extracted offhand from the above mentioned work of Hadamard (1892, §§14–21)"). Among the two conclusions drawn from this generalization, but given without proof, there is one later referred to by Henrici (1974) as *Rutishauser's rule*, which covers the case of poles with equal modulus. With tools from functional analysis this rule was first fully proved by Stewart (1971). A direct but complicated proof requiring an additional assumption was given by Henrici (1974, pp. 642–650), who later in Part II of (Henrici, 1983) admitted that his earlier proof in (Henrici, 1958) is erroneous and promoted a new direct and simple proof due to his student Seewald (1982).

There is another, very elegant derivations of the rhombus rules and the qd algorithm, which is based on various ways to expand $f(z)$ into a continued fraction. This derivation is given in the original paper (Rutishauser, 1954b) and in (Rutishauser, 1957b), but according to footnotes in both papers it was suggested by Stiefel and was not the way the rules were discovered. (However, Henrici (1977, p. 527) guessed that the continued fraction based derivation was the original one.)

5. From qd to LR

In none of Rutishauser's early qd papers (see Rutishauser, 1954b,a,c, 1955a) is there any hint of the LR algorithm yet. Why should there be?

Consider the tridiagonal matrices

$$\mathbf{T}_\nu := \begin{pmatrix} q_1^{(\nu)} & 1 & & & \\ e_1^{(\nu)}q_1^{(\nu)} & e_1^{(\nu)}+q_2^{(\nu)} & 1 & & \\ & e_2^{(\nu)}q_2^{(\nu)} & e_2^{(\nu)}+q_3^{(\nu)} & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & e_{n-1}^{(\nu)}q_{n-1}^{(\nu)} & e_{n-1}^{(\nu)}+q_n^{(\nu)} \end{pmatrix}. \tag{5.1}$$

Note that $\mathbf{T}_\nu$ has the simple LU (in German: LR) decomposition

$$\mathbf{T}_\nu = \mathbf{L}_\nu \mathbf{R}_\nu \tag{5.2}$$

with

$$\mathbf{L}_\nu = \begin{pmatrix} 1 & & & & \\ e_1^{(\nu)} & 1 & & & \\ & e_2^{(\nu)} & \ddots & & \\ & & \ddots & \ddots & \\ & & & e_{n-1}^{(\nu)} & 1 \end{pmatrix}, \qquad \mathbf{R}_\nu = \begin{pmatrix} q_1^{(\nu)} & 1 & & & \\ & q_2^{(\nu)} & 1 & & \\ & & q_3^{(\nu)} & \ddots & \\ & & & \ddots & 1 \\ & & & & q_n^{(\nu)} \end{pmatrix}. \tag{5.3}$$

At some historic moment in 1954 or 1955, Rutishauser must have made the remarkable observation that his rhombus rules (4.4), (4.5) for the qd table could be interpreted as computing this triangular factorization $\mathbf{T}_\nu = \mathbf{L}_\nu \mathbf{R}_\nu$ and then forming a new tridiagonal matrix $\mathbf{T}_{\nu+1}$ by multiplying together the two factors in reverse order. In fact,

$$\mathbf{R}_\nu \mathbf{L}_\nu = \begin{pmatrix} e_1^{(\nu)}+q_1^{(\nu)} & 1 & & & \\ e_1^{(\nu)}q_2^{(\nu)} & e_2^{(\nu)}+q_2^{(\nu)} & 1 & & \\ & e_2^{(\nu)}q_3^{(\nu)} & e_3^{(\nu)}+q_3^{(\nu)} & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & e_{n-1}^{(\nu)}q_n^{(\nu)} & q_n^{(\nu)} \end{pmatrix}. \tag{5.4}$$

In view of $e_n^{(\nu)} = 0$ and the rhombus rules (4.4), (4.5) this product equals exactly $\mathbf{T}_{\nu+1}$ as defined by (5.1) with ν replaced by $\nu+1$. Thus,

$$\mathbf{R}_\nu \mathbf{L}_\nu = \mathbf{T}_{\nu+1} \qquad (\text{if} \quad e_n^{(\nu)} = 0). \tag{5.5}$$

So, the progressive qd algorithm can be viewed as performing the steps

$$\mathbf{T}_\nu = \mathbf{L}_\nu \mathbf{R}_\nu \quad \rightsquigarrow \quad \mathbf{R}_\nu \mathbf{L}_\nu = \mathbf{T}_{\nu+1}. \tag{5.6}$$

Such a step from $\mathbf{T}_\nu$ to $\mathbf{T}_{\nu+1}$ is called an *LR transformation*. Since

$$\mathbf{T}_{\nu+1} = \mathbf{L}_\nu^{-1}\mathbf{T}_\nu\mathbf{L}_\nu\,, \tag{5.7}$$

the tridiagonal matrices are similar. If all the poles have distinct moduli, we know from Hadamard's theorem or from the qd algorithm that $\mathbf{L}_\nu \to \mathbf{I}$ as $\nu \to \infty$, so in the long run $\mathbf{R}_\nu$ will contain the eigenvalues in its diagonal. If there are poles that are multiple or have the same modulus, we can still apply an adaptation of what Henrici called Rutishauser's rule.

That moment in 1954 or 1955 was the birth of the LR algorithm, first only in its form for tridiagonal matrices, which is equivalent to the qd algorithm for finite J–fractions and S–fractions. But it must have taken Rutishauser only a few minutes to see that the tridiagonal form is not necessary and that one could as well start from a full matrix. From this point of view, the LR algorithm is a natural generalization of the qd algorithm. However, for a full matrix the LR algorithm is costly. But it has the most welcome feature to conserve the bandwidth of a banded matrix, in particular of a tridiagonal one. So, Rutishauser ended up with a most elegant and intriguing algorithm, but its most important use was for the tridiagonal case, where it is identical with the progressive qd algorithm. And Rutishauser knew that, for nearly any pair of starting vectors, the reduction to tridiagonal matrices could be achieved with the Lanczos algorithm. The reduction of a symmetric matrix to tridiagonal form by orthogonal Givens rotations (see Givens, 1953) was introduced around the same time, but was no alternative for the general case.

Rutishauser realized from the beginning that fast convergence requires shifting the spectrum appropriately since he had analyzed convergence and introduced spectral shifts for the progressive qd algorithm before; see (Rutishauser, 1954a, § 7) or (Rutishauser, 1957b, Ch. 2, § 8) and (Rutishauser, 1955a, § 4) or (Rutishauser, 1957b, Ch. 3, § 4) . So, in practice, (5.6) is replaced by

$$\mathbf{T}_\nu - \delta_\nu\mathbf{I} = \mathbf{L}_\nu\mathbf{R}_\nu \quad \rightsquigarrow \quad \mathbf{R}_\nu\mathbf{L}_\nu + \delta_\nu\mathbf{I} = \mathbf{T}_{\nu+1}\,, \tag{5.8}$$

where δ_ν is the shift parameter.

Rutishauser's first publication (see Rutishauser, 1955b) on the LR algorithm is, in 1955, a two-page note in French in the *Comptes Rendus*, the premium journal for research announcement in mathematics of the time. In the following year, according to Rutishauser (1957b), he produced an mimeographed 51-page preprint in English, entitled "Report on the Solution of Eigenvalue Problems with the LR–transformation" (see Rutishauser, 1956)[3], but only two years later this article got properly published at the NBS (see Rutishauser, 1958), in the volume that also contained Henrici's review of the qd algorithm (see Henrici, 1958). Only then the qd and LR algorithms made their appearance for the English speaking world, except for the few people who had received the LR preprint before. In 1957, Rutishauser included a 5-page appendix on the LR transformation in his substantial report (see Rutishauser, 1957b), which compiled and updated most of his previous work on qd, but was still in German.

What remains a puzzle is the following. LR is much more general than qd. Despite this, Rutishauser attached much more importance to qd than to LR. For example, in 1968, ten years after the NBS volume, Rutishauser wrote an updated script (see Rutishauser, 1968) of some of his research on qd with no mention of LR.

Two comments are in order to put Rutishauser's work in perspective. Firstly, starting around 1955 his main preoccupation for more than a decade was the development of the Algol60 programming

[3] According to the annual report 1956 of the Institute for Applied Mathematics this report was sent in that year to people interested in the subject.

language (see Backus *et al.*, 1960, 1963), the description of its usage for basic problems of numerical analysis (see Rutishauser, 1967), and the creation of open source high-quality programming libraries (see Rutishauser, 1961; Wilkinson & Reinsch, 1971). At the same time he published a large number of papers on various topics in numerical analysis. Secondly, although the discovery of LR is considered to be one of the most significant moments in matrix computations, nevertheless LR was quickly eclipsed by Francis' QR algorithm (see Francis, 1961, 1962) shortly after its introduction. Francis' original papers on QR explicitly reveal the primal importance of LR for the development of QR, which was, correctly, seen as a backward stable variation of LR. Strangely enough, Rutishauser (1963) presented at the IFIP Congress 1962 a paper entitled *Numerical experiments with the QD–transformation of J.G.F. Francis*. We do not know whether the "QD" in the title was a misprint or an intention (recall that he usually wrote qd in capitals). In any case, one can hardly deny that Rutishauser somehow disliked the QR algorithm: there was no room for it in his lectures (see Rutishauser, 1976, 1990).

References

AITKEN, A. C. (1926) On Bernoulli's numerical solution of algebraic equations. *Proc. Roy. Soc. Edinb.*, **46**, 289–305.

AITKEN, A. C. (1931) Further numerical studies in algebraic equations and matrices. *Proc. Roy. Soc. Edinb.*, **51**, 80–90.

BACKUS, J. W., BAUER, F. L., GREEN, J., KATZ, C., MCCARTHY, J., (ED.), P. N., PERLIS, A. J., RUTISHAUSER, H., SAMELSON, K., VAUQUOIS, B., WEGSTEIN, J. H., VAN WIJNGAARDEN, A. & WOODGER, M. (1960) Report on the algorithmic language ALGOL 60. *Numer. Math.*, **12**, 106–136.

BACKUS, J. W., BAUER, F. L., GREEN, J., KATZ, C., MCCARTHY, J., (ED.), P. N., PERLIS, A. J., RUTISHAUSER, H., SAMELSON, K., VAUQUOIS, B., WEGSTEIN, J. H., VAN WIJNGAARDEN, A. & WOODGER, M. (1963) Revised report on the algorithmic language ALGOL 60. *Numer. Math.*, **4**, 420–453.

BERNOULLI, D. (1732). *Commentarii Acad. Sc. Petropol.*, **III**.

CHEBYSHEV, P. (1859) Sur l'interpolation par la méthode des moindres carrés. *Mém. Acad. Impér. Sciences St. Pétersbourg, série VII*, **1**, 1–24. [*Oeuvres*, vol. 1, pp. 471–498].

FRANCIS, J. G. F. (1961) The QR transformation—a unitary analogue to the LR transformation, Part I. *Comput. J.*, **4**, 265–271.

FRANCIS, J. G. F. (1962) The QR transformation—a unitary analogue to the LR transformation, Part II. *Comput. J.*, **4**, 332–345.

GAUTSCHI, W. (1968) Construction of Gauss–Christoffel quadrature formulas. *Math. Comp.*, **22**, 251–270.

GAUTSCHI, W. (1970) On the construction of Gaussian quadrature rules from modified moments. *Math. Comp.*, **24**, 245–260.

GAUTSCHI, W. (1982) On generating orthogonal polynomials. *SIAM J. Sci. Stat. Comput.*, **3**, 289–317.

GIVENS, W. (1953) A method of computing eigenvalues and eigenvectors suggested by classical results on symmetric matrices. *Simultaneous linear equations and the determination of eigenvalues*. National Bureau of Standards Applied Mathematics Series, No. 29. Washington, D. C.: U. S. Government Printing Office, pp. 117–122.

GOLOMB, M. (1943) Zeros and poles of functions defined by Taylor series. *Bull. Amer. Math. Soc.*, **49**, 581–592.

GRAGG, W. B. (1973) On Hadamard's theory of polar singularities. *Padé Approximants and Their Applications* (P. Graves-Morris ed.). London: Academic Press, pp. 117–123.

GRAGG, W. B. & HOUSEHOLDER, A. S. (1966) On a theorem of Koenig. *Numer. Math.*, **8**, 465–468.

HADAMARD, J. (1892) Essai sur l'étude des fonctions donnée par leurs développement de Taylor. *Ph.D. thesis*, Université de Paris, Paris. Also in J. de Math. Pures Appl., vol. 8, no. 4, pages 101-186 (1992).

HENRICI, P. (1958) The quotient-difference algorithm. *Further Contributions to the Solution of Simultaneous Linear Equations and the Determination of Eigenvalues*. Applied Mathematics Series, vol. 49. National Bureau of Standards, pp. 23–46.

HENRICI, P. (1974) *Applied and Computational Complex Analysis, Vol. 1*. New York: Wiley.

HENRICI, P. (1977) *Applied and Computational Complex Analysis, Vol. 2*. New York: Wiley.

HENRICI, P. (1983) Topics in computational complex analysis. I: Methods of descent for polynomial equations. II: New developments on the qd algorithm. III: The asymptotic behavior of the best approximation to an analytic function in the unit disk. IV: The Lagrange-Bürmann formula for systems of formal power series. *Computational Complex Analysis* (H. Werner, L. Wuytack, E. Ng & H. J. Bünger eds). Dordrecht: D. Reidel Publishing Co.

HESTENES, M. R. & STIEFEL, E. (1952) Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bureau Standards*, **49**, 409–435.

HOUSHOLDER, A. S. (1974) The Koenig-Hadamard theorem again. *Studies in Numerical Analysis* (B. K. P. Scaife ed.). London: Academic Press, pp. 235–240.

KÖNIG, J. (1884) Über eine Eigenschaft der Potenzreihen. *Math. Ann.*, **23**, 447–449.

LANCZOS, C. (1950a) An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bureau Standards*, **45**, 255–281.

LANCZOS, C. (1950b) An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bureau Standards*, **45**, 255–281.

PARLETT, B. N. (1996) What Hadamard missed. *Technical Report* PAM-671. Center for Pure and Applied Mathematics, University of California at Berkeley.

PARLETT, B. N. (2000) The QR algorithm. *Computing in Science and Engineering*, 38–42.

RUTISHAUSER, H. (1953a) Beiträge zur Kenntnis des Biorthogonalisierungs-Algorithmus von Lanczos. *Z. Angew. Math. Phys.*, **4**, 35–56.

RUTISHAUSER, H. (1953b) *Bemerkungen zum programmgesteuerten Rechnen*. Max-Planck-Institut für Physik, Göttingen, pp. 34–37. Vorträge über Rechenanlagen.

RUTISHAUSER, H. (1954a) Anwendungen des Quotienten–Differenzen–Algorithmus. *Z. Angew. Math. Phys.*, **5**, 496–508.

RUTISHAUSER, H. (1954b) Der Quotienten–Differenzen–Algorithmus. *Z. Angew. Math. Phys.*, **5**, 233–251.

RUTISHAUSER, H. (1954c) Ein infinitesimales Analogon zum Quotienten–Differenzen–Algorithmus. *Arch.Math.*, **5**, 132–137.

RUTISHAUSER, H. (1955a) Bestimmung der Eigenwerte und Eigenvektoren einer Matrix mit Hilfe des Quotienten–Differenzen–Algorithmus. *Z. Angew. Math. Phys.*, **6**, 387–401.

RUTISHAUSER, H. (1955b) Une méthode pour la détermination des valeurs propres d'une matrice. *C.R. Acad. Sci. Paris*, **240**, 34–36.

RUTISHAUSER, H. (1956) Report on the solution of eigenvalue problems with the LR–transformation. *Mimeographed report*. Inst. f. angew. Math. ETH Zürich.

RUTISHAUSER, H. (1957a) *Der Quotienten-Differenzen-Algorithmus*. Mitt. Inst. angew. Math. ETH, Nr. 7. Basel: Birkhäuser.

RUTISHAUSER, H. (1957b) Der Quotienten–Differenzen–Algorithmus. *Mitt. Inst. f. angew. Math. ETH Zürich*, **Nr. 7**. (Contains revised versions of Rutishauser (1954b), Rutishauser (1954a), Rutishauser (1955a).).

RUTISHAUSER, H. (1958) Solution of eigenvalue problems with the LR–transformation. *Further Contributions to the Solution of Simultaneous Linear Equations and the Determination of Eigenvalues*. Applied Mathematics Series, vol. 49. National Bureau of Standards, pp. 47–81.

RUTISHAUSER, H. (1961) ALGOL–Manual der ALCOR–Gruppe. *Elektronische Rechenanlagen*, **5**. Heft 6/1961, Heft 2/1962.

RUTISHAUSER, H. (1963) Numerical experiments with the QD–transformation of J.G.F. Francis. *Information Processing 1962 (Proc. IFIP Congress, Munich, 1962)*. Amsterdam: North-Holland Publ. Co., pp. 202–203.

RUTISHAUSER, H. (1967) *Description of ALGOL 60 (Handbook for Automatic Computation Vol. Ia)*. Berlin/Heidelberg/New York: Springer-Verlag.

RUTISHAUSER, H. (1968) Les propriétés numériques de l'algorithme quotient-différence. *Rapport EUR* 4083f. Communauté Européenne de l'Energie Atomique – Euratom. Rapport des conférences données par le Prof. H. Rutishauser; rapporteur: R. - F. Gloden.

RUTISHAUSER, H. (1976) *Vorlesungen über numerische Mathematik (2 Bände)*. Basel/Stuttgart: Birkhäuser Verlag. Herausgegeben von M. Gutknecht unter Mitwirkung von P. Henrici, P. Läuchli, H.R. Schwarz; Engl. Transl.: Rutishauser (1990).

RUTISHAUSER, H. (1990) *Lectures on Numerical Mathematics*. Boston: Birkhäuser.

SACK, R. A. & DONOVAN, A. F. (1972) An algorithm for Gaussian quadrature given modified moments. *Numer. Math.*, **18**, 465–478.

SEEWALD, W. (1982) Quotienten–Differenzen–Algorithmus: Beweis der Regeln von Rutishauser. *Numer. Math.*, **40**, 93–98.

STEWART, G. W. (1971) On a companion operator for analytic functions. *Numer. Math.*, **18**, 26–43.

STIEFEL, E. L. (1955) Kernel polynomials in linear algebra and their numerical applications. *Progress Report*. National Bureau of Standards.

STIEFEL, E. L. (1958) Kernel polynomials in linear algebra and their numerical applications. *Further Contributions to the Solution of Simultaneous Linear Equations and the Determination of Eigenvalues*. Applied Mathematics Series, vol. 49. National Bureau of Standards, pp. 1–22.

WHEELER, J. C. (1974) Modified moments and Gaussian quadratures. *J. Math.*, **4**, 287–295.

WILKINSON, J. H. (1965) *The Algebraic Eigenvalue Problem*. Oxford: Clarendon Press.

WILKINSON, J. & REINSCH, C. (1971) *Linear Algebra (Handbook for Automatic Computation)*, vol. 2. Berlin/Heidelberg/New York: Springer-Verlag.