THE LANCZOS ALGORITHMS, CG, QD, AND A WHOLE CIRCLE OF IDEAS

M. H. Gutknecht
Seminar for Applied Mathematics, ETH Zurich
CH-8092 Zurich, Switzerland
mhg@ethz.ch

In their seminal 1952 paper on the conjugate gradient (CG) method Hestenes and Stiefel pointed out that their method, which is applicable to linear systems of equations with symmetric positive definite matrix only, is closely related to certain orthogonal polynomials, the corresponding Gauss quadrature formulas, certain continued fractions, and their convergents (or 'partial sums'), which are Padé approximants.

Around the same time, in 1950 and 1952, Cornelius Lanczos published two related articles, of which the second one introduced a precursor of the biconjugate gradient (BiCG) method, which generalizes CG to the case of a nonsymmetric system. Here, the residual polynomials are formal orthogonal polynomials only, but the connections to continued fractions and Padé approximants persist. The latter are diagonal ones of the function

$$F(\zeta) := \mathbf{y}_0^H(\zeta\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}_0 = \sum_{k=0}^{\infty} \frac{\mu_k}{\zeta^{k+1}}, \quad \text{where} \quad \mu_k := \mathbf{y}_0^H \mathbf{A}^{\mathbf{k}} \mathbf{x}_0 ,$$

that involves the resolvent of the matrix $\mathbf{A}$ and its moments μ_k with respect to the starting vectors $\mathbf{x}_0$ and $\mathbf{y}_0$. Moreover, there is a relation to the qd algorithm of Rutishauser (1954). The understanding of all these connections became probably the key to Rutishauser's discovery of the LR algorithm (1955, 1958), which was later enhanced by John G. F. Francis to the ubiquitous QR algorithm (1961/62).

But this is not yet the full circle of ideas. E.g., F can be viewed as transfer function of a single-input-single-output linear time-invariant system. Or the Lanczos process can be viewed as one operating on polynomials. It is then seen to be equivalent to the Stieltjes process and delivers an inverse symmetric LDU decomposition of the Hankel moment matrix $\mathbf{M} := (\mu_{k+\ell})$. So, essentially, the Lanczos process gives rise to a fast Hankel solver.