Some pioneers of extrapolation methods

Claude Brezinski

Laboratoire Paul Painlevé, UMR CNRS 8524, UFR de Mathématiques Pures et Appliquées, Université des Sciences et Technologies de Lille, 59655 - Villeneuve d'Ascq cedex, France.
Claude.Brezinski@univ-lille1.fr

Abstract. There are two extrapolation methods methods which are described in almost all numerical analysis books: Richardson's extrapolation method (which forms the basic ingredient for Romberg's method), and Aitken's Δ^2 process. In this paper, we consider the historical roots of these two procedures (in fact, the computation of π) with an emphasis on the pioneers of this domain of numerical analysis. Finally, we will discuss some more recent developments and applications.

Richardson's extrapolation method and Aitken's Δ^2 process are certainly the most well known methods for the acceleration of a slowly converging sequence. Both are based on the idea of extrapolation, and they have their historical roots in the computation of π.

We will first explain what extrapolation methods are, and how they lead to sequence transformations for accelerating the convergence. Then, we will present the history of Richardson's extrapolation method, of Romberg's method, and of Aitken's Δ^2 process, with an emphasis on the lives and the works of the pioneers of these topics.

The study of extrapolation methods and convergence acceleration algorithms now forms an important domain of numerical analysis having many applications; see [15, 24, 71, 77, 78, 80]. More details about its mathematical developments and its history could be found in [11, 13, 26, 34].

1 Interpolation, extrapolation, sequence transformations

Assume that the values of a function f are known at k distinct points x_i, that is

$$y_i = f(x_i), \quad i = 0, \ldots, k-1.$$

Choose a function F_k depending on k parameters $a_0, \ldots, a_{k-1}$, and belonging to some class of functions $\mathcal{F}_k$ (for example polynomials of degree $k-1$).

What is interpolation? Compute $a_0^e, \ldots, a_{k-1}^e$ solution of the system of equations (the meaning of the superscript $.^e$ will appear below)

$$F_k(a_0^e, \ldots, a_{k-1}^e, x_i) = y_i, \quad i = 0, \ldots, k-1. \tag{1.1}$$

THE BIRTH OF NUMERICAL ANALYSIS

Then, for any $\forall x \in I = [\min_i x_i, \max_i x_i]$, we say that f has been *interpolated* by $F_k \in \mathcal{F}_k$, and we have $F_k(a_0^e, \ldots, a_{k-1}^e, x) \simeq f(x)$. Moreover, if $f \in \mathcal{F}_k$, then, for all x, $F_k(a_0^e, \ldots, a_{k-1}^e, x) = f(x)$.

What is extrapolation? Now choose $x^e \notin I$, and compute

$$y^e = F_k(a_0^e, \ldots, a_{k-1}^e, x^e),$$

where the coefficients a_i^e are those computed as the solution of the system (1.1). The function f has been *extrapolated* by $F_k \in \mathcal{F}_k$ at the point x^e, and $y^e \simeq f(x^e)$. Again, if $f \in \mathcal{F}_k$, then $F_k(a_0^e, \ldots, a_{k-1}^e, x^e) = f(x^e)$.

What is a sequence transformation? Assume now, without restricting the generality, that we have an infinite decreasing sequence of points $x_0 > x_1 > x_2 > \cdots > x^*$ such that $\lim_{n\to\infty} x_n = x^*$. We set $y_i = f(x_i)$, $i = 0, 1, \ldots$, and we also assume that $\lim_{n\to\infty} y_n = \lim_{n\to\infty} f(x_n) = y^*$. For any fixed n, compute $a_0^{(n)}, \ldots, a_{k-1}^{(n)}$ solution of the system

$$F_k(a_0^{(n)}, \ldots, a_{k-1}^{(n)}, x_{n+i}) = y_{n+i}, \quad i = 0, \ldots, k-1.$$

Then, compute $F_k(a_0^{(n)}, \ldots, a_{k-1}^{(n)}, x^*)$. This value obviously depends on n and, for that reason, it will be denoted by z_n. Then, the sequence (y_n) has been transformed into the new sequence (z_n), and $T : (y_n) \longmapsto (z_n)$ is called a *sequence transformation*. As we can see, it's a kind of *moving* extrapolation or, if one prefers, it is a sequence of extrapolated values based on different points. Obviously, if $f \in \mathcal{F}_k$, then, for all n, $z_n = y^*$. Instead of a decreasing sequence (x_n), we can take an increasing one, for example $(x_n = n)$.

A sequence transformation can be defined without any reference to a function f, but only to a sequence (y_n). An important concept is the notion of *kernel* of a sequence transformation: it is the set $\mathcal{K}_T$ of sequences (y_n) such that, for all n, $z_n = y^*$. If (y_n) converges, y^* is its limit, otherwise it is called its *antilimit*.

For readers who are not familiar with the topics of interpolation, extrapolation, and sequence transformations, these notions will be explained again at the beginning of Section 3 via the construction of Aitken's Δ^2 process.

When are extrapolation and sequence transformations powerful? As explained above, extrapolation and sequence transformations are based on the choice of the the class of functions $\mathcal{F}_k$. Thus, if the function f to be extrapolated behaves like a function of $\mathcal{F}_k$, the extrapolated value y^e will be a good approximation of $f(x^e)$. Similarly, the sequence (z_n) will converge to y^* faster than the sequence (y_n), that is $\lim_{n\to\infty}(z_n - y^*)/(y_n - y^*) = 0$, if (y_n) is close, in a sense to be defined (an open problem), to $\mathcal{K}_T$.

For each sequence transformation, there are sufficient conditions so that (z_n) converges to the same limit as (y_n), but faster. It was proved that a universal sequence transformation able to accelerate the convergence of *all* converging

THE BIRTH OF NUMERICAL ANALYSIS

sequences cannot exist [25]. This is even true for some classes of sequences such as the monotonically decreasing ones. This negative result does not mean that a particular sequence belonging to such a class cannot be accelerated, but that an algorithm for accelerating all of them cannot exist.

2 Richardson's extrapolation

The procedure named after Richardson consists in extrapolation at $x^e = 0$ by a polynomial. It is carried out by means of the Neville-Aitken scheme for the recursive computation of interpolation polynomials.

2.1 First contributions

The number π can be approximated by considering polygons with n sides inscribed into and circumscribed to a circle. With $n = 96$, Archimedes obtained two significant figures. He also proved, by geometrical arguments, that the area of a circle is equal to $rp/2$, where r is its radius and p its perimeter.

In 1596, Adriaan van Roomen, also called Adrianus Romanus (Leuven, 1561 - Mainz, 1615), obtained 15 figures with $n = 2^{30}$ while, in 1610, Ludolph van Ceulen (Hildesheim, 1540 - Leiden, 1610), with the help of his student Pieter Cornelisz (Amsterdam, 1581 - The Hague, 1647), gave 36 figures by using $n = 2^{62}$. According to the Dutch mathematician David Bierens de Haan (1822 - 1895), these values were, in fact, obtained at about the same time. Van Ceulen's result was carved on his tombstone in the St. Peters church in Leiden. His work was continued in 1616 by Philips van Lansbergen (Ghent, 1561 - Middleburg, 1632) who held him in high esteem. He was a minister who published books on mathematics and astronomy where he supported Copernic's theories. However, he did not accept Kepler's theory of elliptic orbits. He suggested the approximation

$$2\pi \simeq p_{2n} + (p_{2n} - p_n)\frac{p_{2n} - 4}{p_n},$$

where p_n is the perimeter of the regular n-gons inscribed in the unit circle. He obtained π with 28 exacts decimal figures. The Dutch astronomer Willebrord Snel van Royen (Leiden, 1580 - Leiden, 1626), known as Snellius, was the introducer of the method of triangulation for measuring the length of the meridian. He also proposed, in 1621, the lower and upper bounds

$$\frac{3p_{2n}^2}{2p_{2n} + p_n} < 2\pi < \frac{p_{2n}(p_{2n} + 2p_n)}{3p_n}.$$

These formulae were preparing the ground for the next step.

2.2 C. Huygens

In 1654, Christiaan Huygens (The Hague, 1629 - id., 1695), in his *De Circuli Magnitudine Inventa* [33], proved 16 theorems or lemmas on the geometry of inscribed and circumscribed regular polygons. In particular, he gave the difference between the areas of the polygon with $2n$ sides and that with n sides. Let a_c and a_n be the areas of the circle and of the n-gons, respectively. He proved that (see [50] for an analysis of the proof)

$$\begin{aligned} a_c &= a_n + (a_{2n} - a_n) + (a_{4n} - a_{2n}) + (a_{8n} - a_{4n}) + \cdots \\ &> a_n + (a_{2n} - a_n) + (a_{2n} - a_n)/4 + (a_{2n} - a_n)/16 + \cdots \\ &= a_{2n} + (a_{2n} - a_n)(1/4 + 1/16 + \cdots). \end{aligned}$$

Then Huygens refers to Archimedes when stating that the sum of this geometric series is 1/3, thus leading to

$$a_c > a_{2n} + \frac{a_{2n} - a_n}{3}.$$

We have

$$a_n = \frac{1}{2} n \sin\left(\frac{2\pi}{n}\right) < \pi < A_n = n \tan\left(\frac{\pi}{n}\right),$$

where A_n is the area of the circumscribed n-gons. Series expansions of trigonometric functions were not available to Huygens. However

$$a_n = \pi - \frac{2\pi^3}{3n^2} + \frac{2\pi^5}{15n^4} - \frac{4\pi^7}{315n^6} + \cdots,$$

and so Huygens' lower bound is such that

$$a_{2n} + \frac{a_{2n} - a_n}{3} = \pi - \frac{8\pi^5}{15 \cdot 16n^4} + \frac{16\pi^7}{63 \cdot 64n^6} - \cdots$$

Similarly

$$A_n = \pi + \frac{\pi^3}{2n^2} + \frac{2\pi^5}{15n^4} + \frac{17\pi^7}{315n^6} + \cdots$$

and it follows

$$A_{2n} + \frac{A_{2n} - A_n}{3} = \pi - \frac{8\pi^5}{15 \cdot 16n^4} - \frac{68\pi^7}{63 \cdot 64n^6} - \cdots,$$

which is also a lower bound for π, but slightly poorer than the previous one.

Thus, Huygens' formulae for lower bounds are exactly those obtained by the first step of Richardson's extrapolation method. Moreover, in order to obtain an upper bound, he proposed

$$A_n - \frac{A_n - a_n}{3} = \pi + \frac{2\pi^5}{15n^4} + \cdots,$$

whose error is bigger than for the lower bounds but uses only polygons with n sides instead of $2n$. With $n = 2^{30}$, this last formula doubles (up to 35) the number of exact digits of π. In letters to Frans van Schooten (Leiden, 1615 - id., 1660) and Daniel Lipstorp (1631 - 1684), he claimed that he was able to triple the number of exact decimals. Therefore, Huygens achieved approximations of π which are even better than those given by Richardson's extrapolation!

Huygens' method was later used by Jacques Frédéric Saigey in 1856 and 1859 [61]. He considered the three approximations

$$A'_n = A_{2n} + \frac{1}{3}(A_{2n} - A_n)$$
$$A''_n = A'_{2n} + \frac{1}{15}(A'_{2n} - A'_n)$$
$$A'''_n = A''_{2n} + \frac{1}{63}(A''_{2n} - A''_n)$$

which are similar to those that will be given later by Romberg in the context of accelerating the convergence of the trapezoidal rule.

Saigey was born in Montbéliard in 1797. He studied at the École Normale Supérieure in Paris, but the school was closed in June 1822 by the regime of Louis XVIII. Saigey became the secretary of Victor Cousin and helped him to publish the volume V of Descartes' complete works. Then, he became one of the main editors of the journal *Bulletin des Sciences Mathématiques*. He published several papers in mathematics and physics, but he was mostly known for his elementary treatises and memoranda which had several editions. He died in Paris in 1871.

In 1903, Robert Moir Milne (1873 - ?) applied Huygens' ideas for computing π [43], as also did Karl Kommerell (1871 - 1948) in his book of 1936 [36]. As explained in [76], Kommerell can be considered as the real discoverer of Romberg's method since he suggested the repeated use of Richardson's rule, although it was in a different context.

2.3 L.F. Richardson

In 1910, Lewis Fry Richardson (1881 - 1953) suggested to eliminate the first error term in the central differences formulæ given by William Fleetwood Sheppard (Sydney, 1863 - 1936) [69] by using several values of the stepsize. He wrote [52]

> *... the errors of the integral and of any differential expressions derived from it, due to using the simple central differences of §1.1 instead of the differential coefficients, are of the form*
>
> $$h^2 f_2(x,y,z) + h^4 f_4(x,y,z) + h^6 f_6(x,y,z) + \&tc.$$
>
> *Consequently, if the equation be integrated for several different values of h, extrapolation on the supposition that the error is of this form will give numbers very close to the infinitesimal integral.*

In 1927, Richardson called this procedure the *deferred approach to the limit* [55]. Let us quote him

Confining attention to problems involving a single independent variable x, let h be the "step", that is to say, the difference of x which is used in the arithmetic, and let $\phi(x,h)$ be the solution of the problem in differences. Let $f(x)$ be the solution of the analogous problem in the infinitesimal calculus. It is $f(x)$ which we want to know, and $\phi(x,h)$ which is known for several values of h. A theory, published in 1910, but too brief and vague, has suggested that, if the differences are "centered" then

$$\phi(x,h) = f(x) + h^2 f_2(x) + h^4 f_4(x) + h^6 f_6(x) \ldots \textit{to infinity} \ldots \qquad (1)$$

odd powers of h being absent. The functions $f_2(x), f_4(x), f_6(x)$ are usually unknown. Numerous arithmetical examples have confirmed the absence of odd powers, and have shown that it is often easy to perform the arithmetic with several values of h so small that $f(x)+h^2 f_2(x)$ is a good approximation to the sum to infinity of the series in (1).
If generally true, this would be very useful, for it would mean that if we have found two solutions for unequal steps h_1, h_2, then by eliminating $f_2(x)$ we would obtain the desired $f(x)$ in the form

$$f(x) = \frac{h_2^2 \phi(x,h_1) - h_1^2 \phi(x,h_2)}{h_2^2 - h_1^2}. \qquad (2)$$

This process represented by the formula (2) will be named the "h^2-extrapolation".
If the difference problem has been solved for three unequal values of h it is possible to write three equations of the type (1) for h_1, h_2, h_3, retaining the term $h^4 f_4(x)$. Then $f(x)$ is found by eliminating both $f_2(x)$ and $f_4(x)$. This process will be named the "h^4-extrapolation".

Let us mention that Richardson referred to a paper by Nikolai Nikolaevich Bogolyubov (Nijni-Novgorod, 1909 - Moscow, 1992) and Nikolai Mitrofanovich Krylov (Saint Petersburg, 1879 - Moscow, 1955) of 1926 where the deferred approach to the limit can already be found [7].

In the same paper, Richardson used this technique for solving a 6th order differential eigenvalue problem. Richardson extrapolation consists in fact in computing the value at 0, denoted by $T_k^{(n)}$, of the interpolation polynomial of the degree at most k which passes through the points $(x_n, S(x_n)), \ldots, (x_{n+k}, S(x_{n+k}))$. Thus, using the Neville-Aitken scheme for these interpolation polynomials, the numbers $T_k^{(n)}$ can be recursively computed by the formula

$$T_{k+1}^{(n)} = \frac{x_{n+k+1} T_k^{(n)} - x_n T_k^{(n+1)}}{x_{n+k+1} - x_n}, \qquad k, n = 0, 1, \ldots \qquad (2.1)$$

with $T_0^{(n)} = S(x_n)$ for $n = 0, 1, \ldots$.

Extensions of the Richardson extrapolation process are reviewed in [15, 71], and many applications are discussed in [39].

Lewis Fry (the maiden name of his mother) Richardson was born on October 11, 1881 in Newcastle upon Tyne, England, the youngest of seven children in a Quaker family. He early showed an independent mind and had an empirical approach. In 1898, he entered the Durham College of Science where he took courses in mathematics, physics, chemistry, botany, and zoology. Then, in 1900, he went to King's College in Cambridge, and followed the physics lectures of Joseph John Thompson (Cheetham Hill near Manchester, 1856 - Cambridge, 1940), the discoverer of the electron. He graduated with a first-class degree in 1903. He spent the next ten years holding a series of positions in various academic and industrial laboratories. When serving as a chemist at the National Peat Industry Ltd., he had to study the percolation of water. The process was described by the Laplace equation on an irregular domain and Richardson used finite differences, and extrapolation. But it was only after much deliberation and correspondence that his paper was accepted for publication [52]. He submitted this work for a D.Sc. and a fellowship at Cambridge, but it was rejected. The ideas were too new, and the mathematics were considered as "approximate mathematics"! Hence, Richardson never worked in any of the main academic research centers. This isolation probably affected him. For some time, he worked with the well-known statistician Karl Pearson (London, 1857 - Coldharbour, 1936), and became to be interested in "living things".

In 1913, Richardson became Superintendent of the Eskdalemuir Observatory in southern Scotland. He had no experience in meteorology, but was appointed to bring some theory in its understanding. He again made use of finite differences. Although he was certainly aware of the difficulty of the problem since he estimated at 64.000 the number of people that have to be involved in the computations in order to obtain the prediction of tomorrow's weather before day actually began, it seems that he did not realize that the problem was ill-conditioned. He also began to write his book on this topic [53]. The quote at the end of its preface is amusing.

> *This investigation grew out of a study of finite differences and fist took shape in 1911 as the fantasy which is now relegated to Chap. 11/2. Serious attention to the problem was begun in 1913 at Eskdalemuir Observatory with the permission and encouragement of Sir Napier Shaw, then Director of the Meteorological Office, to whom I am greatly indebted for facilities, informations and ideas. I wish to thank Mr. W.H. Dines, F.R.S., for his interest in some early arithmetical experiments, and Dr. Crichton Mitchell, F.R.S.E., for some criticisms of the first draft. The arithmetical reduction of the balloon, and other observations, was done with much help from my wife. In May 1916 the manuscript was communicated by Sir Napier Shaw to the Royal Society, which generously voted £100 towards to cost of its publication. The manuscript was revised and the detailed example of Chap. IX was worked out in France in the intervals of transporting wounded in 1916-1918. During the battle of Champagne in April 1917 the working copy was sent to the rear, where it became lost, to be re-discovered some months later under a heap of*

coal. In 1919, as printing was delayed by the legacy of the war, various excrescences were removed for separate publication, and an introductory example was added. This was done at Benson, where I had again the good fortune to be able to discuss the hypotheses with Mr. W.H. Dines. The whole work has been thoroughly revised in 1920, 1921. As the cost of printing had by this time much increased, an application was made to Dr. G.C. Simpson, F.R.S., for a further grant in aid, and the sum of fifty pounds was provided by the Meteorological Office.

As Richardson wrote, on May 16, 1916 he resigned and joined the Friends' Ambulance Unit (a Quaker organisation) in France. He began to think about the causes of wars and how to prevent them. He suggested that the animosity between two countries could be measured, and that some differential equations are involved into the process. He published a book on these ideas [54], and then returned to weather prediction.

Along the years, Richardson made important contributions to fluid dynamics, in particular to eddy-diffusion in the atmosphere. The so-called "Richardson number" is a fundamental quantity involving gradients of temperature and wind velocity.

In 1920, he became a Lecturer in mathematics and physics at the Westminster Training College, an institution training prospective school teachers up to a bachelor's degree. In 1926, he changed again his field of research to psychology where he wanted to apply the ideas and the methods of mathematics and physics. He established that many sensations are quantifiable, he found methods for measuring them, and modelled them by equations. The same year, he was elected as a Fellow of the Royal Society of London.

Richardson left the Westminster Training College in 1929 for the position of Principal at the Technical College in Paisley, an industrial city near Glasgow. Although he had to teach sixteen hours a week, he continued his research but came back to the study of the causes of wars and their prevention. He prepared a model for the tendencies of nations to prepare for wars, and worked out its applications using historical data from the previous conflicts. He also made predictions for 1935, and showed that the situation was unstable, which could only be prevented by a change in the nation's policies. Richardson wanted to "see whether there is any statistical connection between war, riot and murder". He began to accumulate such data [57], and decided to search for a relation between the probability of two countries going to war and the length of their common border. To his surprise, the lengths of the borders were varying from one source to another. Therefore, he investigated how to measure the length of a border, and he realized that it highly depends on the length of the ruler. Using a small ruler allows to follow more wiggles, more irregularities, than a long one which cuts the details. Thus, the smaller the ruler, the larger the result. The relation between the length of the border and that of the ruler leads to a new mathematical measure of wiggliness. At that time, Richardson's results were ignored by the scientific community, and they were only published posthumously [58]. Today, they are considered to be at the origin of fractals.

In 1943, Richardson and his wife moved to their last home at Kilmun, 25 miles from Glasgow. He returned to his research on differential equations, and solved the associated system of linear equations by the so-called Richardson's method [56]. He mentioned that the idea was suggested to him in 1948 by Arnold Lubin. At home, Richardson was also constructing an analogous computer for his meteorological computations. He died on September 30, 1953 in Kilmun.

Richardson was a very original character whose contributions to many different fields were prominent but, unfortunately, not appreciated at their real values at his epoch; see [32] for details.

2.4 W. Romberg

Let us now come to the procedures for improving the accuracy of the trapezoidal rule for computing approximations to a definite integral. If the function to be integrated is sufficiently differentiable, the error of the trapezoidal rule is given by the Euler-Maclaurin expansion. In 1742, Colin Maclaurin (Kilmodan, 1698 - Edinburgh, 1746) [38] showed that the precision could be improved by linear combinations of the results obtained with various stepsizes. His procedure can be interpreted as a preliminary version of Romberg's method; see [21] for a discussion.

In 1900, Sheppard used an elimination strategy in the Euler-Maclaurin quadrature formula, with $h_n = r_n h$ and $1 = r_0 < r_1 < r_2 < \cdots$, for producing a better approximation [70]. In 1952, Mario Salvadori (Rome, 1907 - 1997), an architect and structural engineer, and Melvin L. Baron (1927 - 1997), a civil engineer, proposed to use Richardson's deferred approach to the limit for improving the trapezoidal rule [62]. This new approximation was obtained as a linear combination of the initial results.

In 1955, Werner Romberg was the first to use repeatedly an elimination approach for improving the accuracy of the trapezoidal rule [59]. He gave the well known formula

$$T_{k+1}^{(n)} = \frac{4^{k+1} T_k^{(n+1)} - T_k^{(n)}}{4^{k+1} - 1},$$

where $T_0^{(n)}$ is the result obtained by the trapezoidal rule with the stepsize $h_0/2^n$. In his paper, Romberg refers to the book of Lothar Collatz (Arnsberg, Westfalia, 1910 - Varna, 1990) of 1951 [22].

In 1960, Eduard L. Stiefel (1909 - 1978), in his inaugural address as the President of the IFIP congress in Munich, draws a line from Archimedes to Romberg. The procedure became widely known after the rigorous error analysis given in 1961 by Friedrich L. Bauer (born 1924 in Regensburg) [6] and the synthesis of Stiefel [75]. Romberg's derivation of his process was mainly heuristic. It was proved by Pierre-Jean Laurent in 1963 [37] that the process comes out, in fact, from the Richardson process when taking $x_n = h_n^2$ and $h_n = h_0/2^n$. Laurent also gave the condition on the sequence (h_n) that there exists $\alpha < 1$ such that $\forall n, h_{n+1}/h_n \le \alpha$ in order that the sequences $(T_k^{(n)})$ tend to the exact value of the definite integral to be computed either when k or n tends to infinity. The case of a harmonic sequence of steps is studied in [23, p. 52]. Romberg's work on the extrapolation of the trapezoidal rule has been continued Tore Håvie for less regular integrands [28].

Werner Romberg was born on May 16, 1909 in Berlin. In 1928, he started to study physics and mathematics in Heidelberg where the Nobel laureate Philip Lenard (Pozsony, Pressburg, 1862 - Messelhausen, 1947) was still quite influential. After two years, Romberg decided to go to the Ludwig-Maximilians University in Munich. He followed the mathematics courses of Constantin Carathéodory (Berlin, 1873 - Munich, 1950) and Oskar Perron (Frankenthal, Pfalz, 1880 - Munich, 1975), and had physics lectures by Arnold Sommerfeld (Königsberg, 1868 - Munich, 1951), who became his advisor. In 1933, he defended his thesis *Zur Polarisation des Kanalstrahllichtes* (On the polarization of canal jet rays). The same year, he had to leave Germany and went to the USSR. He stayed at the Department of Physics and Technology in Dnepropetrovsk from 1934 to 1937 as a theoretical physicist. He was briefly at the Institute of Astrophysics in Prag in 1938, but he had to escape from there. Then, he got a position in Oslo in the autumn of 1938 as the assistant of the physicist Egil Andersen Hylleraas (Engerdal, 1898 - 1965). He also worked for a short period with Johan Holtsmark (1894 - 1975), who built a Van de Graaff generator (the second one in Europe and the first particle accelerator in Scandinavia) for nuclear disintegration between 1933 and 1937 at Norwegian Institute of Technology (NTH) in Trondheim. Romberg had again to escape for some time to Uppsala during the German occupation of Norway. In 1949, he joined the NTH in Trondheim as an associate professor in physics. In 1960, he was appointed head of the Applied Mathematics Department at the NTH. He organized a teaching program in applied mathematics, and began to build a research group in numerical analysis. He was strongly involved in the introduction of digital computers in Norway, and in the installation of the first computer (GIER) at NTH. He became a Norwegian citizen and stayed Norwegian until the end of his life.

In 1968, Romberg came back to Heidelberg where he accepted a professorship. He built up a group in numerical mathematics, at that time quite underdeveloped in Heidelberg, and was the head of the Computing Center of the University from 1969 to 1975. Romberg retired in 1978, and died on February 5, 2003.

3 Aitken's process and Steffensen's method

Let (S_n) be a sequence of scalars converging to S. The most popular nonlinear acceleration method is certainly Aitken's Δ^2 process which consists in building a new sequence (T_n) by

$$T_n = \frac{S_n S_{n+2} - S_{n+1}^2}{S_{n+2} - 2S_{n+1} + S_n}, \quad n = 0, 1, \ldots \tag{3.1}$$

For deriving this formula, Aitken assumed that he had a sequence (S_n) of the form

$$S_n = S + \alpha\lambda^n, \quad n = 0, 1, \ldots \tag{3.2}$$

with $\lambda \neq 1$, and he wanted to compute S (the limit of the sequence if $|\lambda| < 1$, its antilimit otherwise). Then, $\Delta S_n = \alpha\lambda^n(\lambda - 1)$, and $\lambda = \Delta S_{n+1}/\Delta S_n$. It follows

$$S = S_n - \frac{\Delta S_n}{(1-\lambda)} = S_n - \frac{\Delta S_n}{(1 - \Delta S_{n+1}/\Delta S_n)} = \frac{S_n S_{n+2} - S_{n+1}^2}{S_{n+2} - 2S_{n+1} + S_n}.$$

If (S_n) has not the form (3.2), the preceding formula can still be used, but the result is no longer equal to S. It depends on n, and it is denoted by T_n as in (3.1). This construction of Aitken's process illustrates how interpolation, extrapolation, and sequence transformations are related. Indeed, let (S_n) be any sequence. We are looking for S, α and λ satisfying the interpolation conditions $S_i = S + \alpha\lambda^i$ for $i = n, n+1, n+2$. Then, the unknown S is taken as the limit when n tends to infinity of the model sequence $(S + \alpha\lambda^n)$. This is an extrapolation process. But, since the value of S obtained in this procedure depends of n, it has been denoted by (S_n), and, thus, the given sequence (S_n) has been transformed into the new sequence (T_n).

Thus, by construction, the kernel of Aitken's process consists in sequences of the form (3.2), or, in other terms, of sequences satisfying a first order linear difference equation

$$a_0(S_n - S) + a_1(S_{n+1} - S) = 0, \quad n = 0, 1, \ldots$$

with $a_0 + a_1 \neq 0$.

If (S_n) is linearly converging, i.e. if a number $\lambda \neq 1$ exists such that

$$\lim_{n\to\infty} \frac{S_{n+1} - S}{S_n - S} = \lambda,$$

then (T_n) converges to S faster than (S_n). This result illustrates the fact mentioned above that sequences *not too far away from the kernel* (in a meaning to be defined) are accelerated. Acceleration is also obtained for some subclasses of sequences satisfying the preceding property with $\lambda = 1$ (logarithmically converging sequences).

In a paper of 1937 [2], Aitken used his process for accelerating the convergence of the power method (Rayleigh quotients) for computing the dominant eigenvalue of a matrix. A section is entitled *The δ^2-process for accelerating convergence*, and, on pages 291-292, he wrote

> *For practical computation it may be remembered by the following* memoria technica: product of outers minors *[minus]* square of middle, divided by sum of outers minus double of middle.

Aitken's paper [2] also contains almost all the ideas that will be developed later by Heinz Rutishauser (Weinfelden, 1918 - 1970) in his QD-algorithm [60].

Notice that Formula (3.1) is numerically unstable, and that one should prefer the following one

$$T_n = S_{n+1} + \frac{(S_{n+1} - S_n)(S_{n+2} - S_{n+1})}{(S_{n+1} - S_n) - (S_{n+2} - S_{n+1})}. \tag{3.3}$$

It is well known that the fixed point iterative method due to Johan Frederik Steffensen (1873-1961) in 1933 is based on Aitken's process. However, Steffensen does not quote Aitken in his paper, and his discovery seems to have been obtained independently. Consider the computation of x such that $x = f(x)$ and the iterations $x_{\nu+1} = f(x_\nu)$. Steffensen writes [74]

In the linear interpolation formula with divided differences

$$f(x) = f(a_0) + (x - a_0)f(a_0, a_1) + (x - a_0)(x - a_1)f(x, a_0, a_1) \quad (5)$$

we put $a_\nu = x_\nu$ *and obtain*

$$f(x) = x_1 + (x - x_0)\frac{x_1 - x_2}{x_0 - x_1} + R_1$$

where

$$R_1 = (x - x_0)(x - x_1)f(x, x_0, x_1). \quad (6)$$

Replacing, on the left of (6), $f(x)$ *by* x*, we have*

$$x = x_1 + (x - x_0)\frac{x_1 - x_2}{x_0 - x_1} + R_1$$

and solving for x*, as if* R_1 *were a constant, we obtain after a simple reduction*

$$x = x_0 - \frac{(\Delta x_0)^2}{\Delta^2 x_0} + R \quad (7)$$

where

$$R = -(x - x_0)(x - x_1)\frac{\Delta x_0}{\Delta^2 x_0} f(x, x_0, x_1). \quad (8)$$

If $f(x)$ *possesses a continuous second derivative, the remainder may be written*

$$R = -\frac{1}{2}(x - x_0)(x - x_1)\frac{\Delta x_0}{\Delta^2 x_0} f''(\xi). \quad (9)$$

The formula (7) is the desired result. The approximation obtained may often be estimated by (9), but we shall make no use of this formula, preferring to test the result by other methods. We shall therefore use as *working formula the approximation*

$$x = x_\nu - \frac{(\Delta x_\nu)^2}{\Delta^2 x_\nu} \tag{10}$$

where, according to the remark made above, x_ν may be any element of the sequence.

Then, Steffensen gave several numerical examples where, after 3 iterations, he restarted them from the approximation given by (10). In a footnote to page 64 (the first page of his paper), he wrote

> *The present notes had already been written when a paper by H. Holme appeared in this journal (1932, pp. 225-250), covering to some extend the same ground. Mr. Holme's treatment of the subject differs, however, so much from mine that I think there is room for both.*

In his paper [31], Harald Holme was solving a fixed point problem due to Birger Øivind Meidell (1882-1958) and related to the interest rate of loans [41]. He used linear interpolation passing through 3 consecutive iterates, and he obtained a method quite close to Steffensen's but different from it.

3.1 Seki Takakazu

In the fourth volume of his book *Katsuyō Sanpō*, published in 1674, Seki Takakazu considered the perimeters c_i of the polygons with 2^i sides inscribed into a circle of diameter 1. For deriving a better approximation of π, he used a method called *Yenri*, which means *principle* (or *theory*) *of the circle*, and consists in the formula

$$c_{16} + \frac{(c_{16} - c_{15})(c_{17} - c_{16})}{(c_{16} - c_{15}) - (c_{17} - c_{16})}.$$

This is exactly Aitken's Δ^2 process (as given by (3.3)) which leads to 12 exact decimal digits while c_{17} has only 10. With

$$\begin{aligned} c_{15} &= \underline{3.14159264}87769856708 \\ c_{16} &= \underline{3.141592652}3565913571 \\ c_{17} &= \underline{3.1415926532}889027755, \end{aligned}$$

Seki obtained $\underline{3.14159265359}$ ($\pi = 3,14159265358979323846\ldots$). His result is, in fact, exact to 16 places. Seki did not explain how he got his formula but, probably, setting $a = c_{15}, b = c_{16} = a + ar$, and $c = c_{17} = a + ar + ar^2$, he obtained [30]

$$b + \frac{(b-a)(c-b)}{(b-a)-(c-b)} = \frac{a}{1-r} = a + ar + ar^2 + ar^3 + \cdots$$

The same method was used by his student Takaaki Takebe (1661 - 1716), who developed it further. Seki also studied how to compute an arc of a circle, given the chord, and he used again his formula for improving his first approximations.

Seki Takakazu is considered as the greatest Japanese mathematician. He was born in Fujioka in 1637 or in 1642. He was later adopted by the Seki family. However, little is known about his life, but it seems that he was self-educated and an infant prodigy in mathematics. In his book mentioned above, he introduced a notation for representing unknowns and variables in equations, and he solved fifteen problems which had been posed three years earlier by Kazuyuki Sawaguchi (it was the habit to end a book by open problems). He anticipated many discoveries of western mathematicians: determinants (1683, ten years before Leibniz) for solving systems of 2 or 3 linear equations, Bernoulli numbers, Newton-Raphson method, and Newton interpolation formula. He studied the solution of equations with negative and positive zeros, and, in 1685, he solved the cubic equation $30 + 14x - 5x^2 - x^3 = 0$ by the same method as Horner a hundred years later. He was also interested in magic squares, and Diophantine equations. He died in 1708.

3.2 A.C. Aitken

The sequence transformation defined by (3.1) was stated by Alexander Craig Aitken in 1926 [1] who used it for accelerating the convergence of Daniel Bernoulli's method of 1728 for the computation of the dominant zero z_1 of the polynomial $a_0 z^n + \cdots + a_{n-1} z + a_n$. The method imagined by Bernoulli consists in considering the sequence $Z_1(t) = f(t+1)/f(t)$ generated from the recursion $a_0 f(t+n) + \cdots + a_n f(t) = 0$, and whose limit is z_1 (assuming that all other zeros of the polynomial have a modulus strictly smaller than $|z_1|$). With this condition, Aitken writes

> *$\Delta Z_1(t)$ tends to become a geometric sequence... of common ratio z_2/z_1. Hence the derivations of $Z_1(t)$ from z_1 will also tend to become a geometric sequence with the same common ratio. Thus a further approximate solution is suggested, viz.*
>
> $$\frac{z_1 - Z_1(t+2)}{z_1 - Z_1(t+1)} = \frac{\Delta Z_1(t+1)}{\Delta Z_1(t)}$$
>
> *and solving for z_1 we are led to investigate the derived sequence*
>
> $$Z_1^{(1)}(t) = \frac{\begin{vmatrix} Z_1(t+1) & Z_1(t+2) \\ Z_1(t) & Z_1(t+1) \end{vmatrix}}{\Delta^2 Z(t)}. \qquad (8.2)$$

This is exactly (3.1). Aitken claims that this new sequence converges geometrically with the ratio $(z_2/z_1)^2$ or z_3/z_1, and that the process can be repeated on the sequence $(Z_1^{(1)}(t)$. In a footnote, he says that *Naegelsbach, in the course of a very detailed investigation of Fürstenau method of solving equations, obtains the formulæ (8.2) and (8.4), but only incidentally.* The reference for the work of Eduard Fürstenau is [27]. It must be pointed out, that, on page 22 of his second paper [44], Hans von Naegelsbach (1838 - ?) gave the stable formulation (3.3) of the process.

The process was also given by James Clerk Maxwell (Edinburgh, 1831 - Cambridge, 1879) in his *Treatise on Electricity and Magnetism* of 1873 [40]. However, neither Naegelsbach nor Maxwell used it for the purpose of acceleration. Maxwell wanted to find the equilibrium position of a pointer oscillating with an exponentially damped simple harmonic motion from three experimental measurements (as in (3.2)).

Aitken was born in Dunedin, New Zealand, on April 1st, 1895. He attended Otago's High School from 1908 to 1912, where he was not particularly brillant. But, at the age of 15, he realized that he had a real power in mental calculations, and that his memory was extraordinary. He was able to recite the first 1000 decimals of π, and to multiply two numbers of nine digits in a few seconds [72]. He also knew the Aeneid by heart. He was also very good at several sports and began to study violin. He studied mathematics, French and Latin at the University of Otago in 1913 and 1914. It seems that the professor of mathematics there, David J. Richards, a "temperamental, eccentric Welshman", was lacking of the power to communicate his knowledge to the students, and Aitken's interest in mathematics lowered. Richards was trained as an engineer as well as mathematician, and was working as an engineer in Newcastle prior to his appointment to the Chair of Mathematics at Otago in 1907, where he stayed until 1917.

Aitken volunteered in the Otago infantry during World War I, and he took part in the Gallipoli landing and in the campaign in Egypt. Then, he was commissioned in the north of France, and was wounded in the shoulder and foot during the battle on the river Somme. Did he met Richardson at this time? After a stay in a London hospital, he was invalided home in 1917, and spent one year of recovering in Dunedin where he wrote a first account of his memoirs published later [4].

Aitken resumed his studies at Otago University, and graduated with first class honours in languages, but only with second ones in mathematics. He married Winifred Betts in 1920, and became a school teacher at his old Otago High School. Richards' successor in the Chair of Mathematics, Robert John Tainsh Bell was born in 1877. He graduated from the University of Glasgow in 1898, and was appointed Lecturer there three years later. He was awarded a D.Sc. in 1911, and was appointed Professor of Pure and Applied Mathematics at Otago University in 1919. Bell was the only staff member in the Mathematics Department, lecturing five days a week, each day from 8.00 am to 1.00 pm. He retired in 1948, and died in 1963. When Bell required an assistant he called on Aitken. He encouraged him to apply for a scholarship for studying with Edmund Taylor Whittaker (Southport, 1873 - Edinburgh, 1956) at Edinburgh. Aitken left New Zealand in 1923. His Ph.D. on the smoothing of data, completed in 1925, was considered so outstanding that he was awarded a D.Sc. for it. The same year, Aitken was appointed as a Lecturer at the University of Edinburgh where he stayed for the rest of his life. But, the efforts for obtaining his degree led him to a first severe breakdown in 1927, and then he was periodically affected by such crisis. They were certainly in part due to his fantastic memory which did not fade with time, and he was always remembering the horrors he saw during the war [5] (see also the biographical introduction by Peter C. Fenton given in this volume).

In 1936, Aitken became a Reader in statistics, and he was elected a Fellow of the Royal Society. In 1946, he was appointed to Whittaker's Chair in Mathematics. In 1956, he received the prestigious Gunning Victoria Jubilee Prize of the Royal Society of Edinburgh. In 1964, he was elected to the Royal Society of Literature. Aitken died in Edinburgh on November 3, 1967.

3.3 J.F. Steffensen

Since the life of Steffensen is not so well-known, let us give some informations about it following [49]. Johan Frederik Steffensen was born in Copenhagen on February 28, 1873. His father was the Supreme Judge of the Danish Army, and he, himself, took a degree in law at the University of Copenhagen. After a short period in Fredericia in the eastern part of the Jutland peninsula in Denmark, he returned to Copenhagen and began a career in insurance. He was self-taught in mathematics and, in 1912, he got a Ph.D. for a study in number theory. After three years as the managing director of a mutual life assurance society, he turned to teach insurance mathematics at the University of Copenhagen, first as a Lecturer and, from 1923 to 1943, as a Professor. However, he was still continuing to be interested in the world of affairs, and was an active member, and even the Chairman, of several societies. He published around 100 research papers in various fields of mathematics, and his book of 1927 [73] can be considered as one of the first books in numerical analysis since its chapters cover interpolation in one and several variables, numerical derivation, solution of differential equations, and quadrature. Steffensen loved English literature, especially Shakespeare. He died on December 20, 1961. For a photography of Steffensen, see [47].

3.4 D. Shanks

The idea of generalizing Aitken's process is due to Daniel Shanks. He wanted to construct a sequence transformation with a kernel consisting of sequences satisfying, for all n,

$$a_0(S_n - S) + a_1(S_{n+1} - S) + \cdots + a_k(S_{n+k} - S) = 0, \tag{3.4}$$

with $a_0 + a_1 + \cdots + a_k \neq 0$. Let us mention that a particular case of an arbitrary value of k was already studied by Thomas H. O'Beirne in 1947 [48]. Writing the relation (3.4) for the indexes $n, n+1, \ldots, n+k$ leads to

$$\begin{vmatrix} S_n - S & S_{n+1} - S & \cdots & S_{n+k} - S \\ S_{n+1} - S & S_{n+2} - S & \cdots & S_{n+k+1} - S \\ \vdots & \vdots & & \vdots \\ S_{n+k} - S & S_{n+k+1} - S & \cdots & S_{n+2k} - S \end{vmatrix} = 0.$$

After elementary manipulations on the rows and columns of this determinant, Shanks obtained

$$S = H_{k+1}(S_n)/H_k(\Delta^2 S_n), \tag{3.5}$$

where $\Delta^2 S_n = S_{n+2} - 2S_{n+1} + S_n$, and where H_k denotes a Hankel determinant defined as

$$H_k(u_n) = \begin{vmatrix} u_n & u_{n+1} \cdots & u_{n+k-1} \\ u_{n+1} & u_{n+2} \cdots & u_{n+k} \\ \vdots & \vdots & \vdots \\ u_{n+k-1} & u_{n+k} \cdots & u_{n+2k-2} \end{vmatrix}.$$

If (S_n) does not satisfy the relation (3.4), the ratio of determinants in the right hand side of (3.5) could nevertheless be computed but, in this case, the result obtained depends on n, and it is denoted by $e_k(S_n)$. Thus, the sequence (S_n) has been transformed into the new sequence $(e_k(S_n))$ for a fixed value of k or, more generally, into a set of new sequences depending on k and n. This sequence transformation is known as *Shanks' transformation*. Let us also mention that the same ratio of determinants was obtained by R.J. Schmidt in 1941 [63] while studying a method for solving systems of linear equations.

Dan Shanks was born on January 17, 1917 in Chicago. In 1937, he received a B.Sc. in physics. From 1941 to 1957, he was employed by the Naval Ordnance Laboratory (NOL) located in White Oak, Maryland. There, in 1949, he published a Memorandum describing his transformation [65]. Without having done any graduate work, he wanted to present this work to the Department of Mathematics of the University of Maryland as a Ph.D. thesis. But, he had first to complete the degree requirements before his work could be examined as a thesis. Hence, it was only in 1954 that he obtained his Ph.D. which was published in the *Journal of Mathematical Physics* [66]. Dan considered this paper as one of his best two (the second one was his computation of π to 100.000 decimals published with John Wrench [67]). After the NOL, Shanks worked at the David Taylor Model Basin in Bethesda where I met him in December 1976. Then, in 1977, he joined the University of Maryland where he stayed until his death on September 6, 1996. Dan served as an editor of *Mathematics of Computation* from 1959 until his death. He was very influential in this position which also led him to turn to number theory, a domain where his book became a classic [68]. More details on Shanks life and works can be found in [79].

3.5 P. Wynn

The application of Shanks' transformation to a sequence (S_n) needs the computation of the ratios of Hankel determinants given by (3.5). The numerators and the denominators in this formula can be computed separately by the well-known recurrence relation for Hankel determinants (a by-product of Sylvester's determinantal identity). This was the way O'Beirne and Shanks were implementing the transformation. However, in 1956, Peter Wynn (born in 1932) found a recursive algorithm for that purpose, the ε-algorithm [81], whose rules are

$$\varepsilon_{k+1}^{(n)} = \varepsilon_{k-1}^{(n+1)} + \frac{1}{\varepsilon_k^{(n+1)} - \varepsilon_k^{(n)}}, \quad k, n = 0, 1, \ldots$$

with $\varepsilon_{-1}^{(n)} = 0$ and $\varepsilon_0^{(n)} = S_n$, for $n = 0, 1, \ldots$.

These quantities are related to Shanks' transformation by

$$\varepsilon_{2k}^{(n)} = e_k(S_n),$$

and the quantities with an odd lower index satisfy $\varepsilon_{2k+1}^{(n)} = 1/e_k(\Delta S_n)$. When $k = 1$, Aitken's process is recovered. The proof makes use of Schweins' and Sylvester's determinantal identities that could be found, for example, in Aitken's small monograph [3].

Later, Wynn became Bauer's assistant in Mainz, then he went to Amsterdam, participating in the birth of ALGOL, and then he held several researcher's positions in the United States, Canada, and Mexico. Wynn's ε-algorithm is certainly the most important and well-known nonlinear acceleration procedure used so far. Wynn dedicated many papers to the properties and the applications of his ε-algorithm. With a vector generalization of it [82], he also opened the way to special techniques for accelerating the convergence of sequences of vectors. The ε-algorithm also provides a derivative free extension of Steffensen's method for the solution of systems of nonlinear equations [8] (see also [15]).

Let us mention the important connection between the ε-algorithm and Padé approximants (and, thus, also with continued fractions). Let f be a formal power series

$$f(x) = \sum_{i=0}^{\infty} c_i x_i.$$

If the ε-algorithm is applied to its partial sums, that is $S_n = \varepsilon_0^{(n)} = \sum_{i=0}^{n} c_i x_i$, then $\varepsilon_{2k}^{(n)} = [n + k/k]_f(x)$, the Padé approximant of f with a numerator of degree $n + k$ and a denominator of degree k, a property exhibited by Shanks [66]. This connection allowed Wynn to obtain a new relation, known as the *cross rule*, between 5 adjacent approximants in the Padé table [83]. However, the ε-algorithm and the cross rule give the values of the Padé approximants only at the point x where the partial sums S_n were computed, while knowing the coefficients of the numerators and the denominators of the Padé approximants allows to compute them at any point.

4 And now?

In the last twenty years, Richardson's and Romberg's methods, Aitken's process and the ε-algorithm have been extended to more general kernels, or to accelerate new classes of sequences. Very general extrapolation algorithms have been obtained; see, for example, [15, 71]. In particular, the E-algorithm, whose rules obviously extend those of the Richardson process, was devised almost simultaneously by different people in different contexts [9, 29, 42, 64]. These procedures are now used in many physical applications [78, 20]. An important new field of investigation is the connection between some convergence acceleration algorithms and integrable systems, Toda lattices, the KdV equation, and solitons [45, 46, 51].

For the improvement of certain numerical techniques, it is often worth to construct special extrapolation procedures built on the analysis of the process to be accelerated (that is to construct extrapolations methods whose sequences in the kernel mimic as closely as possible the exact behavior of the sequence to be accelerated). For example, this was the methodology recently followed for Tikhonov regularization techniques [18], estimations of the error for systems of linear equations [12], treatment of the Gibbs phenomenon in Fourier and other series [14], and ranking in web search [35, 19, 16, 17].

Acknowledgments: I would like to thank Prof. Douglas Rogers and Prof. Haakon Waadeland for providing me informations on the life of W. Romberg. I am grateful to Prof. Peter C. Fenton for biographical informations on R.J.T. Bell and D.J. Richards. I am indebted to Michela Redivo-Zaglia for her efficient help in web searching. Finally, I acknowledge the clarifying comments of the referee (in particular her/his remarks on the interpolation/extrapolation interpretation of Aitken's process, and on Bierens de Haan, van Ceulen, van Roomen, and Seki) which greatly helped to improve parts of the paper.

References

1. A.C. Aitken, On Bernoulli's numerical solution of algebraic equations, *Proc. R. Soc. Edinb.*, **46** (1926) 289–305.
2. A.C. Aitken, Studies in practical mathematics. II. The evaluation of the latent roots and latent vectors of a matrix, *Proc. R. Soc. Edinb.*, **57** (1937) 269–304.
3. A.C. Aitken, *Determinants and Matrices*, Oliver and Boyd, Edinburgh and London, 1939.
4. A.C. Aitken, *Gallipoli to the Somme. Recollections of a New Zealand Infantryman*, Oxford University Press, London, 1963.
5. A.C. Aitken, *To Catch the Spirit. The Memoir of A.C. Aitken with a Biographical Introduction by P.C. Fenton*, University of Otago Press, Dunedin, 1995.
6. F.L. Bauer, La méthode d'intégration numérique de Romberg, in *Colloque sur l'Analyse Numérique*, Librairie Universitaire, Louvain, 1961, pp.119–129.
7. N. Bogolyubov, N. Krylov, On Rayleigh's principle in the theory of differential equations of mathematical physics and upon Euler's method in the calculus of variation (in Russian), *Acad. Sci. Ukraine (Phys. Math.)*, **3** (1926) 3–22.
8. C. Brezinski, Application de l'ε-algorithme à la résolution des systèmes non linéaires, *C.R. Acad. Sci. Paris*, **271** A (1970) 1174–1177.
9. C. Brezinski, A general extrapolation algorithm, *Numer. Math.*, **35** (1980) 175–187.
10. C. Brezinski, *History of Continued Fractions and Padé Approximants*, Springer, Berlin, 1991.

11. C. Brezinski, Extrapolation algorithms and Padé approximations: a historical survey, *Appl. Numer. Math.*, **20** (1996) 299–318.
12. C. Brezinski, Error estimates for the solution of linear systems., *SIAM J. Sci. Comput.*, **21** (1999) 764–781.
13. C. Brezinski, Convergence acceleration during the 20th century, *J. Comput. Appl. Math.*, **122** (2000) 1–21.
14. C. Brezinski, Extrapolation algorithms for filtering series of functions, and treating the Gibbs phenomenon, *Numer. Algorithms*, **36** (2004) 309–329.
15. C. Brezinski, M. Redivo-Zaglia, *Extrapolation Methods. Theory and Practice*, North-Holland, Amsterdam, 1991.
16. C. Brezinski, M. Redivo-Zaglia, The PageRank vector: properties, computation, approximation, and acceleration, *SIAM J. Matrix Anal. Appl.*, **28** (2006) 551–575.
17. C. Brezinski, M. Redivo-Zaglia, Rational extrapolation for the PageRank vector, *Math. Comp.*, **77** (2008) 1585–1598.
18. C. Brezinski, M. Redivo-Zaglia, G. Rodriguez, S. Seatzu, Extrapolation techniques for ill–conditioned linear systems, *Numer. Math.*, **81** (1998) 1–29.
19. C. Brezinski, M. Redivo-Zaglia, S. Serra-Capizzano, Extrapolation methods for PageRank computations, *C.R. Acad. Sci. Paris, Sér. I*, **340** (2005) 393–397.
20. E. Caliceti, M. Meyer-Hermann, P. Ribeca, A. Surzhykov, U.D. Jentschura, From useful algorithms for slowly convergent series to physical predictions based on divergent perturbative expansions, *Pysics Reports*, **446**, Nos. 1–3, 2007.
21. J.-L. Chabert et al., *A History of Algorithms. From the Pebble to the Microchip*, Springer, Berlin, 1999.
22. L. Collatz, *Numerische Behandlung von Differentialgleichungen*, Springer–Verlag, Berlin, 1951.
23. M. Crouzeix, A. Mignot, *Analyse Numérique des Équations Différentielles*, 2nd ed., Masson, Paris, 1989.
24. J.P. Delahaye, *Sequence Transformations*, Springer–Verlag, Berlin, 1988.
25. J.P. Delahaye, B. Germain-Bonne, Résultats négatifs en accélération de la convergence, *Numer. Math.*, **35** (1980) 443–457.
26. J. Dutka, Richardson extrapolation and Romberg integration, *Historia Math.*, **11** (1984) 3–21.
27. E. Fürstenau, *Darstellung der reellen Würzeln algebraischer Gleichungen durch Determinanten der Coefficienten*, Marburg, 1860.
28. T. Håvie, On a modification of Romberg's algorithm, *BIT*, **6** (1966) 24–30.
29. T. Håvie, Generalized Neville type extrapolation schemes, *BIT*, **19** (1979) 204–213.
30. A. Hirayama, K. Shimodaira, H. Hirose, *Takakazu Seki's Collected Works Edited with Explanations*, Osaka Kyoiku Tosho, Osaka, 1974.
31. H. Holme, Beitrag zur Berechnung des effektiven Zinsfusses bei Anleihen, Skand. *Aktuarietidskr.*, **15** (1932) 225–250.
32. J.C.R. Hunt, Lewis Fry Richardson and his contributions to mathematics, meteorology, and models of conflicts, *Annu. Rev. Fluid Mech.*, **30** (1998) xiii–xxxvi.
33. C. Huygens, *De Circuli Magnitudine Inventa*, Johannem et Danielem Elzevier, Amsterdam, 1654 and *Œuvres*, Société Hollandaise des Sciences, Martinus Nijhoff, 1910, vol. 12, pp. 91–181.
34. D.C. Joyce, Survey of extrapolation processes in numerical analysis, *SIAM Rev.*, **13** (1971) 435–490.
35. S.D. Kamvar, T.H. Haveliwala, C.D. Manning, G.H. Golub, Extrapolations methods for accelerating PageRank computations, in *Proceedings of the 12th International World Wide Web Conference*, ACM Press, New York, 2003, pp. 261-270.

36. K. Kommerell, *Das Grenzgebiet der Elementaren und Höheren Mathematik*, Verlag Köhler, Leipzig, 1936.
37. P.J. Laurent, Un théorème de convergence pour le procédé d'extrapolation de Richardson, *C. R. Acad. Sci. Paris*, **256** (1963) 1435–1437.
38. C. Maclaurin, *Treatise of Fluxions*, Ruddimans, Edinburgh, 1742.
39. G.I. Marchuk, V.V. Shaidurov, *Difference Methods and their Extrapolations*, Springer–Verlag, New York, 1983.
40. J.C. Maxwell, *A Treatise on Electricity and Magnetism*, Oxford University Press, Oxford, 1873.
41. B. Meidell, Betrachtungen über den effektiven Zinsfuss bei Anleihen, *Skand. Aktuarietidskr.*, **15** (1932) 159–174.
42. G. Meinardus, G.D. Taylor, Lower estimates for the error of the best uniform approximation, *J. Approx. Theory*, **16** (1976) 150–161.
43. R.M. Milne, Extension of Huygens' approximation to a circular arc, *Math. Gaz.*,**2** (1903) 309–311.
44. H. Naegelsbach, Studien zu Fürstenau's neuer Methode der Darstellung und Berechnung der Wurzeln algebraischer Gleichungen durch Determinanten der Coefficienten, *Arch. d. Math. u. Phys.*, **59** (1876) 147–192; **61** (1877) 19–85.
45. A. Nagai, J. Satsuma, Discrete soliton equations and convergence acceleration algorithms, *Phys. Letters A*, **209** (1995) 305–312.
46. A. Nagai, T. Tokihiro, J. Satsuma, The Toda molecule equation and the ε–algorithm, *Math. Comp.*, **67** (1998) 1565–1575.
47. N.E. Nørdlund, Johan Frederik Steffensen in memoriam, *Nordisk Matematisk Tidsskrift*, **10** (1962) 105–107.
48. T.H. O'Beirne, On linear iterative processes and on methods of improving the convergence of certain types of iterated sequences, Technical report, Torpedo Experimental Establishment, Greenock, May 1947.
49. M.E. Ogborn, Johan Frederik Steffensen, *J. Inst. Actuaries*, **88** (1962) 251–252.
50. O. Østerby, Archimedes, Huygens, Richardson and Romberg, unpublished manuscript; see `www.daimi.au.dk/~oleby/notes/arch.ps`
51. V. Papageorgiou, B. Grammaticos, A. Ramani, Integrable lattices and convergence acceleration algorithms, *Phys. Letters, A*, **179** (1993) 111–115.
52. L.F. Richardson, The approximate arithmetical solution by finite difference of physical problems involving differential equations, with an application to the stress in a masonry dam, *Philos. Trans. Roy. Soc. London, ser. A*, **210** (1910) 307–357.
53. L.F. Richardson, *Weather Prediction by Numerical Process*, Cambridge University Press, Cambridge, 1922; second edition, 2007.
54. L.F. Richardson, *Mathematical Psychology of War*, Hunt, Oxford, 1919.
55. L.F. Richardson, The deferred approach to the limit. Part I: Single lattice, *Philos. Trans. Roy. Soc. London, ser. A*, **226** (1927) 299–349.
56. L.F. Richardson, A purification method for computing the latent columns of numerical matrices and some integrals of differential equations, *Philos. Trans. Roy. Soc. London, Ser. A*, **242** (1950) 439–491.
57. L.F. Richardson, *Statistics of Deadly Quarrels*, Boxwood Press, Pacific Grove, 1960.
58. L.F. Richardson, The problem of contiguity: an appendix of statisticfs of deadly quarrels, *Gen. Syst. Yearb.*, **6** (1961) 139–187.
59. W. Romberg, Vereinfachte numerische Integration, *Kgl. Norske Vid. Selsk. Forsk.*, **28** (1955) 30–36.
60. H. Rutishauser, *Der Quotienten-Differenzen Algorithmus*, Birkhäuser-Verlag, Basel, 1957.

61. J.F. Saigey, *Problèmes d'Arithmétique et Exercices de Calcul du Second Degré avec les Solutions Raisonnées*, Hachette, Paris, 1859.
62. M.G. Salvadori, M.L. Baron, *Numerical Methods in Engineering*, Prentice–Hall, Englewood Cliffs, 1952.
63. R.J. Schmidt, On the numerical solution of linear simultaneous equations by an iterative method, *Phil. Mag.*, **7** (1941) 369–383.
64. C. Schneider, Vereinfachte Rekursionen zur Richardson–Extrapolation in Spezialfällen, *Numer. Math.*, **24** (1975) 177–184.
65. D. Shanks, An analogy between transient and mathematical sequences and some nonlinear sequence-to-sequence transforms suggested by it. Part I, Memorandum 9994, Naval Ordnance Laboratory, White Oak, July 1949.
66. D. Shanks, Non linear transformations of divergent and slowly convergent sequences, *J. Math. and Phys.*, **34** (1955) 1–42.
67. D. Shanks, J.W. Wrench, Jr., Calculation of π to 100.000 decimals, *Math. Comp.*, **16** (1962) 76–99.
68. D. Shanks, *Solved and Unsolved Problems in Number Theory*, Spartan Books, Washington, 1962.
69. W.F. Sheppard, Central difference formulæ, *Proc. London Math. Soc.*, **31** (1899) 449–488.
70. W.F. Sheppard, Some quadrature formulas, *Proc. London Math. Soc.*, **32** (1900) 258–277.
71. A. Sidi, *Practical Extrapolation Methods: Theory and Applications*, Cambridge University Press, Cambridge, 2003.
72. S.B. Smith, *The Great Mental Calculators. The Psychology, Methods, and Lives of Calculating Prodigies Past and Present*, Columbia University Press, New York, 1983.
73. J.F. Steffensen, *Interpolation*, Williams and Wilkins, Baltimore, 1927. Reprinted by Chelsea, New York, 1950.
74. J.F. Steffensen, Remarks on iteration, *Skand. Aktuarietidskr.*, **16** (1933) 64–72.
75. E. Stiefel, Altes und neues über numerische Quadratur, *Z. Angew. Math. Mech*, **41** (1961) 408–413.
76. G. Walz, The history of extrapolation methods in numerical analysis, Report Nr. 130, Universität Mannheim, Fakultät für Mathematik und Informatik, 1991.
77. G. Walz, *Asymptotics and Extrapolation*, Akademie Verlag, Berlin, 1996.
78. E.J. Weniger, Nonlinear sequence transformations for the acceleration of convergence and the summation of divergent series, *Comput. Physics Reports*, **10** (1989) 189–371.
79. H.C. Williams, Daniel Shanks (1917-1996), *Math. Comp.*, **66** (1997) 929–934 and *Notices of the AMS*, **44** (1997) 813–816.
80. J. Wimp, *Sequence Transformations and their Applications*, Academic Press, New York, 1981.
81. P. Wynn, On a device for computing the $e_m(S_n)$ transformation, *MTAC*, **10** (1956) 91–96.
82. P. Wynn, Acceleration techniques for iterated vector and matrix problems, *Math. Comp.*, **16** (1962) 301–322.
83. P. Wynn, Upon systems of recursions which obtain among the quotients of the Padé table, *Numer. Math.*, **8** (1966) 264–269.