View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by RERO DOC Digital Library

IMA Journal of Numerical Analysis (2011) **31**, 741–754
doi:10.1093/imanum/drq003
Advance Access publication on May 27, 2010

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
*Corresponding author: mhg@math.ethz.ch

AND

BERESFORD N. PARLETT
Department of Mathematics, University of California, Evans Hall, Berkeley, CA 94720-3840, USA
parlett@math.berkeley.edu

[Received on 17 August 2009; revised on 16 December 2009]

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 have 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 honour was well merited as 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, and QR is an improved 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 $\mathbf{A}_2 = \mathbf{L}_1^{-1}\mathbf{A}_1\mathbf{L}_1$ 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 by-product of his qd algorithm. In *The Algebraic Eigenvalue Problem* the eminent numerical analyst J. H. Wilkinson called Rutishauser an 'algorithmic genius' (see Wilkinson, 1965, p. vii) and, regarding the LR algorithm, he wrote in the same book (see Wilkinson, 1965, p. 485) the following: '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

© The author 2010. Published by Oxford University Press on behalf of the Institute of Mathematics and its Applications. All rights reserved.

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 as to 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 (1950, Chapter 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 the 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 the *Eidgenössische Technische Hochschule (ETH)* in 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 (1954b) approached the key problem of determining the poles of the following 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 that $\mathbf{A}$ is an $N \times N$ matrix and that $\mathbf{x}_0$ and $\mathbf{y}_0$ are two N-vectors. Then, for $s_\nu := \mathbf{y}_0^{\mathrm{T}}\mathbf{A}^\nu\mathbf{x}_0$, the series in (2.1) is the Taylor expansion at ∞ of

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

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

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

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 control system given by the state matrix $\mathbf{A}$ and the vectors $\mathbf{x}_0$ and $\mathbf{y}_0$. So Stiefel's proposal for Rutishauser was to determine the eigenvalues of $\mathbf{A}$ given the sequence of moments. Rutishauser (1954b) wrote the following in his introduction: 'Following this suggestion the author developed an algorithm that solves the posed problem.'

We next describe what was previously known 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 the *quotient-difference algorithm* or, briefly, the *QD algorithm*. Nowadays, the abbreviation in lower-case letters, *qd algorithm*, is widely used. It reflects the fact that in English 'quotient' and 'difference' are written in lower case, and it has the additional advantage of emphasizing 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 because 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 a better solution of the matrix eigenvalue problem (see Rutishauser, 1955a), namely, using the Lanczos (1950) algorithm 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 break down and is, in general, unstable, but there are also situations were it works well (see Rutishauser, 1963b, 1976, 1990), 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

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

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

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

This recursion only depends on p. The numerator polynomial q of degree $n-1$ matches the first n terms of the Laurent series of fp, that is,

$$f(z)p(z) - q(z) = \mathcal{O}\left(\frac{1}{z}\right) \quad \text{as } z \to \infty.$$

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 (1732, p. 92) method for finding such a greatest root (see 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) 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 that 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 following double sequence of *Hankel determinants*:

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

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

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

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

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 & Householder, 1966; Gragg, 1973; Householder, 1974). Here are some obvious conclusions.

COROLLARY 3.2. Under the assumptions of Theorem 3.1, we have that $H_{n+1}^{(\nu)} = 0$ $(\forall\, \nu)$. Moreover, if f has n simple poles, then the following holds:

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

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

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

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

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

Naturally, Hadamard was only interested in exact relationships, not in computation. Thus the motivation for developing an efficient algorithm was missing, though, in fact, he had the key in his hands, namely, the striking nonlinear relation

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

among neighbouring Hankel determinants. It appears in expression (14) in Section 17 of Hadamard's (1892, p. 20) thesis, but it is often called *Jacobi's identity* for Hankel determinants (see Henrici, 1974). Note that it expresses the square of $H_k^{(\nu)}$ as a difference of products of next neighbours.

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

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

Unfortunately, when one of the determinants vanishes both the horizontal and the vertical recursions break down.

Another constructive tool that was available when Rutishauser solved Stiefel's problem was the *Chebyshev (1859) algorithm* 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, the Rutishauser qd algorithm serves the same purpose, as we will see. The Chebyshev algorithm was later revived, analysed 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 (1950) 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, 1954a,b,c) in 1954, the first and most fundamental of which was received by the journal *Zeitschrift für Ange-wandte Mathematik und Physik* in 5 August 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 was collected in Rutishauser (1957), which also covered some additional material, in particular, a short appendix on the LR algorithm. Another appendix contained a shortened version of Rutishauser (1954c).

A first announcement on the qd algorithm had been made by Stiefel (1953). Of great importance for the dissemination of the qd algorithm was Henrici's review article (see Henrici, 1958), the first publication on qd in English. It appeared in Volume 49 of the Applied Mathematics Series of the National Bureau of Standards (NBS). The only two other papers in that 81-page volume are Rutishauser's main publication on the LR algorithm (see Rutishauser, 1958) and Stiefel's paper on kernel polynomials (see Stiefel, 1958), which is also related to cg and qd. The volume was issued on 15 January 1958, but it seems to have been compiled long before. In fact, the preface is dated 26 June 1956. Moreover, Rutishauser (1957) cited preprints dated 1956 and 1955 of his and Stiefel's contributions, respectively, to the volume.

Nowhere does Rutishauser clearly state how he found the qd algorithm. He only gives an indication, not a complete derivation. Henrici (1958) wrote that the qd algorithm 'by a simple but ingenious modification of Aitken's method, entirely bypasses the computation of Hankel determinants', and that '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 were hinted at by Henrici (1958) and worked out by Householder (1970) and 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) centred at $H_{k-1}^{(\nu+1)}$, namely,

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

by

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

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

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

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

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

and multiply it by

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

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

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

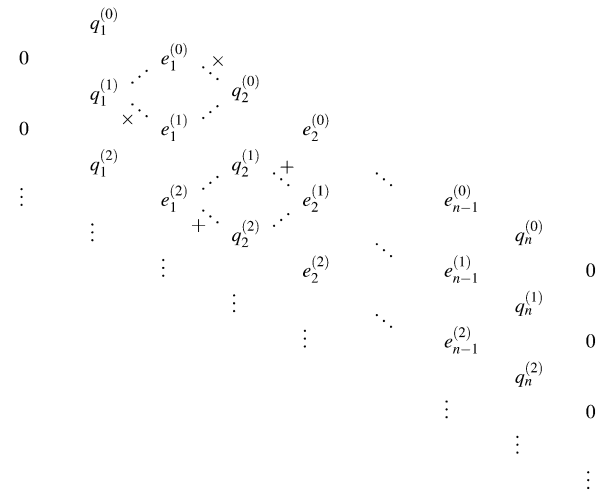

FIG. 1. The qd table of a rational function f of degree n and manifestations of the two rhombus rules (4.5) and (4.4), namely, $q_2^{(0)} \times e_1^{(0)} = q_1^{(1)} \times e_1^{(1)}$ and $q_2^{(1)} + e_2^{(1)} = q_2^{(2)} + e_1^{(2)}$.

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 writing down the quantities $e_k^{(\nu)}$ and $q_k^{(\nu)}$ in a triangular scheme called a *qd scheme* (also known as a *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, and 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 as follows:

$$e_k^{(\nu)} = q_k^{(\nu+1)} + e_{k-1}^{(\nu+1)} - q_k^{(\nu)}, \quad q_{k+1}^{(\nu)} = q_k^{(\nu+1)} e_k^{(\nu+1)} / e_k^{(\nu)}. \tag{4.6}$$

Alternatively, they can be used to build up the scheme from its top diagonal as follows:

$$q_k^{(\nu+1)} = e_{k-1}^{(\nu+1)} - e_k^{(\nu)} - q_k^{(\nu)}, \quad e_k^{(\nu+1)} = (q_{k+1}^{(\nu)} / q_k^{(\nu+1)}) e_k^{(\nu)}. \tag{4.7}$$

The latter application is called the *progressive qd algorithm*. Rutishauser soon noted that, for stability reasons, this version is much more useful. But it requires some preparatory work to come up with

[1] According to the footnotes of Henrici (1958) and Rutishauser (1957), the name 'rhombus rules' was coined by Stiefel (see Stiefel, 1955, p. 42 or Stiefel, 1958, p. 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 also an 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.

the right data before it can be started. Rutishauser's knowledge of the Lanczos algorithm, tridiagonal matrices and continued fractions enabled him to see ways of how to do that. We present here some of the details as we think they are important from a historical perspective, but they may be hard to digest for many readers.

Rutishauser (1954b, Section 8, (19)) knew that the top diagonal of the qd scheme can be computed with what is now called the Lanczos algorithm and was described in Lanczos (1950) and Rutishauser (1953), since it reduces the matrix to tridiagonal form, which is directly reflected by the top diagonal of the qd scheme. He also mentioned that, if $\mathbf{A}$ is symmetric positive definite, then the conjugate gradient method (see Hestenes & Stiefel, 1952) provides the top diagonal too (see Rutishauser, 1954b, Section 9, (24)). And in the final summary of Rutishauser (1955a) he listed 'the methods of C. Lanczos or W. Givens' as tools for this preprocessing step. Additionally, for the task of finding the poles of a rational function f, that is, the zeros of its denominator polynomial p, Rutishauser (1954b, Section 6) pointed out how to obtain the top diagonal by computing the finite continued fraction of f.

However, soon he discovered a better way to find the zeros of a polynomial. In fact, in a footnote that Rutishauser (1954a) added when proofreading, it is mentioned that Alston Householder pointed out to him that Aitken (1931) had also presented a progressive algorithm, and this was for an extension of the scheme he had introduced in Aitken (1926): the triangular table (3.9) of Hankel matrices can be extended upwards so that it can be started from a horizontal row that is readily computed from the coefficients of a given polynomial p whose zeros have to be determined. In the sequel Rutishauser used a formula of Wronski (1811) that relates the power series of a function with one of its reciprocal to show that his progressive qd algorithm can be applied in an analogue extension of his qd scheme. Again, the first rows of q-values and e-values can then be easily computed from the coefficients of the given polynomial p (see Rutishauser, 1956a, 1957, Section 6). Recall that, in contrast, the original qd algorithm would require us to compute the Laurent series of $1/p$.

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

$$\left.\begin{aligned} \lim_{\nu\to\infty} q_k^{(\nu)} &= \lambda_k, \\ \lim_{\nu\to\infty} e_k^{(\nu)} &= 0 \end{aligned}\right\} \quad (k = 1, 2, \ldots, n). \tag{4.8}$$

This behaviour means that the original qd algorithm (4.6) for building up the table from its first column is a computational disaster as the first formula of (4.6) inevitably leads to the cancellation of leading digits. In contrast, the progressive form (4.7) is a version that is still of importance. It avoids the highly ill-conditioned computation of the table from the moments (see, e.g., Gautschi, 1968, 1982).

Rutishauser (1954b) also gave the generalization of (4.8) to meromorphic functions, referring for its proof to Hadamard ('The proof of the proposition can be easily deduced 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) was erroneous and promoted a new direct and simple proof due to his student Seewald (1982).

There is another very elegant derivation of the rhombus rules and the qd algorithm that is based on various ways to expand $f(z)$ into a continued fraction. This derivation was given in the original paper Rutishauser (1954b) and in Rutishauser (1957), 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 (1954a,b,c, 1955a) early qd papers was there any hint of the LR algorithm. (The last of these four papers, on finding eigenvalues with the qd algorithm, was received on 19 July 1954.) What follows is the high point of the drama of the discovery of LR.

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

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

At some historic moment in the second half of 1954, Rutishauser must have made the remarkable observation that his rhombus rules (4.4) and (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) and (4.5), this product equals exactly $\mathbf{T}_{\nu+1}$ as defined by (5.1) with ν replaced by $\nu + 1$. Thus

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

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 then we know from Hadamard's theorem or from the qd algorithm that $\mathbf{L}_\nu \to \mathbf{I}$ as $\nu \to \infty$, and 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 then we can still apply an adaptation of what Henrici called Rutishauser's rule.

That moment in 1954 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 of conserving 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 (1953) rotations 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 analysed convergence and introduced spectral shifts for the progressive qd algorithm before (see Rutishauser (1954a, Section 7) or Rutishauser (1957, Chapter 2, Section 8) and Rutishauser (1955a, Section 4) or Rutishauser (1957, Chapter 3, Section 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 (1955b) first publication on the LR algorithm was, in 1955, a two-page note in French in the *Comptes Rendus*, the primary journal for research announcement in mathematics of the time. In the following year, according to Rutishauser (1957), he produced a mimeographed 51-page preprint in English, entitled 'Report on the solution of eigenvalue problems with the LR–transformation' (see Rutishauser, 1956b),[3] but it was two years later that this article was properly published by the National Bureau of Standards (see Rutishauser, 1958) in the volume that also contained Henrici's (1958) review of the qd algorithm. Only then did the qd and LR algorithms make their appearance for the English speaking world, except for those few people who had received the LR preprint before. Rutishauser (1957) included a five-page appendix on the LR transformation in this substantial report, which compiled and updated most of his previous work on qd but was still in German.

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

6. LR abandoned

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, 10 years after the National Bureau of Standards volume, Rutishauser wrote an updated script of some of his research on qd with no mention of LR (see Rutishauser, 1968). We now give our perspective on the puzzle.

A few 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 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's (1961, 1962) QR algorithm shortly after its introduction. Francis's 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 (1963a) presented at the IFIP Congress in 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 (1976, 1990) somehow disliked the QR algorithm—there was no room for it in his lectures.

Our quotation above shows that Wilkinson, among others, appreciated the depth and originality of the LR algorithm. In order to understand Rutishauser's neglect of his amazing discovery, let us consider what might have deterred him.

(A) Today we take it for granted that a given full matrix would be transformed into a more compact banded form that would be invariant under LR. The Hessenberg form is one option. There is no evidence in Rutishauser (1958) that he gave this any thought. Yet, without it, the cost of each step is prohibitive.

(B) The LR factorization had a bad reputation. Row pivoting, at least, would be needed. We know that a PLR algorithm (P for permutation) is a viable option, particularly on the Hessenberg form. However, there is no evidence in Rutishauser (1958) that he contemplated it. More generally, permutations for stability would usually spoil the bandwidth.

(C) The 'implicit L' property of triangular reduction to Hessenberg form was not part of the intellectual landscape. It says that, if $\mathbf{H} = \mathbf{L}^{-1}\mathbf{A}\mathbf{L}$ is unreduced Hessenberg, then all quantities are determined by $\mathbf{A}$ and the first column of $\mathbf{L}$. However, an implicit LR transform on a general matrix is not possible.

(D) Rutishauser suggested an ingenious but expensive shift strategy for LR on a full matrix. However, even with this improvement, the early stages would have been prohibitive. What a contrast to the qd algorithm for which he had an analysis of two effective shift strategies, namely, to use either the last q-value or a more expensive but cubically convergent variant.

The features mentioned above are familiar to the numerical linear algebra community now because of their role in making a success of the QR algorithm. In his explicit shift LR algorithm, Rutishauser would have been keenly aware of the possible loss of information in subtracting the shift from the diagonal entries. The stable descendant of LR, namely, QR, came so quickly on the heels of the (much delayed) presentation of the LR algorithm to the English speaking community that there was no time for these four features to be developed in the context of LR.

We think that the intellectual successor to Rutishauser's LR algorithm was Kublanovskaja's (1961) QR algorithm. This was one of the four backward stable algorithms for a general full matrix presented by her. We must emphasize that, in contrast, the implicit version of the QR algorithm presented by Francis (1961, 1962), right from the start, was a novel contribution in the same league as Rutishauser's idea of reversing the order of factors. The reason that Francis deserves the credit for the success of QR is that several crucial features were never mentioned by Kublanovskaja, namely, the following:

(1) the invariance of the Hessenberg form;
(2) clever shift strategies for complex eigenvalues of real matrices;
(3) the implicit implementation of shifts.

Acknowledgement

The authors appreciated the referees' very helpful comments, in particular, the mentioning of the footnote on Householder's hint to Rutishauser regarding the paper of Aitken (1931), the authors also need to thank Gerhard Wanner for verifying the Bernoulli reference.

REFERENCES

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

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

BACKUS, J. W., BAUER, F. L., GREEN, J., KATZ, C., MCCARTHY, J., NAUR, P. (ed.), 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., NAUR, P. (ed.), 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) Observationes de Seriebus quae formantur ex additione vel subtractione quacunque terminorum se mutuo consequentium, ubi praesertim earundem insignis usus pro inveniendis radic[ibus] omnium Aequationum Algebraicarum ostenditur. *Commentarii Acad. Sci. Imper. Petropol.*, **III**, pp. 85–100.

CHEBYSHEV, P. (1859) Sur l'interpolation par la méthode des moindres carrés. *Mém. Acad. Impér. Sci. St. Pétersbourg, Sér. VII*, **1**, 1–24; *Oeuvres*, **1**, 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. Comput.*, **22**, 251–270.

GAUTSCHI, W. (1970) On the construction of Gaussian quadrature rules from modified moments. *Math. Comput.*, **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, vol. 29. Washington, DC: US 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, France. Also in *J. Math. Pures Appl.*, **8**, 101–186 (1992).

HENRICI, P. (1958) The quotient-difference algorithm. *Nat. Bur. Standards Appl. Math. Ser. no*., **49**, 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, pp. 133–147, 149–168, 169–191, 193–215.

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

HOUSEHOLDER, A. S. (1970) *The Numerical Treatment of a Single Non-linear Equation*. New York: McGraw-Hill.

HOUSEHOLDER, 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.

KUBLANOVSKAJA, V. N. (1961) Certain algorithms for the solution of the complete problem of eigenvalues. *Sov. Math. Dokl.*, **2**, 17–19.

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

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

PARLETT, B. N. (2000) The QR algorithm. *Comput. Sci. Eng.*, **2000**, 38–42.

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

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. (1956a) Eine Formel von Wronski und ihre Bedeutung für den Quotienten–Differenzen–Algorithmus. *Z. Angew. Math. Phys.*, **7**, 164–169.

RUTISHAUSER, H. (1956b) Report on the solution of eigenvalue problems with the LR–transformation. *Mimeographed Report*. Institut für Angewandte Mathematik, Zürich: ETH.

RUTISHAUSER, H. (1957) *Der Quotienten-Differenzen-Algorithmus*. Mitteilungen aus dem Institut für Angewandte Mathematik, ETH, Nr. 7. Basel, Switzerland: Birkhäuser.

RUTISHAUSER, H. (1958) Solution of eigenvalue problems with the LR-transformation. *Nat. Bur. Standards Appl. Math. Ser.*, **49**, 47–81.

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

RUTISHAUSER, H. (1963a) Numerical experiments with the QD-transformation of J.G.F. Francis. *Information Processing 1962*, Proceedings of the IFIP Congress, Munich, 1962 (C. M. Popplewell ed.). Amsterdam: North-Holland, pp. 202–203.

RUTISHAUSER, H. (1963b) Stabile Sonderfälle des Quotienten–Differenzen–Algorithmus. *Numer. Math.*, **5**, 95–112.

RUTISHAUSER, H. (1967) *Description of ALGOL 60*. Handbook for Automatic Computation, vol. Ia. Berlin: Springer.

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), p. 32.

RUTISHAUSER, H. (1976) *Vorlesungen über numerische Mathematik (2 Bände)*. Basel, Switzerland: Birkhäuser. Herausgegeben von M. Gutknecht unter Mitwirkung von P. Henrici, P. Läuchli, H. R. Schwarz; English Translation: Rutishauser (1990).

RUTISHAUSER, H. (1990) *Lectures on Numerical Mathematics*. Boston, MA: 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. *Nat. Bur. Standards Appl. Math. Ser*., **49**, 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) *Handbook for Automatic Computation: Volume 2: Linear Algebra*. Berlin: Springer.

WRONSKI (1811) *Introduction à la Philosophie des Mathématiques*, et technie de l'algorithmie. Paris: Courcier. Available at http://books.google.de/books?id=MC8VAAAAQAAJ&printsec=frontcover& dq=Introduction +%C3%A0+la+philosophie+des+math%C3%A9matiques,+et+technie+de+l%27algorithmie+%281811%29 &source=bl&ots=8rQ9iaEA2d&sig=_PGK2QrKqrxu6zHRobsJ2eyh9Ts&hl=de&ei=54_pS_z8LpajOMSMxf0 K&sa=X &oi=book_result&ct=result&resnum=2&ved=0CCEQ6AEwAQ#v=onepage&q&f=false.