Rend. Sem. Mat. Univ. Politec. Torino
Vol. 70, 4 (2012), 347 – 368

W. Gautschi

INTERPOLATION BEFORE AND AFTER LAGRANGE

Abstract. A brief history of interpolation is given and selected aspects thereof, centered around Lagrange's contribution to the subject.

1. Introduction

The idea and practice of interpolation has a long history going back to antiquity and extending to modern times. We will briefly sketch the early development of the subject in ancient times and the middle ages through the 17th century, culminating in the work of Newton. We next draw attention to a little-known paper of Waring predating Lagrange's interpolation formula by 16 years. The rest of the paper deals with a few selected contributions made after Lagrange till recent times. They include the computationally more attractive barycentric form of Lagrange's formula, the theory of error and convergence based on real-variable and complex-variable analyses, Hermite and Hermite-Fejér as well as nonpolynomial interpolation. Applications to numerical quadrature and the solution of ordinary and partial differential equations are briefly indicated.

As seems appropriate for this auspicious occasion, however, we begin with Lagrange himself.

2. The Lagrange interpolation formula

In the words of Lagrange [67, v. 7, p. 535]: "La méthode des interpolations est une des plus ingénieuses et des plus utiles que l'Astronomie possède," echoing what Newton wrote to Oldenburg in 1676 about the problem of interpolation [76, p. 137]: "...it ranks among the most beautiful of all that I could wish to solve." Both citations indicate the high regard in which the problem of interpolation was held by scholars of the 17th and 18th century, and the former the principal area — astronomy — in which it was applied. Of course, times have changed, but the utility aspect of interpolation is as valid today as it was three hundred years ago.

There are two memoirs of Lagrange, [47, 48], in which he deals with special questions of interpolation, polynomial as well as trigonometric, with equally spaced values of the independent variable. They involve differences of one sort or another, in the tradition of Newtonian mathematics. What today is called the "Lagrange interpolation formula" is not dealt with in any of Lagrange's memoirs, but appears in the fifth lecture of his "Leçons élémentaires sur les mathématiques" delivered in 1795 at the École Normale, which were published in two issues of the *Séances des Écoles Normales*, year III (1794–1795) and reprinted in 1812, at the request of Lagrange, in the *Journal de l'École Polytechnique*; see also [67, v. 7, 271–287]. In this fifth lecture

"Sur l'usage des courbes dans la solution des problèmes," Lagrange discusses two geometric problems and uses analysis to solve them. He then continues with: "En effet, tout se réduit à d'écrire ou faire passer une courbe par plusieurs points, soit que ces points soient donnés par le calcul ou par une construction, ou même par des observations ou des expériences isolées et indépendantes les unes des autres. Ce Problème est, à la vérité, indéterminé; car on peut, à la rigueur, faire passer par des points donnés une infinité de courbes différentes, régulières ou irrégulières, c'est-à-dire soumises à des équations, ou tracées arbitrairement à la main; mais il ne s'agit pas de trouver des solutions quelconques, mais les plus simples et les plus aisées à employer."

Taking for these "simplest solutions" polynomials of appropriate degree, Lagrange then goes on to derive Newton's form of the interpolation polynomial. Thus, he sets up a table of (what we now call) divided differences, which he writes as

$$\begin{array}{cc|ccccc} x & y & & & & & \\ \hline p & P & & & & & \\ q & Q & Q_1 & & & & \\ r & R & R_1 & R_2 & & & \\ s & S & S_1 & S_2 & S_3 & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

where

$$Q_1 = \frac{Q-P}{q-p}, \quad R_1 = \frac{R-Q}{r-q}, \quad S_1 = \frac{S-R}{s-r}, \quad \dots$$

and[1]

$$R_2 = \frac{R_1-Q_1}{r-p}, \quad S_2 = \frac{S_1-R_1}{s-q}, \quad S_3 = \frac{S_2-R_2}{s-p},$$

and so on. (Note Lagrange's elegant and uncomplicated notation; the quantities indexed by ν are the νth divided differences, a name not used by Newton but apparently first introduced by de Morgan [58, Ch.XVII, p. 550] in 1842; see [54, footnote 15 on p. 322].) From this table of divided differences, Lagrange obtains, "après les réductions", Newton's formula,

$$y = P + Q_1(x-p) + R_2(x-p)(x-q) + S_3(x-p)(x-q)(x-r) + \cdots, \tag{1}$$

"qu'il est aisé de continuer aussi loin qu'on voudra." Thus, the diagonal entries in the table of divided differences serve as coefficients in the factorial expansion (1) of the interpolation polynomial.

Lagrange then continues with: "Mais on peut réduire cette solution à une plus grande simplicité par la considération suivante." And, using elementary algebraic manipulations, he arrives at the formula now bearing his name,

$$y = AP + BQ + CR + DS + \cdots, \tag{2}$$

[1]The third equation, for S_3, added by the author, does not appear in Lagrange's text.

where

$$A = \frac{(x-q)(x-r)(x-s)\cdots}{(p-q)(p-r)(p-s)\cdots},$$
$$B = \frac{(x-p)(x-r)(x-s)\cdots}{(q-p)(q-r)(q-s)\cdots},$$
$$C = \frac{(x-p)(x-q)(x-s)\cdots}{(r-p)(r-q)(r-s)\cdots},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,$$

"en prenant autant de facteurs, dans les numérateurs et dans les dénominateurs, qu'il y aura de points donnés de la courbe, moins un." Lagrange then says about his formula (2) that "elle est préférable par la simplicité de l'Analyse sur laquelle elle est fondée, et par sa forme même, qui est beaucoup plus commode pour le calcul." (Here, "calcul" is probably intended to mean analytic, not necessarily numerical, calculation.) Notice that the first function A has the value 1 at the first point p and vanishes at all the other points. Similarly for the other functions. This is a characteristic property of these coefficient functions, known as the elementary Lagrange interpolation polynomials.

Interestingly, Lagrange then suggests — and this is motivated by the two geometrical problems discussed previously — what is now called "inverse interpolation", i.e., reversing the roles of $p,q,r,s,\ldots$ and $P,Q,R,S,\ldots$ and finding an x for which y has a given value, specifically $y = 0$.

3. The period before Lagrange

3.1. The pre-Newtonian period

Instances of interpolation, especially linear interpolation, can be traced back to ancient Babylonian and Greek times, i.e., several centuries BC and the first centuries AD. They relate to astronomical observations and trigonometric tables used in astronomy. Evidences of higher-order interpolation appear in the early medieval ephemerides of China and India, and also in Arabic and Persian sources. In the Western countries, interpolation theory developed in the late 16th century quite independently from the much earlier work elsewhere, some of the early contributors being Harriot, Napier, Bürgi, and Briggs, of logarithms fame. See [54, Sec. II] for more details.

3.2. Newton

It was in the 17th century when in the Western world, especially England and France, the theory of interpolation began to flourish, driven not only by the exigencies of astronomy, but also by the need for preparing and using logarithmic and other tables. Initially, most attention was given to interpolation at equally spaced points, which gave rise to the calculus of finite differences. Notable contributors include Gregory, Newton, Stirling, and many others. For historical accounts, see, e.g., Joffe [42], Fraser [24], Turnbull [75], and Goldstine [35]. The culmination of this development, however, was

Newton's general formula for interpolation at arbitrary (distinct) points $x_0, x_1, x_2, \dots$, which in modern notation can be written in the form

$$f(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)(x-x_1) + \cdots + a_n(x-x_0)(x-x_1) \cdots (x-x_{n-1}) + R_n(f;x), \tag{3}$$

where

$$a_k = [x_0, x_1, \dots, x_k]f, \quad k = 0, 1, 2, \dots, n, \tag{4}$$

are the kth-order divided differences of f with respect to the interpolation abscissae $x_0, x_1, \dots, x_k$ and $R_n(f;x)$ is the remainder term. Neither Newton nor Lagrange said much about the remainder. The formula (3) (without the remainder) is contained in Lemma V, Book III, of Newton's *Principia* [59, p. 499], occupying not much more than half a page. In a corollary at the end of the lemma, Newton proposes integrating the polynomial part in (3) to approximate a definite integral of f; cf. §4.5.

It may be amusing to record here the curious use of Newton's formula made by Euler in a letter of 1734 to Daniel Bernoulli. Euler sought to compute the common logarithm $\log x$ for some x between 1 and 10 by applying (3) with $x_k = 10^k$, $\log x_k = k$, $k = 0, 1, 2, \dots$. (Only Euler would dare to try such a thing!) He calculates the divided differences to be

$$a_k = [x_0, x_1, \dots, x_k]\log = \frac{(-1)^{k-1}}{10^{k(k-1)/2}(10^k - 1)}, \quad k = 1, 2, 3, \dots,$$

and thus, from Newton's formula extended to infinity, obtains

$$S(x) = \sum_{k=1}^{\infty} a_k (x-1)(x-10)\cdots(x-10^{k-1}). \tag{5}$$

Amazingly, the series converges quite fast but, alas, not to $\log x$. Indeed, for $x = 9$, Euler finds $S(9) = .897778586588\dots$ instead of the correct value $\log 9 = .954242509439\dots$. In writing this to his friend Daniel Bernoulli, Euler presumably hoped to get from him some sensible explanation of this phenomenon. Unfortunately, Bernoulli's reply has not survived. It is unlikely, though, that he would have been able to shed some light on the matter. Analysis, particularly complex analysis, just hasn't yet sufficiently advanced to provide an answer. What happens, in a nutshell, is that the fast convergence in (5) renders $S(x)$ to be an entire function and therefore incapable of representing the logarithm; cf. [32].

Once a subject has caught Euler's fancy, he usually never abandoned it, but returned to it time and again. In this case, however, it took almost twenty years until (in [20]; see also [19, ser. 1, v. 14, 516–541]) he revisited the function $S(x)$, now using logarithms to an arbitrary base, and developed many interesting identites, among them two special cases of what today is known as the q-binomial theorem. The function studied by Euler has recently been suggested as a way to define a q-analogue of the logarithm [45].

3.3. Waring

Although this may not be the most opportune occasion, it must be said, in the interest of historical accuracy, that Waring [81], a Lucasian Professor at Cambridge University, predates Lagrange by 16 years.

In the cited paper, Waring starts out by saying that "Mr. Briggs was the first person, I believe, that invented a method of differences for interpolating logarithms at small intervals from each other: his principles were followed by Reginald[2] and Mouton in France. Sir Isaac Newton, from the same principles, discovered a general and elegant solution of the abovementioned problem: perhaps a still more elegant one on some accounts has been since discovered by Mess. Nichole and Stirling. In the following theorems the same problem is resolved and rendered somewhat more general, without having any recourse to finding the successive differences." He then formulates and proves a first theorem that reads as follows (in his original wording and notation): "Assume an equation $a+bx+cx^2+dx^3\ldots.x^{n-1}=y$, in which the coefficients a,b,c,d,e, &c. are invariable; let $\alpha,\beta,\gamma,\delta,\varepsilon$, &c. denote n values of the unknown quantity x, whose correspondent values of y let be represented by $s^\alpha,s^\beta,s^\gamma,s^\delta,s^\varepsilon$, &c. Then will be the equation $a+bx+cx^2+dx^3+ex^4\ldots.x^{n-1}=y=$

$$\begin{aligned}
&\frac{\overline{x-\beta}\times\overline{x-\gamma}\times\overline{x-\delta}\times\overline{x-\varepsilon}\times\&c.}{\alpha-\beta\times\alpha-\gamma\times\alpha-\delta\times\alpha-\varepsilon\times\&c.}\times s^\alpha+\frac{\overline{x-\alpha}\times\overline{x-\gamma}\times\overline{x-\delta}\times\overline{x-\varepsilon}\times\&c.}{\beta-\alpha\times\beta-\gamma\times\beta-\delta\times\beta-\varepsilon\times\&c.}\times s^\beta\\
&+\frac{\overline{x-\alpha}\times\overline{x-\beta}\times\overline{x-\delta}\times\overline{x-\varepsilon}\times\&c.}{\gamma-\alpha\times\gamma-\beta\times\gamma-\delta\times\gamma-\varepsilon\times\&c.}\times s^\gamma+\frac{\overline{x-\alpha}\times\overline{x-\beta}\times\overline{x-\gamma}\times\overline{x-\varepsilon}\times\&c.}{\delta-\alpha\times\delta-\beta\times\delta-\gamma\times\delta-\varepsilon\times\&c.}\times s^\delta\\
&+\frac{\overline{x-\alpha}\times\overline{x-\beta}\times\overline{x-\gamma}\times\overline{x-\delta}\times\&c.}{\varepsilon-\alpha\times\varepsilon-\beta\times\varepsilon-\gamma\times\varepsilon-\delta\times\&c.}\times s^\varepsilon+\&c.''
\end{aligned}$$

Clearly, this is the formula of Lagrange. Waring then states a second theorem that can be derived from the one just formulated.

4. The period after Lagrange

4.1. The barycentric form of Lagrange's formula

The original Lagrange interpolation formula, written in modern form, is

$$f(x)=p_n(f;x)+R_n(f;x),\quad p_n(f;x)=\sum_{k=0}^{n}f(x_k)\ell_k(x), \tag{6}$$

where

$$\ell_k(x)=\prod_{\substack{j=0\\ j\neq k}}^{n}\frac{x-x_j}{x_k-x_j},\quad k=0,1,\ldots,n, \tag{7}$$

are the elementary Lagrange interpolation polynomials (cf. §2). It is attractive more for theoretical than for computational purposes. Only relatively recently, in the mid to

[2]Probably François Regnaud; see [48, p. 664] (author's note).

late 1940s ([74, 17]), it was given another form — the barycentric form — which, as Rutishauser [64, §6.2] put it, "is more palatable for numerical computation". Indeed, its many advantages and extraordinary robustness have been emphazised recently by Berrut and Trefethen [5] and Higham [40].

Having in mind the situation where one is given an infinite sequence $x_0, x_1, x_2, \ldots$ of points and seeks to successively, for $n = 1, 2, 3, \ldots$, interpolate at the first $n+1$ points $x_0, x_1, \ldots, x_n$, we introduce a triangular array of auxiliary quantities defined by

$$\lambda_0^{(0)} = 1, \quad \lambda_k^{(n)} = \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{1}{x_k - x_j}, \quad k = 0, 1, \ldots, n; \; n = 1, 2, 3, \ldots .$$

Then (7) can be written in the form

$$\ell_k(x) = \frac{\lambda_k^{(n)}}{x - x_k} w_n(x), \quad k = 0, 1, \ldots, n; \quad w_n(x) = \prod_{j=0}^{n} (x - x_j). \tag{8}$$

Dividing $p_n(f;x)$ in (6) through by $1 \equiv \sum_{k=0}^{n} \ell_k(x)$ and using (8), one finds

$$p_n(f;x) = \frac{\displaystyle\sum_{k=0}^{n} \frac{\lambda_k^{(n)}}{x - x_k} f_k}{\displaystyle\sum_{k=0}^{n} \frac{\lambda_k^{(n)}}{x - x_k}}, \quad x \neq x_k \text{ for } k = 0, 1, \ldots, n. \tag{9}$$

This expresses the interpolation polynomial as a "weighted" average of function values $f_k = f(x_k)$ and is therefore called the barycentric formula. Note, however, that the "weights" are not necessarily all positive.

Comparison with (6), incidentally, reveals that

$$\ell_k(x) = \frac{\dfrac{\lambda_k^{(n)}}{x - x_k}}{\displaystyle\sum_{j=0}^{n} \frac{\lambda_j^{(n)}}{x - x_j}}, \quad k = 0, 1, \ldots, n. \tag{10}$$

It may be noted that in both formulae, (9) and (10), any factor of $\lambda_k^{(n)}$ that does not depend on k can be ignored since it cancels out in the numerator and denominator of (9) resp. (10). This can be used to advantage, for example, when $x_k = \cos(k\pi/n)$, and $\lambda_k^{(n)} = (-1)^k 2^{n-1}/n$, if $0 < k < n$, and half this value otherwise [66], so that $\lambda_k^{(n)}$ can be replaced by $(-1)^k$ resp. $\frac{1}{2}(-1)^k$.

4.2. Error and convergence

With regard to his interpolation process, Lagrange writes [67, v. 7, p. 284]: "Or il est clair que, quelle que puisse être la courbe proposée [the given function (or curve)], la

courbe parabolique [the polynomial curve] ainsi tracée en différera toujours d'autant moins que le nombre des points donnés sera plus grand, et leur distance moindre," suggesting that the *n*th-degree interpolation polynomial converges to the correct value of the given function as n goes to infinity and the largest distance between consecutive interpolation points goes to zero. Yet, convergence is by no means guaranteed, as will be shown further on by a remarkable example of Runge.

Real-variable analysis

A serious study of the error of interpolation has been taken up only about 25 years after Lagrange's death. Cauchy [11], in 1840, investigates divided differences (called "fonctions interpolaires", following Ampère), particularly the effect of confluences, $[\underbrace{x_0,x_0,\ldots,x_0}_{n+1 \text{ times}}]f = f^{(n)}(x_0)/n!$. He then obtains the mean-value formula $[x_0,x_1,...,x_n]f = f^{(n)}(\bar{\xi})/n!$ with $\bar{\xi} \in \text{span}(x_0,x_1,\ldots,x_n)$, and the remainder term in Newton's formula (3), first in the form $[x_0,x_1,\ldots,x_n,x]f \cdot \prod_{k=0}^{n}(x-x_k)$, and then, using his mean-value formula for divided differences, in the form generally known today,

$$R_n(f;x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)(x-x_1)\cdots(x-x_n) \tag{11}$$

with $\xi \in \text{span}(x_0,x_1,\ldots,x_n,x)$. We note that (11) reduces to Lagrange's form of the remainder term in Taylor's expansion [67, v. 9, p. 84], which is the special case of (3) when $x_0 = x_1 = \cdots = x_n$ (cf. Hermite interpolation in §4.3).

The representation (11) of the remainder is useful only if sufficient information is known about higher derivatives of f. If we know only, for example, that f is continuous on the interval $[a,b]$ (containing the points $x_0,x_1,\ldots,x_n$), then another representation must be used that involves the best (uniform) approximation

$$\mathcal{E}_n = \min_{p\in\mathbb{P}_n} \|f-p\|_\infty = \|f-\hat{p}_n(f;\cdot)\|_\infty$$

of f on $[a,b]$ by polynomials of degree $\le n$. Writing $R_n(f;x) = f(x) - p_n(f;x) = f(x) - \hat{p}_n(f;x) + \hat{p}_n(f;x) - p_n(f;x)$ and noting that $\hat{p}_n - p_n$, being a polynomial of degree $\le n$, can be represented by Lagrange's formula, and $p_n(f;x_k) = f(x_k)$, one gets

$$R_n(f;x) = f(x) - \hat{p}_n(f;x) + \sum_{k=0}^{n} \ell_k(x)[\hat{p}_n(x_k) - f(x_k)].$$

By the triangle inequality, therefore,

$$\|R_n(f;\cdot)\|_\infty \le (1+\Lambda_n)\mathcal{E}_n, \tag{12}$$

where

$$\Lambda_n = \|\lambda_n\|_\infty, \quad \lambda_n(x) = \sum_{k=0}^{n} |\ell_k(x)|$$

are respectively the Lebesgue constant and the Lebesgue function of interpolation (so named by the analogy with the Lebesgue constant in Fourier analysis [49, §8], [12, Ch. 4, §6, p. 147]).

Although $\mathcal{E}_n \to 0$ as $n \to \infty$, by Weierstrass's theorem, one cannot conclude from (12) the same for $\|R_n(f;\cdot)\|_\infty$ since Λ_n is unbounded, at least logarithmically, as $n \to \infty$. Faber [21] indeed has shown in 1914 that no matter how the interpolation points are selected on $[a,b]$, there always exists a continuous function $f \in C[a,b]$ for which Lagrange interpolation diverges. From Jackson's theorem in the theory of best uniform aproximation it is known, however, that in the case of Chebyshev points (for these, see Hermite-Fejér interpolation in §4.3), $\mathcal{E}_n(f)\log n \to 0$ as $n \to \infty$ whenever $f \in C^1[-1,1]$, implying uniform convergence in this case. The same holds for functions that are continuous and of bounded variation in the case of Jacobi points (zeros of the Jacobi polynomial $P_n^{(\alpha,\beta)}$) when $-1 < \max(\alpha,\beta) < 1/2$ [78].

An interesting problem is the determination of optimal nodes minimizing Λ_n over all possible choices of $n+1$ distinct nodes. While characterizations for optimal nodes, originally conjectured by Bernstein and Erdős, have been established by Kilgore [43, 44] and de Boor and Pinkus [8], explicit representations are not known, though good approximations for them (on $[-1,1]$) are the stretched Chebyshev points (stretched linearly so that the smallest and largest of the Chebyshev points come to lie respectively on -1 and $+1$); see [9, §2.3]. Even better nodes exist which are in agreement with the known estimate [79, eq. (2.11)]

$$\Lambda_n = \frac{2}{\pi}\left(\log n + \gamma + \log\frac{4}{\pi}\right) + o(1) \quad \text{as } n \to \infty,$$

where $\gamma = .577215\ldots$ is Euler's constant. The Lebesgue constant Λ_n also measures the stability of Lagrange interpolation with respect to small changes in the data. For many other interesting properties of Λ_n see [79, Sec. 2].

For convergence in the mean, there are more positive results. For example, if one interpolates a function f on $[-1,1]$ at the n zeros $x_k^{(n)}$, $k = 1,2,\ldots,n$, of a polynomial π_n orthogonal (relative to a weight function w) on the interval $[-1,1]$, then one has the classical result of Erdős and Turán [18] that the process converges in the mean for any continuous function f, i.e.,

$$\|f - p_n(f;\cdot)\|_w \to 0 \quad \text{as } n \to \infty, \text{ for all } f \in C[-1,1], \tag{13}$$

where $\|u\|_w^2 = \int_{-1}^1 u^2(x)w(x)dx$. For a survey of results on mean convergence, see [72].

An intriguing and largely open problem (see, however, [53, §4.2.3.3]) is extended Lagrange interpolation, that is, interpolation by a polynomial of degree $\le 2n$ at the n zeros $x_k^{(n)}$ of π_n and $n+1$ additional points τ_j, $j = 1,2,\ldots,n+1$, suitably interlaced with the $x_k^{(n)}$. The most natural choice for the τ_j would be the zeros of the orthogonal polynomial π_{n+1}, but nothing seems to be known about convergence in the mean in this case. Another choice suggested by Bellen [2] are the zeros of the polynomial of degree $n+1$ orthogonal on $[-1,1]$ relative to the weight function $\pi_n^2 w$. Here sufficient conditions are known for convergence in the mean [2] and have been studied in [30]. For a related problem of "quadrature convergence", see also [34].

Complex-variable analysis

An analogue of (11) for analytic functions f has been obtained by Hermite [39] (or [62, pp. 432–443]) in 1878, even in the more general case of Hermite interpolation (cf. §4.3). Assuming f analytic in a domain $\mathcal{D}$ in the complex plane that contains all the points $x_0, x_1, \ldots, x_n$ in its interior, and defining

$$w_n(z) = \prod_{k=0}^{n} (z - x_k), \quad z \in \mathcal{D},$$

there holds

$$R_n(f;z) = \frac{1}{2\pi i} \int_\Gamma \frac{w_n(z)}{w_n(\zeta)} \frac{f(\zeta)}{\zeta - z} d\zeta, \tag{14}$$

where Γ is a contour in $\mathcal{D}$ containing all the x_k and z.

For equidistant points $x_k^{(n)} = a + k(b-a)/n$, $k = 0, 1, \ldots, n$, in $[a,b]$, Runge [63, p. 243] used (14) to obtain a family of concentric contours Γ_ρ, $\rho > 0$, and domains $\mathcal{D}_\rho$ bounded by them,

$$\Gamma_\rho = \{z \in \mathbb{C} : \ \sigma(z) = \rho\}, \quad \mathcal{D}_\rho = \{z \in \mathbb{C} : \ \sigma(z) \le \rho\}, \tag{15}$$

where

$$\sigma(z) = \exp\left\{ \frac{1}{b-a} \int_a^b \ln|z-t| dt \right\}, \tag{16}$$

with the property: if $f \in A(\mathcal{D}_\rho)$ is analytic in $\mathcal{D}_\rho$, then $R_n(f;z) \to 0$ as $n \to \infty$ for any z in the interior of $\mathcal{D}_\rho$, and if $\rho = \rho^*$ is the largest value of ρ for which $f \in A(\mathcal{D}_\rho)$, then $\sup |R_n(f;z)| = \infty$ for z outside of $\mathcal{D}_{\rho^*}$.

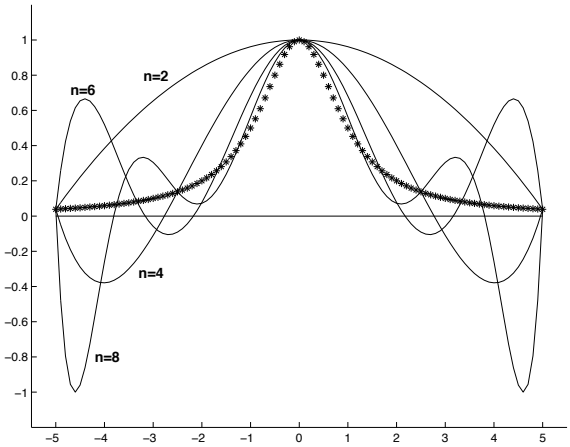

Figure 1: Runge's example

The example given by Runge is the function

$$f(x) = \frac{1}{1+x^2}, \quad -5 \le x \le 5,$$

and the $x_k^{(n)}$ equally spaced on $[-5,5]$,

$$x_k^{(n)} = -5 + k\frac{10}{n}, \quad k = 0,1,2,\ldots,n.$$

What one finds, when one tries to interpolate f for $n = 2,4,6,\ldots$ by polynomials p_n of degree n, is shown in Fig. 1. The approximation provided by p_n, as n increases, becomes better and better in the central part of the interval $[-5,5]$, but worse and worse in the lateral parts of the interval.

Runge, in fact, proves that for real x convergence takes place if $|x| < x^*$ and divergence if $|x| > x^*$, where $x^* = 3.633\ldots$. One can show that $x^* = 3.633384302388339\ldots$ is the solution of the transcendental equation

$$(5+x)\ln(5+x) + (5-x)\ln(5-x) = 10 + 5\int_0^1 \ln(1+25t^2)dt.$$

The domains $\mathcal{D}_\rho$ for this example are shown in Fig. 2 for several values of ρ determined in such a way that Γ_ρ intersects the positive real axis resp. at x^*, 5, 6–8. Note that the first of them, $\mathcal{D}_{\rho^*}$, must pass through the poles $\pm i$. It bounds the largest domain $\mathcal{D}_\rho$

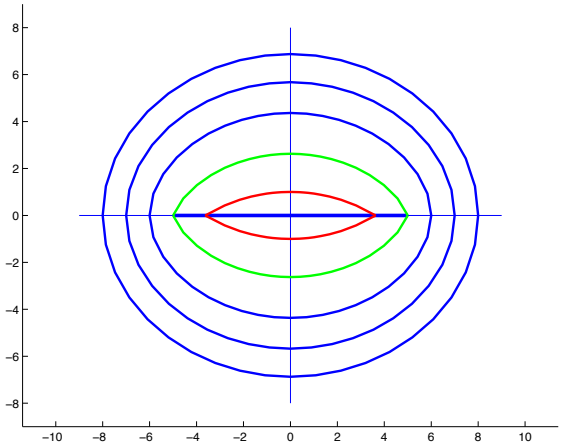

Figure 2: Runge's convergence domains $\mathcal{D}_\rho$

inside of which f is analytic. The interpolation process therefore converges inside this domain but not outside, which explains the behavior shown in Fig. 1. The next domain, the smallest one containing the interval $[-5,5]$, is a critical domain, denoted here by $\mathcal{D}_{\rho^0}$. It has the property that if f is analytic in any domain $\mathcal{D}$ that contains $\mathcal{D}_{\rho^0}$ in

its interior, but can be arbitrarily close to $\mathcal{D}_{\rho^0}$, and if all points $x_k^{(n)}$ are in $\mathcal{D}$, then $p_n(f;z) \to f(z)$ for any $z \in \mathcal{D}$.

If the points $x_k^{(n)}$ are not necessarily equally distributed, but have a limit distribution $d\mu$ on a finite interval $[a,b]$, i.e. $\int_a^x d\mu(t)$, $a < x \leq b$, is the ratio of the number of points $x_k^{(n)}$ in $[a,x]$ to the total number, $n+1$, of points, as $n \to \infty$, then Runge's theory remains valid if dt in (16) is replaced by $d\mu(t)$. (Cf., e.g., Krylov [46, Ch. 12, Sect. 2].) For example, in the case of the arcsin-distribution $d\mu(t) = \dfrac{1}{\pi} \dfrac{dt}{(1-t^2)^{1/2}}$ on $[-1,1]$, typical for points $x_k^{(n)}$ that are zeros of an orthogonal polynomial of degree n, one finds that $\mathcal{D}_{\rho^0} = [-1,1]$; cf. Fig. 3. Thus, in this case, Lagrange interpolation converges for any function f analytic on $[-1,1]$. Figure 3 shows the domains $\mathcal{D}_\rho$ that intersect the real axis at 2, 1.5, 1.1, 1.01 and correspond to the ρ-values 1.36602…, 1.14412…, .88268…, and .75887… .

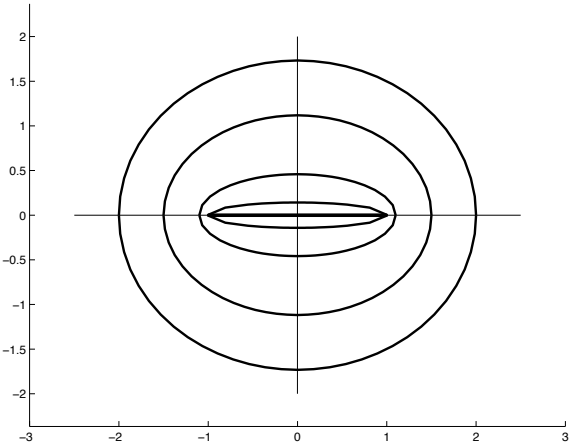

Figure 3: Convergence domains $\mathcal{D}_\rho$ for points with an arcsin-distribution

4.3. Hermite and Hermite–Fejér interpolation

Hermite interpolation

The Hermite interpolation problem consists in obtaining the polynomial of lowest degree that interpolates not only to the function values $f(x_k)$ at the (distinct) points x_k, but also to the successive derivatives $f^{(\mu)}(x_k), \mu = 1,2,\ldots,m_k-1, k = 1,2,\ldots,K$. Hermite in [39] expresses the polynomial (of degree $n = m_1 + m_2 + \cdots + m_K - 1$) as a sum of residues of a contour integral in the complex plane. In practice, it can be obtained as a limit case of Newton's formula, by setting up a table of divided differences in which each point x_k is listed m_k times and Cauchy's limit formulae for confluent divided differences (cf. Real-variable analysis in §4.2) are used, when $m_k > 1$, to initialize the

table. All remaining divided differences are computed by the same rules as in Newton's table of divided differences. The remainder (cf. (11)), accordingly, will assume the form

$$R_n(f;x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{k=1}^{K} (x-x_k)^{m_k}. \tag{17}$$

For $K = 1$, the Hermite interpolation polynomial becomes Taylor's polynomial of degree $m_1 - 1$ and (17) Lagrange's form of the remainder.

A large body of literature is devoted to lacunary Hermite interpolation, also called Birkhoff interpolation, where the derivatives $f^{(\mu)}(x_k)$ involved are not necessarily for successive values of $\mu = 1, 2, 3, \ldots$, but for an arbitrary sequence of μ-values. For this, see, e.g., [50]. An important topic here is the study of existence and uniqueness, which is no longer guaranteed.

Hermite–Fejér interpolation

Soon after Faber's negative result (cf. Real-variable analysis in §4.2) was published, Fejér asked himself whether a polynomial interpolation process exists which would converge for any continuous function on a finite interval. This led him to consider the special case of Hermite interpolation in which $m_k = 2$ and $p'(x_k) = 0$, $k = 1, 2, \ldots, n$, now called Hermite-Fejér interpolation. The corresponding polynomial p of degree $2n-1$ can be represented similarly as in the Lagrange formula (6) (with n replaced by $2n-1$), involving elementary Hermite-Fejér polynomials $h_k(x)$ of degree $2n-1$ in place of the $\ell_k(x)$. The case of Chebyshev points $x_k = \cos((2k-1)\pi/(2n))$ is again very favorable. Fejér [22] (or [77, pp. 25–48]), in 1916, indeed proved that, in contrast to Lagrange interpolation, the interpolation error tends to zero uniformly on $[-1,1]$ for any function continuous on $[-1,1]$.

The Hermite-Fejér interpolation process has since been studied from a number of different angles. One is to look at points other than Chebyshev points. Fejér himself already considered what he calls Gauss points (the zeros of the Legendre polynomial P_n). For $f \in C[-1,1]$ he proves uniform convergence to f on any compact subinterval of $[-1,1]$ and convergence to $\frac{1}{2}\int_{-1}^{1} f(x)dx$ at ± 1. Szegő considers more generally the zeros of the nth-degree Jacobi polynomial with parameters $\alpha, \beta > -1$. He obtains [73, Theorem 14.6], [71, Ch. V, §1.2] uniform convergence to f for arbitrary $f \in C[-1,1]$ whenever $-1 < \alpha, \beta < 0$ but not if $\max(\alpha, \beta) \geq 0$. Grünwald [37] proves the same for another class of points, called ρ-normal, when $\rho > 0$ (cf. [71, Ch. 5, §1.3]). Error estimates are discussed in [71, Ch. 5, §2], and convergence when the function f is suitably restricted, in [71, Ch. 5, §3]. Comparisons with Lagrange interpolation are made in [71, Ch. 6]. For a comprehensive review of the literature on Hermite–Fejér interpolation up to 1987, see [36].

One talks of Hermite–Fejér type interpolation when all derivatives up to some order m are required to vanish at the interpolation points. If these are again Chebyshev points, uniform convergence on $[-1,1]$ for all continuous functions f then holds whenever m is odd, but not for m even ([79, Sec. 4.5, Remark 6]). The convergence behavior

for m odd and m even indeed is similar to respectively Hermite–Fejér and Lagrange interpolation [70].

4.4. Other interpolants and extensions

Functions other than polynomials can of course also serve as interpolants, for example rational functions, trigonometric polynomials, and sinc functions. In each of these cases there exists a barycentric form of the interpolant; for rational functions, see [6], for trigonometric polynomials, [38] when the points x_k are equally spaced and [65, 3] otherwise, and for cardinal sinc-interpolation [4]. Piecewise polynomial interpolants (spline functions, especially cubic splines) are popular in computer-aided geometric design and are treated extensively in [7]. In the area of signal and image processing, a number of techniques are in use, collectively named convolution-based interpolation, for which we refer to [54, Sec. IV].

An important extension is multivariate Lagrange interpolation. This also has a long history [25, 26] and is a topic currently receiving renewed attention [61]. For multivariate Hermite and Birkhoff interpolation, see [52, 51].

4.5. Applications

Quadrature

Newton–Cotes quadrature. The idea of approximating a definite integral of a function f by replacing f by an interpolating polynomial $p_n(f;\cdot)$ and integrating the polynomial p_n instead of f goes back to Newton (cf. §3.2) and has been implemented numerically by Roger Cotes, a protégé of Newton, who passed away at a young age. It provides here a first opportunity to apply Lagrange's interpolation formula.

Considering a general weighted integral $\int_{\mathbb{R}} f(x)d\lambda(x)$, where typically $d\lambda(x) = w(x)dx$ and w is a positive weight function supported on a finite or infinite interval, we integrate (6) (with n instead of $n+1$ points) to get

$$\int_{\mathbb{R}} f(x)d\lambda(x) = \sum_{k=1}^{n} \lambda_k f(x_k) + E_n(f), \tag{18}$$

where

$$\lambda_k = \lambda_k^{(n)} = \int_{\mathbb{R}} \ell_k(x)d\lambda(x), \quad \ell_k(x) = \prod_{\substack{j=1 \\ j\neq k}}^{n} \frac{x-x_j}{x_k-x_j}, \quad k = 1,2,\ldots,n, \tag{19}$$

and $E_n(f) = \int_{\mathbb{R}} R_n(f;x)d\lambda(x)$ is the remainder term. This is called a weighted Newton–Cotes formula, the classical Newton–Cotes formula being the special case with $d\lambda(x) = dx$ on $[-1,1]$ and x_k equally spaced on $[-1,1]$ with $x_1 = 1$ and $x_n = -1$. In this case, Cotes around 1712 published the respective coefficients $\lambda_k^{(n)}$ for $n \leq 11$. For much larger values of n, these formulae become numerically unattractive because of the coefficients $\lambda_k^{(n)}$ becoming large and oscillatory in sign.

It is of interest, therefore, to know of integration measures $d\lambda$ and/or quadrature points x_k which give rise to weighted Newton–Cotes formulae (18) whose coefficients λ_k are all nonnegative. Here again, in the case $d\lambda(x) = dx$ on $[-1,1]$, the Chebyshev points (of the first kind) come to our aid. In this case, and also in the case of Chebyshev points of the second kind, giving rise to what are called Filippi rules, Fejér [23] has shown that for all n the coefficients are positive. The same is true for Clenshaw–Curtis rules [15] whose points x_k are the extreme points (including ± 1) of the Chebyshev polynomial T_{n-1}; the positivity of the quadrature weights in this case has been proved by Imhof [41].

To compute weighted Newton–Cotes formulae, it is convenient to use Gaussian quadrature relative to the measure $d\lambda$ (see below) in (19) combined with the barycentric form (10) of $\ell_k(x)$; see, e.g., Machine Assignment 3 in [33, Ch. 3, pp. 215–217] and its solution therein on pp. 232–247 for some relevant experimentation with x_k the zeros of a Jacobi polynomial and $d\lambda$ a Jacobi measure relative to a different pair of Jacobi parameters. In this case, quasi-positivity, i.e.,

$$\lim_{n\to\infty} \sum_{\substack{k=1 \\ \lambda_k<0}}^{n} |\lambda_k^{(n)}| = 0,$$

has also been studied in [80].

Gaussian quadrature. By construction, n-point Newton–Cotes formulae have polynomial degree of exactness $n-1$, i.e., are exact whenever f is a polynomial of degree $n-1$. As Gauss [27] discovered in 1816, in the case of $d\lambda(x) = dx$, the degree of exactness can be made as large as $2n-1$ (which is optimal when $d\lambda$ is a positive measure) by taking the nodes x_k to be the zeros of the nth-degree orthogonal polynomial π_n relative to the measure $d\lambda$, and the weights λ_k such as to make the formula "interpolatory". The resulting formula has been called the Gauss–Christoffel formula [28], since Christoffel [13] in 1877 generalized Gauss's original formula to arbitrary positive weight functions. What is nice about these formulae is that all weights λ_k are positive, as has been shown very elegantly by Stieltjes [69, pp. 384–385].

The standard way nowadays of computing the n-point Gauss–Christoffel formula is by eigenvalue/vector techniques based on the Jacobi matrix of order n [31, Sec. 3.1.1.1], wich is defined in terms of the recurrence coefficients of the respective orthogonal polynomials. If these are not known explicitly, there are various methods available to compute them numerically [31, Ch. 2].

If $d\lambda$ is supported on a finite interval $[a,b]$, minor extensions of the Gauss–Christoffel formula are the Gauss–Radau and Gauss–Lobatto formulae, having one or both of the end points of $[a,b]$ as nodes. A more substantial extension is the $(2n+1)$-point Gauss–Kronrod formula whose nodes x_k are the n Gauss–Christoffel nodes together with $n+1$ additional real nodes chosen so as to achieve maximum degree of exactness [31, Sec. 3.1.2]. This poses intriguing problems of existence and challenges with regard to constructive methods. For reviews of Gauss–Kronrod quadrature, see Monegato [55, 56, 57], Gautschi [29], and Notaris [60].

Multistep methods in ordinary differential equations

Quadrature, and therefore interpolation, is implicit also in the numerical solution of initial value problems for ordinary differential equations,

$$y' = f(x,y), \quad y(0) = y_0, \tag{20}$$

especially when multistep methods are being used. Basically, one has information (given by f) about the derivative of a real- or vector-valued function y and wants to obtain from this some information about the function itself. In particular, one may want to know how to advance the solution y from some point x to another point $x+h$, where $h > 0$ is a small increment. An answer to this is provided by the fundamental theorem of calculus,

$$y(x+h) = y(x) + \int_x^{x+h} y'(\xi)d\xi. \tag{21}$$

Now suppose we already know the values of the derivative at k points $x, x-h, x-2h, \ldots, x-(k-1)h$, where $k > 1$. Then we may approximate the integral in (21) by passing through the points $y'_s = y'(x-sh), s = 0,1,\ldots,k-1$, a polynomial of degree $k-1$ and then (following Newton!) integrate the polynomial instead of y'. Using the Lagrange interpolation formula, one gets

$$y(x+h) \approx y(x) + h\sum_{s=0}^{k-1} \beta_{k,s} y'_s, \tag{22}$$

where

$$\beta_{k,s} = \int_0^1 \prod_{\substack{r=0 \\ r\neq s}}^{k-1} \left(\frac{t+r}{r-s}\right) dt, \quad s = 0,1,\ldots,k-1.$$

To develop this into a computational algorithm, we take a grid $x_n = nh$, $n = 0,1,2,\ldots$, and denote approximations to $y(x_n)$ by u_n. Then formula (22) suggests the following method,

$$u_{n+1} = u_n + h\sum_{s=0}^{k-1} \beta_{k,s} f_{n-s}, \quad n = k-1, k, k+1, \ldots, \tag{23}$$

where

$$f_m = f(x_m, u_m), \quad m = 0,1,2,\ldots,$$

and it is assumed that the first $k-1$ approximations $u_1, u_2, \ldots, u_{k-1}$ are obtained in some other way. (Of course, $u_0 = y_0$.)

Had we used Newton's interpolation formula rather than Lagrange's, we would have obtained the difference form of the method,

$$u_{n+1} = u_n + h\sum_{s=0}^{k-1} \gamma_s \nabla^s f_n, \quad n = k-1, k, k+1, \ldots, \tag{24}$$

where

$$\gamma_s = \int_0^1 \binom{t+s-1}{s} dt, \quad s = 0,1,\ldots,k-1,$$

and $\nabla f_n = f_n - f_{n-1}$, $\nabla^2 f_n = \nabla(\nabla f_n), \ldots$ are the backward differences of f.

What we have obtained are the explicit k-step Adams–Bashforth method in the Lagrange resp. Newton form, published in 1883 by the astronomer John Couch Adams in the book [1]. There is a companion method—the Adams–Moulton method—which on the right-hand sides of (23) and (24) involves also f_{n+1}, and therefore is an implicit method requiring the solution of a nonlinear equation resp. a system of nonlinear equations. Its derivation, as well as the derivation of many similar multistep methods, uses the same techniques as outlined above.

Lagrangian bases in finite element methods

Classical finite element methods for approximating the solution of a boundary value problem use piecewise polynomial functions over a partition of the domain, with suitable inter-element continuity. For instance, a second-order self-adjoint problem is formulated in a variational (or weak) form, which after "integration by parts" involves only integrals of functions and their first derivatives. Then a corresponding discrete variational formulation is developed by selecting a finite-dimensional space of trial and test functions. Typically, these are polynomials within each element, glued together to form globally C^0 functions.

This paradigmatic scenario suggests the importance of having simple and efficient ways to "glue together" pieces of polynomials, usually having the same local degree. Here, the concept of Lagrange interpolation plays a natural role, leading to the construction of Lagrangian bases in the spaces of trial and test functions.

The simplest example, in one dimension, is given by piecewise linear interpolation on the partition of an interval $[a,b]$ defined by nodes $a = x_0 < x_1 < \cdots < x_\nu < \cdots < x_{n-1} < x_n = b$. Here the basis functions are the elementary Lagrange piecewise linear interpolation functions—the "hat" functions

$$u_\nu(x) = \begin{cases} \dfrac{x - x_{\nu-1}}{x_\nu - x_{\nu-1}} & \text{if } x_{\nu-1} \le x \le x_\nu\,, \\ \dfrac{x_{\nu+1} - x}{x_{\nu+1} - x_\nu} & \text{if } x_\nu \le x \le x_{\nu+1}\,, \\ 0 & \text{elsewhere}\,, \end{cases} \qquad \nu = 0,1,\ldots,n,$$

which like the polynomial counterpart $\ell_\nu(x)$ in (7) satisfy $u_\nu(x_\mu) = \delta_{\nu\mu} = 0$ if $\nu \neq \mu$ and $= 1$ otherwise (cf. Fig. 4). Note that inter-element continuity of the piecewise linear interpolant $p_n(f;\cdot)$ is trivially guaranteed by the fact that the interpolation nodes are the common points between consecutive elements $[x_{\nu-1}, x_\nu]$ and $[x_\nu, x_{\nu+1}]$.

Errors in finite element methods are naturally measured in Sobolev norms, such

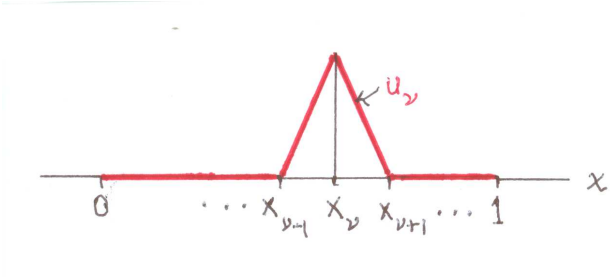

Figure 4: The hat function

as $|f|_m = \left(\int_a^b |f^{(m)}(x)|^2\,\mathrm{d}x\right)^{1/2}$, $m \geq 0$. If $|f|_2$ is finite, one easily proves that

$$|f - p_n(f;\cdot)|_m \leq Ch^{2-m}|f|_2, \quad m = 0,1, \tag{25}$$

where $h = \max_\nu(x_\nu - x_{\nu-1})$. Higher-order approximations are achieved by increasing the polynomial degree in each element.

In two dimensions, the analogue of the previous example is obtained by introducing a triangulation on the domain and requiring that the interpolant restricted to each triangle is an affine function $a + bx + cy$. Inter-element C^0-continuity is naturally

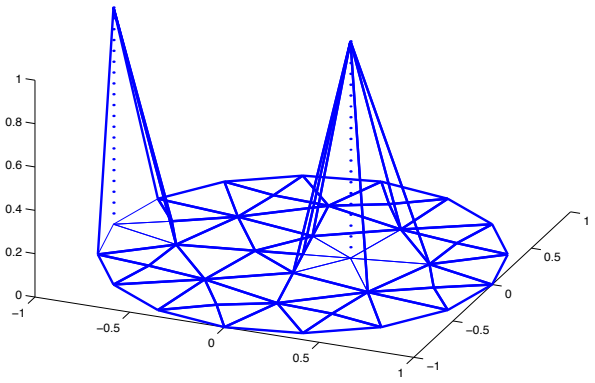

Figure 5: The pyramidal functions

achieved by choosing for the freely variable entities (to be subject to linear conditions to uniquely identify the interpolant) the values at all vertices of the triangulation. Indeed, the restrictions of two affine functions to an edge shared by two triangles coincide if and only if they coincide at the end points of the edge. The Lagrangian basis functions u_ν are now "pyramidal" functions associated with the vertices of the triangulation, each

of them vanishing identically outside the patch of triangles sharing the corresponding vertex (cf. Fig. 5). The element just described is called the Courant element, since its origin dates back to a 1943 paper by R. Courant [16], where it was used in the numerical treatment of a torsion problem.

In general, a Lagrange finite element method is such that all variable entities are values of the approximating functions at a selected set of nodes in the domain. The corresponding Lagrangian basis is obtained by setting to zero all but one of these values. If some of the entities involve certain derivatives of the aproximating functions, one speaks of a Hermite finite element method.

One of the major properties of Lagrangian bases is that they have minimum support (within the linear space of all continuous piecewise polynomial functions over the given domain mesh). Therefore, in the case of linear problems, for example, the final linear systems generated by the finite element methods have maximum degree of sparsity. Moreover, also the numerical evaluation of all the integrals needed to define these linear systems is greatly simplified when one uses such bases.

The mathematical analysis of the finite element method dates back to the late 1960s. The 1972 paper by P. G. Ciarlet and P.-A. Raviart [14] was one of the most influential in the field, since it provided error estimates similar to (25) in a very general setting and with a unifying approach.

Acknowledgements. The author is indebted to Claudio Canuto and Giovanni Monegato for contributing the section Lagrangian bases in finite element methods of §4.5, to Giovanni Monegato for Fig. 5, and to Peter Vértesi for helpful comments to the section Hermite-Fejér interpolation of §4.3.

References

[1] BASHFORTH, FRANCIS AND ADAMS, J. C. *An attempt to test the theories of capillary action by comparing the theoretical and measured forms of drops of fluid, with an explanation of the method of integration employed in constructing the tables which give the theoretical forms of such drops*, Cambridge University Press, Cambridge, 1883.

[2] BELLEN, ALFREDO. Alcuni problemi aperti sulla convergenza in media dell'interpolazione Lagrangiana estesa, *Rend. Istit. Mat. Univ. Trieste* 20 (1988), fasc. suppl., 1–9.

[3] BERRUT, JEAN-PAUL. Baryzentrische Formeln zur trigonometrischen Interpolation I., *Z. Angew. Math. Phys.* 35 (1984), 91–105; Baryzentrische Formeln zur trigonometrischen Interpolation II. Stabilität und Anwendung auf die Fourieranalyse bei ungleichabständigen Stützstellen, *ibid.* 193–205.

[4] BERRUT, JEAN-PAUL. Barycentric formulae for cardinal (SINC-)interpolants, *Numer. Math.* 54 (1989), 703–718. [Erratum: *ibid.* 55 (1989), 747.]

[5] BERRUT, JEAN-PAUL AND TREFETHEN, LLOYD N. Barycentric Lagrange interpolation, *SIAM Rev.* 46 (2004), 501–517.

[6] BERRUT, JEAN-PAUL; BALTENSPERGER, RICHARD; AND MITTELMANN, HANS D. Recent developments in barycentric rational interpolation. Trends and applications in constructive approximation, 27–51, *Internat. Ser. Numer. Math.* 151, Birkhäuser, Basel, 2005.

[7] DE BOOR, CARL. *A practical guide to splines*, rev. ed., Applied Mathematical Sciences 27, Springer, New York, 2001.

[8] DE BOOR, CARL AND PINKUS, ALLAN. Proof of the conjectures of Bernstein and Erdős concerning the optimal nodes for polynomial interpolation, *J. Approx. Theory* 24 (1978), 289–303.

[9] BRUTMAN, L. Lebesgue functions for polynomial interpolation—a survey, The heritage of P. L. Chebyshev: a Festschrift in honor of the 70th birthday of T. J. Rivlin, *Ann. Numer. Math.* 4 (1997), 111–127.

[10] CAUCHY, A. *Œuvres complètes d'Augustin Cauchy*, ser. 1, vols. 1–12 (1882–1900), ser. 2, vols. 1–15 (1905–1974), Académie des Sciences, Gauthier-Villars, Paris.

[11] CAUCHY, A. Sur les fonctions interpolaires, (1840), in [10, ser. 1, v. 5, 409–424].

[12] CHENEY, E. W. *Introduction to approximation theory*, Reprint of the second (1982) edition, AMS Chelsea Publishing, Providence, RI, 1998.

[13] CHRISTOFFEL, E. B. Sur une classe particulière de fonctions entières et de fractions continues, *Ann. Mat. Pura Appl.* (2)8(1877), 1–10. [Also in *Ges. Math. Abhandlungen* II, 42–50.]

[14] CIARLET, P. G. AND RAVIART, P.-A. General Lagrange and Hermite interpolation in R^n with applications to finite element methods, *Arch. Rational Mech. Anal.* 46 (1972), 177–199.

[15] CLENSHAW, C. W. AND CURTIS, A. R. A method for numerical integration on an automatic computer, *Numer. Math.* 2 (1960), 197–205.

[16] COURANT, R. Variational methods for the solution of problems of equilibrium and vibrations, *Bull. Amer. Math. Soc.* 49 (1943), 1–23.

[17] DUBUY, MICHEL. Le calcul numérique des fonctions par l'interpolation barycentrique, *C. R. Acad. Sci. Paris* 226 (1948), 158–159.

[18] ERDŐS, P. AND TURÁN. P. On interpolation.I. Quadrature- and mean-convergence in the Lagrange-interpolation, *Ann. of Math.* 38 (1937), 142–155.

[19] EULER, LEONHARD. *Opera omnia*, ser. 1, v. 1–29 (1911–1956), ser. 2, v. 1–30 (1912–1964), ser. 3, v. 1–12 (1926–1960), B. G. Teubner, Leipzig and Berlin, and Orell Füssli, Zürich.

[20] EULER, LEONHARD. Consideratio quarumdam serierum quae singularibus proprietatibus sunt praeditae, *Novi Comment. Acad. Sci. Petropolitanae* 3 (1750/51) 1753, 10–12, 86–108. [An English translation of this memoir can be downloaded from the E190 page of the Euler Archive at `http://www.math.dartmouth.edu/~euler`.]

[21] FABER, GEORG. Über die interpolatorische Darstellung stetiger Funktionen, Jahresber. Deutsch. Math.-Verein 23 (1914), 192–210.

[22] FEJÉR, L. Über Interpolation, *Göttinger Nachr.* (1916), 66–91.

[23] FEJÉR, L. Mechanische Quadraturen mit positiven Cotesschen Zahlen, *Math. Z.* 37 (1933), 287–309.

[24] FRASER, D. C. *Newton's interpolation formulas*, C. & E. Layton, London. 1927.

[25] GASCA, MARIANO AND SAUER, THOMAS. Polynomial interpolation in several variables. Multivariate polynomial interpolation, *Adv. Comput. Math.* 12 (2000), 377-410.

[26] GASCA, MARIANO AND SAUER, THOMAS. On the history of multivariate polynomial interpolation. Numerical analysis 2000, Vol. II: Interpolation and extrapolaton, *J. Comput. Appl. Math.* 122 (2000), 23–35.

[27] GAUSS, C. F. Methodus nova integralium valores per approximationem inveniendi, *Commentationes Societatis Regiae Scientarium Gottingensis Recentiores* 3 (1816). [Also in *Werke* III, 163–196.]

[28] GAUTSCHI, WALTER. A survey of Gauss–Christoffel quadrature formulae, in *E. B. Christoffel—the influence of his work in mathematics and the physical sciences* (P. L. Butzer and F. Fehér, eds.), 72–147, Birkhäuser, Basel, 1981.

[29] GAUTSCHI, WALTER. Gauss–Kronrod quadrature—a survey, in *Numerical methods and approximation III (Niš, 1987)*, 39–66, Univ. Niš, Niš, 1988.

[30] GAUTSCHI, WALTER. On mean convergence of extended Lagrange interpolation, *J. Comput. Appl. Math.* 43 (1992), 19–35.

[31] GAUTSCHI, WALTER. *Orthogonal polynomials: computation and approximation*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2004.

[32] GAUTSCHI, WALTER. On Euler's attempt to compute logarithms by interpolation: a commentary to his letter of February 16, 1734 to Daniel Bernoulli, *J. Comput. Appl. Math.* 219 (2008), 408–415.

[33] GAUTSCHI, WALTER. *Numerical analysis*, 2d ed., Birkhäuser, New York, 2012.

[34] GAUTSCHI, WALTER AND LI, SHIKANG. On quadrature convergence of extended Lagrange interpolation, *Math. Comp.* 65 (1996), 1249–1256.

[35] GOLDSTINE, H. H. *A history of numerical analysis from the 16th through the 19th century*, Springer, Berlin, 1977.

[36] GONSKA, H. H. AND KNOOP, H.-B. On Hermite–Fejér interpolation: a bibliography (1914–1987), *Studia Sci. Math. Hungar.* 25 (1990), 147–198.

[37] GRÜNWALD, G. On the theory of interpolation, *Acta Math.* 75 (1942), 219–245.

[38] HENRICI, PETER. Barycentric formulas for interpolating trigonometric polynomials and their conjugates, *Numer. Math.* 33 (1979), 225–234.

[39] HERMITE, CH. Sur la formule d'interpolation de Lagrange, *J. Reine Angew. Math.* 84 (1878), 70–79.

[40] HIGHAM, NICHOLAS J. The numerical stability of barycentric Lagrange interpolation, *IMA J. Numer. Anal.* 24 (2004), 547–556.

[41] IMHOF, J. P. On the method for numerical integration of Clenshaw and Curtis, *Numer. Math.* 5 (1963), 138–141.

[42] JOFFE, S. A. Interpolation-formulae and central-difference notation, *Trans. Actuar. Soc. Amer.* 18 (1917), 72–98.

[43] KILGORE, T. A. Optimization of the norm of the Lagrange interpolation operator, *Bull. Amer. Math. Soc.* 83 (1977), 1069–1071.

[44] KILGORE, THEODORE A. A characterization of the Lagrange interpolating projection with minimal Tchebycheff norm, *J. Approx. Theory* 24 (1978), 273–288.

[45] KOELINK, ERIK AND VAN ASSCHE, WALTER. Leonhard Euler and a q-analogue of the logarithm, *Proc. Amer. Math. Soc.* 137 (2009), 1663–1676.

[46] KRYLOV, VLADIMIR IVANOVICH. *Approximate calculation of integrals* [translated from the Russian by Arthur H. Stroud], ACM Monograph Ser., Macmillan, New York, 1962.

[47] LAGRANGE, J. -L. Sur les interpolations, (1778), in [67, v. 7, 535–553].

[48] LAGRANGE, J. -L. La méthode d'interpolation, (1792/93), in [67, v. 5, 663–684].

[49] LEBESGUE, HENRI. Sur la représentation trigonométrique approchée des fonctions satisfaisant à une condition de Lipschitz, *Bull. Soc. Math. France* 38 (1910), 184–210. [Also in *Œuvres Scientifiques*, III, 363–389.]

[50] LORENTZ, GEORGE G., JETTER, KURT, AND RIEMENSCHNEIDER, SHERMAN D. *Birkhoff interpolation*, Encyclopedia of Mathematics and its Applications, 19, Addison Wesley Publ. Co., Reading, MA, 1983.

[51] LORENTZ, RUDOLPH A. Multivariate Birkhoff interpolation, *Lecture Notes in Mathematics* 1516, Springer, Berlin, 1992.

[52] LORENTZ, R. A. Multivariate Hermite interpolation by algebraic polynomials: a survey. Numerical analysis 2000, Vol. II: Interpolation and extrapolation, *J. Comput. Appl. Math.* 122 (2000),167–201.

[53] MASTROIANNI, GIUSEPPE AND MILOVANOVIĆ, GRADIMIR V. *Interpolation processes: basic theory and applications*, Springer Monographs in Mathematics, Springer, Berlin, 2008.

[54] MEIJERING, ERIK. A chronology of interpolation: from ancient astronomy to modern signal and image processing, *Proc. IEEE* 90(3)(2002), 319–342.

[55] MONEGATO, GIOVANNI. An overview of results and questions related to Kronrod schemes, in G. Hämmerlin, ed., *Numerische Integration, Internat. Ser. Numer. Math.* 45, 231–240, Birkhäuser Basel, 1979.

[56] MONEGATO, GIOVANNI. Stieltjes polynomials and related quadrature rules, *SIAM Rev.* 24 (1982), 137–158.

[57] MONEGATO, GIOVANNI. An overview of the computational aspects of Kronrod quadrature rules, *Numer. Algorithms* 26 (2001), 173–196.

[58] DE MORGAN, A. *The differential and integral calculus*, Baldwin and Cradock, London, 1842.

[59] NEWTON, I. *Philosophiæ naturalis principia mathematica* (1687). [English translation in *Sir Isaac Newton's mathematical principles of natural philosophy and his system of the world* (Florian Cajori, ed.), University of California Press, Berkeley, CA, 1946.]

[60] NOTARIS, SOTIRIOS E. An overview of results on the existence or nonexistence and the error term of Gauss–Kronrod quadrature formulae, in R. V. W. Zahar, ed., *Approximation and computation (West Lafayette, IN, 1993), Internat. Ser. Numer. Math.* 119, 485–496, Birkhäuser Boston, 1994.

[61] OLVER, PETER J. On multivariate interpolation, *Stud. Appl. Math.* 116 (2006), 201–240.

[62] PICARD, ÉMILE, ed. *Œuvres de Charles Hermite*, v. 3, Gauthier-Villars, Paris, 1912.

[63] RUNGE, CARL. Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten, *Z. Math. und Phys.* 46 (1901), 224–243.

[64] RUTISHAUSER, HEINZ. *Vorlesungen über numerische Mathematik*, vol. 1: Gleichungssysteme, Interpolation und Approximation, Birkhäuser, Basel, 1976. [Annotated English translation in *Lectures on numerical mathematics*, Birkhäuser, Boston, MA, 1990.]

[65] SALZER, HERBERT E. Coefficients for facilitating trigonometric interpolation, *J. Math. Physics* 27 (1949), 274–278.

[66] SALZER, H. E. Lagrangian interpolation at the Chebyshev points $x_{n,\nu} \equiv \cos(\nu\pi/n)$, $\nu = 0(1)n$; some unnoted advantages, *Comput. J.* 15 (1972), 156–159.

[67] SERRET, J. -A. , ed. *Œuvres de Lagrange*, vols. 1–14, Gauthier-Villars, Paris, 1867–1892.

[68] SMITH, SIMON J. Lebesgue constants in polynomial interpolation, *Ann. Math. Inform.* 33 (2006), 109–123.

[69] STIELTJES, T. J. Quelques recherches sur la théorie des quadratures dites mécaniques, *Ann. Sci. Éc. Norm. Paris*, Sér. 3, 1 (1884), 409–426. [Also in *Œuvres* I, 377–396.]

[70] SZABADOS, J. On the order of magnitude of fundamental polynomials of Hermite interpolation, *Acta Math. Hungar.* 61 (1993), 357–368.

[71] SZABADOS, J. AND VÉRTESI, P. *Interpolation of functions*, World Scientific, Singapore, 1990.

[72] SZABADOS, J. AND VÉRTESI, P. A survey on mean convergence of interpolatory processes, *J. Comput. Appl. Math.* 43 (1992), 3–18.

[73] SZEGŐ, GABOR. *Orthogonal polynomials*, American Mathematical Society, Colloquium Publications 23, 4th ed., American Mathematical Society, Providence, RI, 1975.

[74] TAYLOR, WILLIAM J. Method of Lagrangian curvilinear interpolation, *J. Research Nat. Bur. Standards* 35 (1945), 151–155.

[75] TURNBULL, H. W. James Gregory: a study in the early history of interpolation, *Proc. Edinburgh Math. Soc.*, ser. 2, v. 3 (1932), 151–172.

[76] TURNBULL, H. W. (ed.) *The correspondence of Isaac Newton*, v. 2, 1676–1687, Cambrige University Press, 1960.

[77] TURÁN, PÁL, ed. *Leopold Fejér—Gesammelte Arbeiten*, vol. 2, Akadémiai Kiadó, Budapest, 1970.

[78] VÉRTESI, PÉTER. One sided convergence conditions for Lagrange interpolation, *Acta Sci. Math. (Szeged)* 45 (1983), 419–428.

[79] VÉRTESI, PÉTER. Classical (unweighted) and weighted interpolation, in *A panorama of Hungarian mathematics in the twentieth century*, 71–117, Bolyai Soc. Math. Stud. 14, Springer, Berlin, 2006.

[80] VÉRTESI, PÉTER. Some remarks on quadrature formulae, *Jaen J. Approx.* 2 (2010), 31–49.

[81] WARING, EDWARD. Problems concerning interpolations, *Philos. Trans. R. Soc. Lond.* 69 (1779), 59–67.

AMS Subject Classification: 33C47, 65D30

Walter GAUTSCHI
Department of Computer Sciences, Purdue University
West Lafayette, IN 47907-2066, USA
e-mail: wgautschi@purdue.edu

Lavoro pervenuto in redazione il 07.06.2013