From qd to LR and QR

Martin H. Gutknecht and Beresford N. Parlett

1 Seminar for Applied Mathematics
   ETH Zurich

2 Department of Mathematics
   University of California
   Berkeley

Tsinghua University, Beijing, 22 Sep. 2006
Numerical Linear Algebra in 1952: $Ax = b$

$Ax = b$ with full or banded $A$:

**Gauss elimination** (in various versions)

$Ax = b$ with sparse spd $A$:

**Chebyshev iteration, SOR, Conjugate Gradients**

$Ax = b$ with sparse nonsym. or sym. indef. $A$:

various “relaxation methods” with limited applicability:

**Jacobi iteration, Gauss-Seidel, Richardson’s iteration, SOR, ...**
Numerical Linear Algebra in 1952: $Ax = x\lambda$

$Ax = x\lambda$ with full sym. $A$:

**Jacobi’s method** (rotations $\leadsto$ diagonal form)

$Ax = x\lambda$ with full or sparse $A$:

**power method** (dominant eigenpair only)

“fractional”, **inverse iteration** (any single eigenpair)

various methods for computing the charact. polynomial, including:

**Krylov’s method** (reduction to companion form, char. pol.)

**Lanczos’ method** (reduction to tridiagonal form, char. pol.)

**Arnoldi’s method** (red. to Hessenberg form, char. pol.)
Eduard Stiefel founded the IAM in January 1948.

His aim was to build an electronic computer and to start up numerical analysis in Zurich.

He hired two assistants (postdocs): Heinz Rutishauser and Ambros P. Speiser.

Oct. 18, 1948, – March 19, 1949, he made a five months trip to Amsterdam (van Wijngarden), New York (Eckert & Thomas at IBM, Courant & Friedrichs at NYU, Lowan & Salzer at NBS), Washington (Mina Rees at ONR), Boston (Aiken, Birkhoff, Young), Princeton (von Neumann).

Rutishauser and Speiser were both sent to the USA for a full year (1949). Each of them spent half the time with Aiken and von Neumann.
Eduard Stiefel founded the IAM in January 1948.

His aim was to build an electronic computer and to start up numerical analysis in Zurich.

He hired two assistants (postdocs): Heinz Rutishauser and Ambros P. Speiser.

Oct. 18, 1948, – March 19, 1949, he made a five months trip to Amsterdam (van Wijngarden), New York (Eckert & Thomas at IBM, Courant & Friedrichs at NYU, Lowan & Salzer at NBS), Washington (Mina Rees at ONR), Boston (Aiken, Birkhoff, Young), Princeton (von Neumann).

Rutishauser and Speiser were both sent to the USA for a full year (1949). Each of them spent half the time with Aiken and von Neumann.
In August 1950 the **Zuse Z4 computer** was installed. Research for developing a new electronic computer continued. Both with respect to hardware and software seminal contributions were made. In 1952 Rutishauser designed a **compiler** (Habil. thesis). But only in July 1956 the **ERMETH** was running.

In numerical analysis:

Stiefel (simult. with Hestenes) discovered the **CG method**. Rutishauser (simultaneously with Dahlquist) investigated the **stability of ODE methods**. Rutishauser (after Lanczos) worked on the (nonsym.) **Lanczos biorthogonalisation (BO, B̂O) algorithm**.
In August 1950 the **Zuse Z4 computer** was installed.

Research for developing a new electronic computer continued. Both with respect to hardware and software seminal contributions were made.

In 1952 Rutishauser designed a **compiler** (Habil. thesis). But only in July 1956 the **ERMETH** was running.

In numerical analysis:

Stiefel (simult. with Hestenes) discovered the **CG method**.

Rutishauser (simultaneously with Dahlquist) investigated the **stability of ODE methods**.

Rutishauser (after Lanczos) worked on the (nonsym.) **Lanczos biorthogonalisation (BO, BIO) algorithm**.
Moments and their generating function

Given: \( N \times N \) matrix \( A \) and \( x_0, y_0 \in \mathbb{R}^N \), let

\[
f(z) : \equiv \langle y_0, (zI - A)^{-1} x_0 \rangle = \langle y_0, \frac{1}{z} (I - \frac{1}{z} A)^{-1} x_0 \rangle \tag{1}
\]

\( f \) is a rational function of type \((N - 1, N)\), so \( f(\infty) = 0 \).

The poles of \( f \) are eigenvalues of \( A \).

\( f \) can be expanded into a power series in \( z^{-1} \):

\[
f(z) = \sum_{k=0}^{\infty} \frac{S_k}{z^{k+1}} = \frac{S_0}{z} + \frac{S_1}{z^2} + \frac{S_2}{z^3} + \ldots . \tag{2}
\]

where

\[
S_k = y_0^T A^k x_0 \tag{3}
\]
is the \textit{kth moment} (= Markov parameter, Schwarz constant).
Moments and their generating function

Given: $N \times N$ matrix $A$ and $x_0, y_0 \in \mathbb{R}^N$, let

$$f(z) \equiv \left\langle y_0, (zI - A)^{-1} x_0 \right\rangle = \left\langle y_0, \frac{1}{z} (I - \frac{1}{z} A)^{-1} x_0 \right\rangle \quad (1)$$

$f$ is a rational function of type $(N - 1, N)$, so $f(\infty) = 0$.

The poles of $f$ are eigenvalues of $A$.

$f$ can be expanded into a power series in $z^{-1}$:

$$f(z) = \sum_{k=0}^{\infty} \frac{S_k}{z^{k+1}} = \frac{S_0}{z} + \frac{S_1}{z^2} + \frac{S_2}{z^3} + \ldots \quad (2)$$

where

$$S_k = y_0^T A^k x_0 \quad (3)$$

is the $k$th moment (= Markov parameter, Schwarz constant).
Moments and their generating function
(cont’d)

Clearly we could also write \( zf(z) \) as a function of \( \zeta \equiv z^{-1} \):

\[
\phi(\zeta) \equiv \zeta^{-1} f(\zeta^{-1}) = \left\langle y_0, (I - \zeta A)^{-1} x_0 \right\rangle = s_0 + s_1 \zeta + s_2 \zeta^2 + \ldots
\]

(4)

\( \phi \) is also a rational function of type \((N - 1, N)\).

Assume the eigenvalues \( \lambda_k \) of \( A \) are ordered such that

\[
|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{N-1}| \geq |\lambda_N|
\]

The series of \( f \) converges for \(|z| > |\lambda_1|\).

The series of \( \phi \) converges for \(|\zeta| < |\lambda_1|^{-1}\).

Could as well look for zeros of a polynomial \( b_0 + \cdots + b_n \zeta^N \):

\[
\phi(\zeta) \equiv \frac{1}{b_0 + b_1 \zeta + \cdots + b_n \zeta^N} = s_0 + s_1 \zeta + s_2 \zeta^2 + \ldots
\]

Martin H. Gutknecht and Beresford N. Parlett

From qd to LR and QR
Moments and their generating function (cont’d)

Clearly we could also write $zf(z)$ as a function of $\zeta := z^{-1}$:

$$
\varphi(\zeta) \equiv \zeta^{-1}f(\zeta^{-1}) = \left\langle y_0, (I - \zeta A)^{-1} x_0 \right\rangle = s_0 + s_1\zeta + s_2\zeta^2 + \ldots
$$

(4)

$\varphi$ is also a rational function of type $(N - 1, N)$.

Assume the eigenvalues $\lambda_k$ of $A$ are ordered such that

$$
|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{N-1}| \geq |\lambda_N|
$$

The series of $f$ converges for $|z| > |\lambda_1|$.

The series of $\varphi$ converges for $|\zeta| < |\lambda_1|^{-1}$.

Could as well look for zeros of a polynomial $b_0 + \cdots + b_n\zeta^N$:

$$
\varphi(\zeta) \equiv \frac{1}{b_0 + b_1\zeta + \cdots + b_n\zeta^N} = s_0 + s_1\zeta + s_2\zeta^2 + \ldots
$$
Moments and their generating function

Clearly we could also write $zf(z)$ as a function of $\zeta :\equiv z^{-1}$:

$$
\varphi(\zeta) :\equiv \zeta^{-1} f(\zeta^{-1}) = \langle y_0, (I - \zeta A)^{-1} x_0 \rangle = s_0 + s_1 \zeta + s_2 \zeta^2 + \ldots
$$

(4)

$\varphi$ is also a rational function of type $(N - 1, N)$.

Assume the eigenvalues $\lambda_k$ of $A$ are ordered such that

$$
|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{N-1}| \geq |\lambda_N|
$$

The series of $f$ converges for $|z| > |\lambda_1|$.

The series of $\varphi$ converges for $|\zeta| < |\lambda_1|^{-1}$.

Could as well look for zeros of a polynomial $b_0 + \cdots + b_n \zeta^N$:

$$
\varphi(\zeta) :\equiv \frac{1}{b_0 + b_1 \zeta + \cdots + b_n \zeta^N} = s_0 + s_1 \zeta + s_2 \zeta^2 + \ldots
$$
Assume, $\lambda_1$ is a **single dominant eigenvalue**:

$$|\lambda_1| > |\lambda_2|$$

Then, according to D. Bernoulli, 1732, and J. König, 1884,

$$\frac{S_{\nu+1}}{S_{\nu}} \rightarrow \lambda_1 \quad \text{as} \quad \nu \rightarrow \infty$$

Around 1953 Stiefel suggested to Rutishauser to try to compute all eigenvalues from the moments.

By now we know that this was a **bad idea**, since this dependence is highly ill-conditioned.

In theory, the problem had been solved before by Hadamard, 1892 (his PhD thesis!) and Aitken, 1926, 1931.
Assume, $\lambda_1$ is a \textit{single dominant eigenvalue}:

\[ |\lambda_1| > |\lambda_2| \]

Then, according to D. Bernoulli, 1732, and J. König, 1884,

\[ \frac{s_{\nu+1}}{s_{\nu}} \rightarrow \lambda_1 \quad \text{as} \quad \nu \rightarrow \infty \]

\textit{Around 1953 Stiefel suggested to Rutishauser to try to compute all eigenvalues from the moments.}

By now we know that this was a \textit{bad idea}, since this dependence is highly ill-conditioned.

In theory, the problem had been solved before by Hadamard, 1892 (his PhD thesis!) and Aitken, 1926, 1931.
Given the power series of $f$ in (2) or of $\varphi$ in (4), let $H_0^{(\nu)} \equiv 1$, and define the *Hankel determinants*

\[
H_k^{(\nu)} = \begin{vmatrix}
S_\nu & S_{\nu+1} & \cdots & S_{\nu+k-1} \\
S_{\nu+1} & S_{\nu+2} & \cdots & S_{\nu+k} \\
\vdots & \vdots & \ddots & \vdots \\
S_{\nu+k-1} & S_{\nu+k} & \cdots & S_{\nu+2k-2}
\end{vmatrix}
\]

\[(k = 1, 2, \ldots; \quad \nu = 0, 1, \ldots)\]

**Theorem**

*[Hadamard (1892)]* If $|\lambda_{k+1}| < \Lambda < |\lambda_k|$, then, as $\nu \to \infty$,

\[
H_k^{(\nu)} = \text{const} \cdot (\lambda_1 \cdots \lambda_k)^\nu \left[ 1 + \mathcal{O} \left( \frac{\Lambda}{|\lambda_k|} \right)^\nu \right]
\]

For a simpler proof see Henrici (1958) or Henrici (1974).
**Corollary**

If $f$ has $N$ simple poles, then

1. $H_k^{(ν)} \neq 0 \ (k = 1, \ldots, N)$ for large enough $ν$, and $H_{N+1}^{(ν)} = 0 \ (∀ ν)$.

2. If $|λ_k| > |λ_{k+1}|$ then

\[
\frac{H_k^{(ν+1)}}{H_k^{(ν)}} \rightarrow \lambda_1 \lambda_2 \cdots \lambda_k \quad \text{as} \quad ν \rightarrow \infty. \quad (5)
\]

3. If $|λ_{k−1}| > |λ_k| > |λ_{k+1}|$ then

\[
q_k^{(ν)} :≡ \frac{H_k^{(ν+1)}}{H_k^{(ν)}} \cdot \frac{H_{k−1}^{(ν)}}{H_k^{(ν+1)}} \rightarrow \lambda_k \quad \text{as} \quad ν \rightarrow \infty. \quad (6)
\]
Aitken’s scheme (1931) (cont’d)

Computing, for fixed $\nu$, the Hankel determinants $H_1^{(\nu)}, \ldots, H_N^{(\nu)}$ (if nonzero) requires the LU decomposition of the matrix $H_N^{(\nu)}$.

Aitken (1926, 1931) knew what is now called “Jacobi identity” ("theorem of compound determinants")

\[
\left( H_k^{(\nu)} \right)^2 = H_k^{(\nu-1)} H_k^{(\nu+1)} + H_k^{(\nu-1)} H_{k+1}^{(\nu+1)}. \tag{7}
\]

It had also been known to Hadamard, but Aitken used to build up — from the left or from the top — the table

\[
\begin{array}{ccccccc}
1 \\
1 & H_1^{(0)} \\
1 & H_1^{(1)} & H_2^{(0)} \\
1 & H_1^{(2)} & H_2^{(1)} & H_3^{(0)} \\
1 & H_1^{(3)} & H_2^{(2)} & H_3^{(1)} & H_4^{(0)} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
\]
Rutishauser’s qd algorithm (QD-Algorithmus)

Rutishauser (1954a) knew Aitken’s work and refers to (??),

\[ \frac{H_k^{(\nu+1)}}{H_k^{(\nu)}} \rightarrow \lambda_1 \lambda_2 \cdots \lambda_k \quad \text{as} \quad \nu \rightarrow \infty \]

as the key to computing non-dominant poles.

But instead of computing the \( H_k^{(\nu)} \)–table, he headed directly for recurrences for

\[ q_k^{(\nu)} : \equiv \frac{H_k^{(\nu+1)}}{H_k^{(\nu)}} \cdot \frac{H_k^{(\nu)}}{H_k^{(\nu+1)}} \quad \text{and} \quad e_k^{(\nu)} : \equiv \frac{H_k^{(\nu+1)}}{H_k^{(\nu+1)}} \cdot \frac{H_k^{(\nu)}}{H_k^{(\nu+1)}} \]

\[ (8) \]

In Rutishauser (1954a) he derives the formulas needed for \( q_2^{(\nu)} \), and then states recursions for general \( k \).

The correctness of the general recursions follows later in that paper from connections to continued fractions.
Rutishauser’s qd algorithm (cont’d)

qd table (QD–Schema):

\[
\begin{array}{cccccc}
q_1^{(0)} & e_1^{(0)} & q_2^{(0)} & \times & \vdots & \\
0 & e_1^{(1)} & e_2^{(0)} & + & \vdots & e_{N-1}^{(0)} \\
q_1^{(1)} & q_2^{(1)} & \vdots & + & e_{N-1}^{(1)} & q_N^{(0)} \\
0 & e_1^{(2)} & e_2^{(1)} & \vdots & e_{N-1}^{(2)} & 0 \\
q_1^{(2)} & q_2^{(2)} & \vdots & + & e_{N-1}^{(2)} & 0 \\
0 & e_1^{(3)} & e_2^{(2)} & \vdots & e_{N-1}^{(3)} & 0 \\
q_1^{(3)} & q_2^{(3)} & \vdots & + & e_{N-1}^{(3)} & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
q_1^{(0)} & q_2^{(0)} & \times & \vdots & \vdots & \vdots \\
0 & e_1^{(0)} & e_2^{(1)} & \vdots & e_{N-1}^{(0)} & 0 \\
\end{array}
\]

\[
e_1^{(0)} \cdot q_2^{(0)} = q_1^{(1)} \cdot e_1^{(1)} \quad q_2^{(1)} + e_2^{(1)} = e_1^{(2)} + q_2^{(2)}
\]
Rhombus rules (called so by Stiefel, 1955) of qd algorithm:

For building up the table columnwise from left to right:

\[
\begin{align*}
e^{(\nu)}_k &:= e^{(\nu+1)}_{k-1} + q^{(\nu+1)}_k - q^{(\nu)}_k \\
q^{(\nu)}_{k+1} &:= q^{(\nu+1)}_k \frac{e^{(\nu+1)}_k}{e^{(\nu)}_k} \\
& \text{for } k = 1, 2, \ldots 
\end{align*}
\]  \hspace{1cm} (9)

For building up the table rowwise, from top to bottom:

\[
\begin{align*}
q^{(\nu+1)}_k &:= q^{(\nu)}_k + e^{(\nu)}_k - e^{(\nu+1)}_{k-1} \\
e^{(\nu+1)}_k &:= e^{(\nu)}_k \frac{q^{(\nu+1)}_k}{q^{(\nu)}_{k+1}} \\
& \text{for } k = 1, 2, \ldots 
\end{align*}
\]  \hspace{1cm} (10)

Recursions (9,10) are the basis of the progressive qd algorithm (the relevant version).
Rhombus rules (called so by Stiefel, 1955) of qd algorithm:

For building up the table columnwise from left to right:

\[
\begin{align*}
e^{(\nu)}_k & := e^{(\nu+1)}_{k-1} + q_k^{(\nu+1)} - q_k^{(\nu)} \\
q_{k+1}^{(\nu)} & := q_k^{(\nu+1)} \frac{e^{(\nu+1)}_k}{e^{(\nu)}_k}
\end{align*}
\]  

\( (k = 1, 2, \ldots) \)  

(9)

For building up the table rowwise, from top to bottom:

\[
\begin{align*}
q_k^{(\nu+1)} & := q_k^{(\nu)} + e_k^{(\nu)} - e_k^{(\nu+1)} \\
e^{(\nu+1)}_k & := e_k^{(\nu)} \frac{q_k^{(\nu+1)}}{q_k^{(\nu)} q_{k+1}^{(\nu)}}
\end{align*}
\]  

\( (k = 1, 2, \ldots) \)  

(10)

Recursions (??) are the basis of the progressive qd algorithm (the relevant version).
**$\rho$–table (P–Schema):**

<table>
<thead>
<tr>
<th></th>
<th>$p_0^{(0)}$</th>
<th>$p_1^{(0)}$</th>
<th>$q_2^{(0)}$</th>
<th>$p_3^{(0)}$</th>
<th>$p_{N-1}^{(0)}$</th>
<th>$p_N^{(0)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 \equiv p_0^{(1)}$</td>
<td>$p_0^{(1)}$</td>
<td>$p_2^{(1)}$</td>
<td>$p_3^{(1)}$</td>
<td>$\ldots$</td>
<td>$p_{N-1}^{(1)}$</td>
<td>$p_N^{(1)}$</td>
</tr>
<tr>
<td>$1 \equiv p_0^{(2)}$</td>
<td>$p_0^{(2)}$</td>
<td>$p_2^{(2)}$</td>
<td>$p_3^{(2)}$</td>
<td>$\ldots$</td>
<td>$p_{N-1}^{(2)}$</td>
<td>$p_N^{(2)}$</td>
</tr>
<tr>
<td>$1 \equiv p_0^{(3)}$</td>
<td>$p_0^{(3)}$</td>
<td>$p_2^{(3)}$</td>
<td>$p_3^{(3)}$</td>
<td>$\ldots$</td>
<td>$p_{N-1}^{(3)}$</td>
<td>$p_N^{(3)}$</td>
</tr>
<tr>
<td>$1 \equiv p_0^{(4)}$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
</tbody>
</table>

$p_2^{(0)}(z) := z p_1^{(1)}(z) - q_2^{(0)} p_1^{(0)}(z)$

In the last column, $p_N^{(0)} = p_N^{(1)} = \ldots$ is the minimal polynomial.
In our tables we assumed the generic case where the qd algorithm does not break down and the vectors \((y_0, x_0)\) in the definition of the moments, \(s_k = y_0^T A^k x_0\), are in general position.

These vectors are also the starting vectors in the nonsymmetric Lanczos algorithm (Lanczos, 1950).

When introducing the \(p\)-table Rutishauser (1954a, Sect. 4) points out that \(p_k^{(0)} \ (k = 0, 1, \ldots)\) in the top diagonal appear in Lanczos (1950) in an algorithm for computing the characteristic polynomial from the moments. (This algorithm is basically the staircase recurrence for the Padé denominators.)

Later, in Sect. 8, he proved that these polynomials are equal to the *Lanczos polynomials* constructed implicitly in the B\(\hat{V}\)O algorithm.
In our tables we assumed the generic case where the qd algorithm does not break down and the vectors \((y_0, x_0)\) in the definition of the moments, \(s_k = y_0^T A^k x_0\), are in general position. These vectors are also the starting vectors in the nonsymmetric Lanczos algorithm (Lanczos, 1950).

When introducing the \(p\)-table Rutishauser (1954a, Sect. 4) points out that \(p_k^{(0)} (k = 0, 1, \ldots)\) in the top diagonal appear in Lanczos (1950) in an algorithm for computing the characteristic polynomial from the moments. (This algorithm is basically the staircase recurrence for the Padé denominators.) Later, in Sect. 8, he proved that these polynomials are equal to the \textit{Lanczos polynomials} constructed implicitly in the BiO algorithm.
The $p$–table is defined by the initial column $p^{(\nu)}_0 \equiv 1$ and the left-to-right recurrence

$$p^{(\nu)}_k(z) := zp^{(\nu+1)}_{k-1}(z) - q^{(\nu)}_k p^{(\nu)}_{k-1}(z) \quad (11)$$

Rutishauser (1954a) derived also a top-to-bottom recurrence

$$p^{(\nu+1)}_k(z) := p^{(\nu)}_k(z) - e^{(\nu)}_k p^{(\nu+1)}_{k-1}(z) \quad (12)$$

and the diagonal 3-term recurrence (with $e^{(\nu)}_0 \equiv 0$, $p^{(\nu)}_0 \equiv 1$)

$$p^{(\nu)}_{k+1}(z) := \left[ z - q^{(\nu)}_{k+1} - e^{(\nu)}_k \right] p^{(\nu)}_k(z) + e^{(\nu)}_k q^{(\nu)}_k p^{(\nu)}_{k-1}(z) \quad (13)$$
A polynomial 3-term recurrence of the form (eq) means that the polynomials \( p_k^{(\nu)} (k = 0, 1, \ldots; \ \nu \ \text{fixed}) \) are (in today’s lingo) formal orthogonal polynomials (FOPs).

These FOPs are the denominators of the convergents = approximants = “partial sums” of a continued fraction.

These convergents are \textit{Padé approximants} (at \( \infty \)) associated with the moments \( s_{k+\nu} (k = 0, 1, \ldots; \ \nu \ \text{fixed}) \) of the function \( f_{\nu}(z) \) defined by

\[
f(z) = \frac{s_0}{z} + \frac{s_1}{z^2} + \cdots + \frac{s_{\nu-1}}{z^\nu} + \frac{f_{\nu}(z)}{z^\nu}.
\]  

These FOPs are also the \textit{Lanczos polynomials} for the \textit{Lanczos BIO process} for \( A \) started with the pair \((y_0, A^\nu x_0)\).
A polynomial 3-term recurrence of the form $(\ldots)$ means that the polynomials $p_k^{(v)}$ ($k = 0, 1, \ldots$; $v$ fixed) are (in today’s lingo) **formal orthogonal polynomials (FOPs)**.

These FOPs are the denominators of the *convergents = approximants = “partial sums”* of a continued fraction.

These convergents are *Padé approximants* (at $\infty$) associated with the moments $s_{k+\nu}$ ($k = 0, 1, \ldots$; $\nu$ fixed) of the function $f_\nu(z)$ defined by

$$f(z) = \frac{s_0}{z} + \frac{s_1}{z^2} + \cdots + \frac{s_{\nu-1}}{z^\nu} + \frac{f_\nu(z)}{z^\nu}.$$  \hspace{1cm} (14)

These FOPs are also the *Lanczos polynomials* for the *Lanczos BiO process* for $A$ started with the pair $(y_0, A^\nu x_0)$. 

---

**From qd to LR and QR**

_Martin H. Gutknecht and Beresford N. Parlett_
A polynomial 3-term recurrence of the form \((\ldots)\) means that the polynomials \(p_k^{(\nu)} (k = 0, 1, \ldots; \ \nu \text{ fixed})\) are (in today’s lingo) \textit{formal orthogonal polynomials (FOPs)}. 

These FOPs are the denominators of the \textit{convergents = approximants = “partial sums”} of a continued fraction.

These convergents are \textit{Padé approximants} (at \(\infty\)) associated with the moments \(s_{k+\nu} (k = 0, 1, \ldots; \ \nu \text{ fixed})\) of the function \(f_{\nu}(z)\) defined by

\[
f(z) = \frac{s_0}{z} + \frac{s_1}{z^2} + \cdots + \frac{s_{\nu-1}}{z^\nu} + \frac{f_{\nu}(z)}{z^\nu} . \quad (14)
\]

These FOPs are also the \textit{Lanczos polynomials} for the \textit{Lanczos BiO process} for \(A\) started with the pair \((y_0, A^\nu x_0)\).
A polynomial 3-term recurrence of the form \( p_k^{(\nu)} \) means that the polynomials \( p_k^{(\nu)} \) \((k = 0, 1, \ldots; \ \nu \text{ fixed})\) are (in today’s lingo) formal orthogonal polynomials (FOPs).

These FOPs are the denominators of the convergents = approximants = “partial sums” of a continued fraction.

These convergents are Padé approximants (at \( \infty \)) associated with the moments \( s_{k+\nu} \) \((k = 0, 1, \ldots; \ \nu \text{ fixed})\) of the function \( f_{\nu}(z) \) defined by

\[
f(z) = \frac{S_0}{z} + \frac{S_1}{z^2} + \cdots + \frac{S_{\nu-1}}{z^{\nu}} + \frac{f_{\nu}(z)}{z^{\nu}}.
\] (14)

These FOPs are also the Lanczos polynomials for the Lanczos BIO process for \( A \) started with the pair \((y_0, A^{\nu}x_0)\).
Rutishauser and Stiefel knew about orthogonal polynomials (also present in CG), continued fractions, and the Lanczos polynomials in the top diagonal.

Rutishauser derived the result on the Lanczos polynomials on the other diagonals.

Neither Rutishauser nor Stiefel ever seem to mention Padé approximants, but actually only the FOPs in the denominator matter.

Later, 1966–74, Householder, Gragg, and Stewart stress the connection to Padé approximants in several papers.

N.B.: Hadamard’s theorem (1892) $\sim$ de Montessus de Ballore’s theorem (1902/1905).
Continued fractions: $J$–fractions and $S$–fractions

\[
f_{\nu}(z) \equiv \sum_{k=0}^{\infty} \frac{s_{\nu+k}}{z^{k+1}} = z^{\nu} \left( f(z) - \sum_{k=0}^{\nu-1} \frac{s_k}{z^{k+1}} \right) \tag{15}
\]

can be expanded both into a **Jacobi fraction** or **J–fraction**

\[
f_{\nu}(z) = \frac{s_{\nu}}{z - q_{1}(\nu)} - \frac{e_{1}(\nu)q_{1}(\nu)}{z - q_{2}(\nu) - e_{1}(\nu)} - \frac{e_{2}(\nu)q_{2}(\nu)}{z - q_{3}(\nu) - e_{2}(\nu)} - \ldots \tag{16}
\]

and into a formal **Stieltjes fraction** or **S–fraction**

\[
f_{\nu}(z) = \frac{s_{\nu}}{z} - \frac{q_{1}(\nu)}{1} - \frac{e_{1}(\nu)}{z} - \frac{q_{2}(\nu)}{1} - \frac{e_{2}(\nu)}{z} - \ldots. \tag{17}
\]

The J–fraction is the so-called **even part** of the S–fraction obtained by merging two successive terms into one.
The odd part of the S–fraction is another formal J–fraction, obtained by merging the two differently chosen successive terms into one,

\[
f_\nu(z) = \frac{s_\nu}{z} \left\{ 1 + \frac{q^{(\nu)}_1}{z - q^{(\nu)}_1 - e^{(\nu)}_1} - \frac{e^{(\nu)}_1 q^{(\nu)}_2}{z - q^{(\nu)}_2 - e^{(\nu)}_2} - \frac{e^{(\nu)}_2 q^{(\nu)}_3}{z - q^{(\nu)}_3 - e^{(\nu)}_3} - \cdots \right\}.
\]

(18)

By comparing this J–fraction with the one for

\[
f_{\nu+1}(z) = zf_\nu(z) - s_\nu,
\]

(19)

one recovers Rutishauser’s rhombus rules of the qd algorithm.

This is the simplest derivation of the qd algorithm, but not the original one.

Rutishauser (1954a) indicates that it may have been suggested to him by Stiefel.
Further relations and applications

In addition to introducing and investigating the qd algorithm Rutishauser (1954a) [rec. 5 Aug. 1953] (1954b) [rec. 18 Sep. 1953], (1955a) [rec. 19 July 1954]) explained many connections to other topics and gave many applications; e.g., in (1954a):

- the connection to continued fractions,
- the connection to the Lanczos BÌO algorithm,
- the connection to the CG algorithm,
- computing partial fraction decompositions of rational fcts.

In (1954b):

- summation of badly converging series,
- solving algebraic equations = computing zeros of polynomials,
- quadratic convergence by using shifts / double shifts.
In (1955a):

- computing EVals by *combining Lanczos’ B&O alg. and the progressive qd algorithm*,
- computing EVecs (several new algorithms are suggested),
- EVals and EVecs of infinite matrices.

*Still missing:*

- *tridiagonal matrices* (except for computing shifts),
- *LU decomposition* of these tridiagonal matrices,
- *LR algorithm*.
Rutishauser knew well (see Rutishauser (1953) on the Lanczos B\textsc{io} algorithm) that associated to the \( 3 \)-term recurrence (??),

\[
p_{k+1}(z) := \left[z - q_{k+1} - e_k^{(\nu)}\right] p_k^{(\nu)}(z) + e_k^{(\nu)} q_k^{(\nu)} p_{k-1}^{(\nu)}(z)
\]

(with fixed \( \nu \)) there is a nested set of tridiagonal matrices

\[
T^{(\nu)}_n = \begin{pmatrix}
q_1^{(\nu)} & 1 & & & \\
e_1^{(\nu)} q_1^{(\nu)} & e_1^{(\nu)} + q_2^{(\nu)} & 1 & & \\
e_2^{(\nu)} q_2^{(\nu)} & e_2^{(\nu)} + q_3^{(\nu)} & \ddots & & \\
& \ddots & \ddots & 1 & \\
e_{n-1}^{(\nu)} q_{n-1}^{(\nu)} & e_{n-1}^{(\nu)} + q_n^{(\nu)} & & &
\end{pmatrix}
\]

such that \( p_n^{(\nu)}(z) \) is the characteristic polynomial of \( T^{(\nu)}_n \).

Since he was interested in the limit of the zeros of \( p_n^{(\nu)} \) as \( \nu \to \infty \) it was natural to look at \( T^{(\nu)}_n \). \textit{Though: what for?}
Clearly, $T_n^{(\nu)}$ has the \textit{LU decomposition (LR-Zerlegung)}

\[ T_n^{(\nu)} = L_n^{(\nu)} R_n^{(\nu)} \quad (20) \]

with

\[
L_n^{(\nu)} = \begin{pmatrix}
1 & & & \\
\nu & 1 & & \\
& \nu & \ddots & \\
& & \ddots & 1 \\
& & & \nu & 1
\end{pmatrix}, \quad
R_n^{(\nu)} = \begin{pmatrix}
\nu & 1 & & & \\
& \nu & 1 & & \\
& & \ddots & 1 & \\
& & & \ddots & 1 \\
& & & & \nu & 1
\end{pmatrix}.
\]
The LU transformation

At some historic moment in 1954, Rutishauser must have realized that his progressive qd algorithm \((\ldots)\) can be interpreted as computing this LU factorization \(T_{n}^{(\nu)} = L_{n}^{(\nu)} R_{n}^{(\nu)}\) and then forming

\[
R_{n}^{(\nu)} L_{n}^{(\nu)} = \begin{pmatrix}
  e_{1}^{(\nu)} + q_{1}^{(\nu)} & 1 \\
  e_{1}^{(\nu)} q_{2}^{(\nu)} & e_{2}^{(\nu)} + q_{2}^{(\nu)} & 1 \\
  e_{2}^{(\nu)} q_{3}^{(\nu)} & e_{3}^{(\nu)} + q_{3}^{(\nu)} & \ddots \\
  \vdots & \ddots & \ddots & 1 \\
  e_{n-1}^{(\nu)} q_{n}^{(\nu)} & q_{n}^{(\nu)}
\end{pmatrix}
\]

\[= T_{n}^{(\nu+1)}\]
So, the qd algorithm consists of performing the step

\[
T^{(\nu)}_n = L^{(\nu)}_n R^{(\nu)}_n \quad \rightsquigarrow \quad R^{(\nu)}_n L^{(\nu)}_n = T^{(\nu+1)}_n
\]

called **LR transformation**, which is a similarity transformation since

\[
T^{(\nu+1)}_n = R^{(\nu)}_n T^{(\nu)}_n \left( R^{(\nu)}_n \right)^{-1}.
\]

If we shift by half a step,

\[
L^{(\nu)}_n, R^{(\nu)}_n \quad \rightsquigarrow \quad R^{(\nu)}_n L^{(\nu)}_n = T^{(\nu+1)}_n = L^{(\nu+1)}_n R^{(\nu+1)}_n \quad \rightsquigarrow \quad L^{(\nu+1)}_n, R^{(\nu+1)}_n.
\]

Today we prefer this second view as we know that the eigenvalues are better determined by the pairs \((L^{(\nu)}_n, R^{(\nu)}_n)\) of bidiagonal matrices than by the tridiagonal matrices \(T^{(\nu)}_n\).
LR algorithm: succession of LR transformations (LR steps).

Convergence of $e_k^{(\nu)} \to 0$ ($k = 1, \ldots, n$) as $\nu \to \infty$ means:

Convergence of $L_n^{(\nu)}$ to diagonal matrix as $\nu \to \infty$,

Convergence of $T_n^{(\nu)}$ to upper bidiagonal matrix as $\nu \to \infty$,

The diagonals of $T_n^{(\nu)}$ and $R_n^{(\nu)}$ ultimately contain eigenvalues of $A$.

Generalization to full matrices is immediate, but unimportant.
The LR algorithm

**LR algorithm**: succession of LR transformations (LR steps).

Convergence of \( e_k^{(\nu)} \rightarrow 0 \) (\( k = 1, \ldots, n \)) as \( \nu \rightarrow \infty \) means:

- Convergence of \( L_n^{(\nu)} \) to diagonal matrix as \( \nu \rightarrow \infty \),
- Convergence of \( T_n^{(\nu)} \) to upper bidiagonal matrix as \( \nu \rightarrow \infty \),
- The diagonals of \( T_n^{(\nu)} \) and \( R_n^{(\nu)} \) ultimately contain eigenvalues of \( A \),
- Generalization to full matrices is immediate, but unimportant.
The first two publication on the LR algorithm were in French, two two-page notes in the *Comptes Rendues*: Rutishauser (1955e) [séance du 3 janvier 1955], Rutishauser/Bauer (1955) [séance du 25 avril 1955].


In 1957, Rutishauser included a 5-page appendix on the LR transformation in a booklet that compiled and updated most of his previous work on qd (Rutishauser, 1957a).
Rutishauser kept on publishing articles on the qd and LR algorithms and their applications.

In particular, he studied the *qd algorithm in finite precision arithmetic* and could prove its stability under certain assumptions.

Before his death he was working on a long manuscript that included the finite precision results and the *differential qd algorithm*, which was later rediscovered by Fernando and Parlett (1994). The finished parts were published as an appendix of the posthumous book on his lectures (1976/1990).
The discovery of the qd and the LR algorithms probably evolved in the following steps:

- Generalizing Aitken’s work $\leadsto$ qd table / algorithm.
- Considering the corresponding $p$–table (gen. Lanczos’s work) and finding the diagonal 3-term recurrence for this table.
- Making the connection to continued fractions and Lanczos polynomials (and as well to many other topics).
- Making the connection to tridiagonal matrices.
- Noticing their extremely simple LU decomposition.
- Noticing that

\[ \text{qd algorithm} = \text{LR algorithm for tridiagonal matrices} \]

- Generalizing the LR algorithm to full matrices.
Unless $A$ is spd or Hpd the LR algorithm may break down, because an LU decomposition (without pivoting) may not exist. Using shifts a symmetric $A$ can be turned into an spd matrix. But, the LR algorithm may be unstable for nonsymmetric $A$. Stability is gained by replacing LU decomposition by QR decomposition $\Rightarrow$ QR algorithm.

Both LR and QR conserve the symmetry and the band structure of $A$.

In symmetric (Hermitian) case: two steps of LR yield the same $R_n^{(2\nu)}$ as one step of QR.

Independently, it was also discovered by V.N. Kublanovskaya (1961).

Francis’ papers contain a full theory, including the double-shift for approaching complex pairs of EVals.
Thanks for listening and come to ...

6th International Congress on Industrial and Applied Mathematics

Zurich, Switzerland
16 - 20 July 2007

www.iciam07.ch