Lectures on Multivariate Polynomial Interpolation

Jean-Paul Calvi

Hanoi, summer 2005

Instructions for the students

The study of the following preliminary chapter is not a necessary prerequisite for my lectures on multivariate polynomial interpolation. However, I strongly recommand to study it. The reading of this chapter should make easier the course and make its content more natural. The proofs (especially, some computations) are not always written in full details. The reader should try to provide the missing details. Also, he or she should go through the text with a critical eye for it might contain some errors and is certainly full of misprints. I would be very grateful to be informed of any problem at calvi@picard.ups-tlse.fr.

Toulouse, June 2005
JPC

Notation and symbols

We shall freely use the commonest notation such as $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.

Sets and spaces of functions

$\mathsf{C}(X)$	space of continuous functions on X
$\mathsf{C}^d(X)$	space of functions d times differentiable on the open set X with a continuous d-th (Fréchet) derivative
$\mathcal{GL}_n(\mathbb{K})$	invertible matrices of order n and coefficients in $\mathbb{K}$
$\mathcal{H}_d(\mathbb{K}^n)$	homogeneous polynomials of n variables and degree d
$\mathsf{H}(\Omega)$	holomorphic (analytic) functions on $\Omega \subset \mathbb{C}$ (or $\mathbb{C}^n$)
$\mathbb{K}$	either $\mathbb{R}$ or $\mathbb{C}$
$\mathbb{K}[X_1, X_2, \cdots, X_n]$	(algebraic) polynomials of n variables with coefficients in $\mathbb{K}$
$\mathbb{N}^\star$	$\{1, 2, 3, \dots\}$
$\mathcal{P}(\mathbb{K}^n)$	polynomials of n variables with coefficients in $\mathbb{K}$
$\mathcal{P}_d(\mathbb{K}^n)$	polynomials in $\mathcal{P}(\mathbb{K}^n)$ of degree at most d
$\mathbb{R}_+$	the set of non negative real numbers

Various symbols

$\mathrm{cv}(X)$	convex hull of X
$\mathbf{e}^\alpha$	monomial function defined by $\mathbf{e}^\alpha(x) = x^\alpha$
$\mathrm{Ind}_\gamma(a)$	index (or contour winding number) of a with respect to the (closed) curve γ
J_ψ	jacobian (determinant) of the mapping ψ
$\mathbf{L}_A$ (or $\mathbf{L}[A,]$)	Lagrange (or Lagrange-Hermite) interpolation polynomial at (the points of) A
$\|\|f\|\|_I$	sup-norm of f on I
$\mathrm{span}\,\{f_1,\ldots,f_n\}$	(vector) space spanned by $f_1,\ldots f_n$
$\langle f_1,\ldots,f_n\rangle$	ideal spanned by $f_1,\ldots,f_n$
$\mathbf{T}_a^d$	Taylor polynomial of f at a and to the order d
VDM	Vandermonde determinant
$\omega(f,\cdot)$	the modulus of continuity of the function f

CHAPTER 1

A short overview of univariate Lagrange-Hermite interpolation

We use $\mathcal{P}_d(\mathbb{K})$ for the space of univariate polynomials of degree at most d with coefficients in $\mathbb{K}$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. The space of (all) polynomials is $\mathcal{P}(\mathbb{K}) := \cup_{d=0}^{\infty}\mathcal{P}_d(\mathbb{K})$. When there is no danger of confusion we write $\mathcal{P}_d$ (resp. $\mathcal{P}$) instead of $\mathcal{P}_d(\mathbb{K})$ (resp. $\mathcal{P}(\mathbb{K})$).

Given $w, p, q \in \mathcal{P}$, $p \equiv q[w]$ means that w divides $p - q$, that is to say, there exists $s \in \mathcal{P}$ such that $p - q = s.w$. Recall that

$$\left.\begin{matrix} p_1 \equiv q_1 \quad [w] \\ p_2 \equiv q_2 \quad [w] \end{matrix}\right\} \implies \left\{\begin{matrix} p_1 + p_2 \equiv q_1 + q_2 \quad [w] \\ p_1 p_2 \equiv q_1 q_2 \qquad [w] \end{matrix}\right.$$

§ 1 Definition and Basic Properties

1.1 Lagrange Interpolation

The problem of (univariate) interpolation is the following simple one : given, say, $d + 1$ points $(x_i, y_i) \subset \mathbb{K}^2$, find a polynomial p, as simple as possible, the graph of which passes through these $d+1$ points. Equivalently, given a function f defined on $A = \{x_0, x_1, \ldots, x_d\} \subset \mathbb{C}$, find $p = L(f)$ in $\mathcal{P}(\mathbb{K})$ such that the restriction of $L(f)$ to A is exactly f. Here simplicity is naturally interpreted in terms of degree and the starting point of the theory lies in the following theorem.

Theorem 1.1. *Let $A = \{x_0, \ldots, x_d\} \subset \mathbb{K}$ be a set of $d+1$ (distinct) points and f a function defined on A. There exists a unique polynomial $p \in \mathcal{P}_d$ such that*

$$p(x_i) = f(x_i) \qquad (i = 0, 1, \ldots, d). \tag{1.1}$$

Proof. A polynomial $p(x) = \sum_{j=0}^{d} a_j x^j$ satisfies (1.1) if and only if its coefficients (a_j) form a solution of the linear system

$$\begin{cases} a_0 & + & a_1x_0 & + & \dots & + & a_dx_0^d & = & f(x_0) \\ a_0 & + & a_1x_1 & + & \dots & + & a_dx_1^d & = & f(x_1) \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_0 & + & a_1x_d & + & \dots & + & a_dx_d^d & = & f(x_d) \end{cases}$$

This system has a unique solution if and only if its determinant does not vanish. Such a determinant is called a **Vandermonde's determinant**, or **vandermondian**, and is denoted by $\mathrm{VDM}(x_0, \dots, x_d)$. Thus

$$\mathrm{VDM}(x_0, \dots, x_d) = \begin{vmatrix} 1 & x_0 & \cdots & x_0^d \\ 1 & x_1 & \cdots & x_1^d \\ \vdots & \vdots & & \vdots \\ 1 & x_d & \cdots & x_d^d \end{vmatrix}.$$

It is not difficult to verify[1] that

$$\mathrm{VDM}(x_0, \dots, x_d) = \prod_{0 \le i < j \le d} (x_i - x_j). \tag{1.2}$$

Hence, since the points are pairwise distinct, the determinant does not vanish. □

Kramer's rules further give

$$p(x) = \sum_{j=0}^{d} f(x_j) \frac{\mathrm{VDM}(x_0, \dots, x_{j-1}, x, x_{j+1}, \dots, x_d)}{\mathrm{VDM}(x_0, \dots, x_d)} \tag{1.3}$$

which, in view of (1.2), easily simplifies to

$$p(x) = \sum_{i=0}^{d} f(x_i) \prod_{j=0, j \ne i}^{d} \frac{x - x_j}{x_i - x_j}. \tag{1.4}$$

A shorter proof of Theorem 1.1 is easily obtained if we guess at the beginning that (1.4) provides a polynomial for which (1.1) holds true. For then, it just remains to prove uniqueness and this can be done as follows. If p_1 and p_2 both satisfy the required property then $p_2 - p_1 \in \mathcal{P}_d$ and has at least $d+1$ roots which implies $p_1 - p_2 = 0$. Two other ways of proving Theorem 1.1 will be presented below in the more general setting of Lagrange-Hermite interpolation.

The unique solution p of 1.1 will be denoted by $\mathbf{L}_A(f)$ and called the **Lagrange interpolation polynomial** of f at A. The elements of A are the **interpolation points** sometimes also called **nodes**. Formula (1.4) is the **Lagrange interpolation formula** and the polynomials

$$\ell_i(x) = \prod_{j=0, j \ne i}^{d} \frac{x - x_j}{x_i - x_j} \qquad (i = 0, 1, \dots, d) \tag{1.5}$$

[1] It is a classical exercise on determinants. We may for example subtract the last line from the first d lines then factor out $(x_i - x_d)$ in the i-th line for $i = 0, \dots, d-1$ to realize that the remaining determinant is but $\mathrm{VDM}(a_0, a_1, \cdots, a_{d-1})$.

are the **fundamental Lagrange interpolation polynomials** for A, hereafter abbreviated to FLIP.

If used for practical computations, the Lagrange interpolation formula (1.4) should be transformed in (1.6) below. Setting

$$w_j = \Big(\prod_{k\neq j}(a_j - a_k)\Big)^{-1} \quad (j = 0, \ldots, d),$$

the (naturally so-called) **barycentric Lagrange interpolation formula** is

$$\mathbf{L}_A(f)(x) = \frac{\sum_{j=0}^{d} \frac{w_j}{x-a_j} f(a_j)}{\sum_{j=0}^{d} \frac{w_j}{x-a_j}}. \tag{1.6}$$

To prove this, we first observe that

$$\ell_i(x) = w(x).\frac{w_i}{(x-a_i)} \quad \text{with } w(x) = \prod_{i=0}^{d}(x-a_i) \tag{1.7}$$

and $1 = \sum_{i=0}^{d} \ell_i$ (for every constant polynomial is equal to its interpolation polynomial). Then we compute $\mathbf{L}_A(f)(x)$ as $\mathbf{L}_A(f)(x)/1$ applying (1.4) both to $\mathbf{L}_A(f)(x)$ and to 1 and conclude by canceling out the common factor $w(x)$ in the quotient. Observe finally that $w_i = w'(a_i)$ so that (1.7) gives the following often useful expression of the FLIP's,

$$\ell_i(x) = \frac{w(x)}{w'(a_i)(x-a_i)} \quad (i = 0, \ldots, d). \tag{1.8}$$

1.2 Lagrange-Hermite interpolation

So far we looked for a graph passing through $d+1$ given points. It is very natural (and often necessary) to control the local geometry of the graph near the interpolation points. This amounts to impose conditions on the derivatives of the polynomial at the interpolation points and thus leads to the concept of Lagrange-Hermite interpolation[2].

Let $B = \{b_0, \ldots, b_k\} \subset \mathbb{K}$ be a set of k distinct points and a **multiplicity function** $n : B \to \mathbb{N}^\star$ such that $\sum_{i=0}^{k} n(b_i) = d+1$. For simplicity, we shall write $n(i)$ instead of $n(b_i)$. The number $n(i)$ is the **multiplicity** of b_i. We set $w_i(x) = (x-b_i)^{n(i)}$.

Theorem 1.2. *Given a function f for which $f^{(n(i))}(b_i)$ exists for $i = 0, \ldots, k$, there exists a unique $p \in \mathcal{P}_d(\mathbb{K})$ such that*

$$p^{(j)}(b_i) = f^{(j)}(b_i) \qquad (0 \le j \le n(i)-1,\ 0 \le i \le k). \tag{1.9}$$

Of course, we obtain Theorem 1.1 as a particular case. Another important particular case is obtained on taking only one point, say $b_0 = b$, and $n(0) = d$. In this case, the polynomial p of the theorem is but the Taylor polynomial of f at b to the order d. Note that when $\mathbb{K} = \mathbb{C}$, we deal with the complex derivatives of f. Conditions (1.9) will be referred to as the **interpolation conditions**.

[2]However, historically, this is not the way Lagrange-Hermite interpolation was introduced, see the notes at the end of this chapter.

Proof. We prove that the linear map

$$\Phi : p \in \mathcal{P}_d \to (p(b_0), \ldots, p^{(n(0)-1)}(b_0), \ldots, p(b_k), \ldots, p^{(n(k)-1)}(b_k)) \in \mathbb{K}^{d+1}$$

is one-to-one. Since $\mathcal{P}_d$ and $\mathbb{K}^{d+1}$ have the same dimension $d+1$, this will show that Φ is a linear isomorphism. The required polynomial is then given by the (unique) pre-image of $(f(b_0), \ldots, f^{(n(0)-1)}(b_0), \ldots, f(b_k), \ldots, f^{(n(k)-1)}(b_k))$.

Now, to prove that Φ is one-to-one we show $\ker \Phi = \{0\}$ as follows. If $\Phi(p) = 0$ then b_0 is a zero of order $n(0)$ of p so that p can be written as $p = w_0 Q_0$. Similarly w_1 divides p. Since $b_0 \neq b_1$, w_1 et w_0 are relatively prime, hence w_1 must divide Q_0 and $p = w_0 w_1 Q_1$. Continuing in this way, we obtain $p = w_0 w_1 \ldots w_k Q_k$. Comparing the degrees of both sides, we see that Q_k must be equal to 0 which in turn gives $p = 0$. □

The unique polynomial p satisfying the conditions in Theorem 1.2 is denoted by $\mathbf{L}_{(B,n)}(f)$ and called the **Lagrange-Hermite interpolation polynomial** of f at (B, n). To simplify the notation, it often helps to use sets in which elements may be repeated. To every such set corresponds uniquely a couple (B, n). For example if $A = \{1, 2, 1, 3, 4, 1, 2\}$ then $B = \{1, 2, 3, 4\}$ and $n(1) = 3$, $n(2) = 2$, $n(3) = 1$ and $n(4) = 1$ so that $n(b_i)$ is but the number of appearance of b_i. It will be soon more convenient to write $\mathbf{L}[A; f]$ or $\mathbf{L}[a_0, \ldots, a_d; f]$ or $\mathbf{L}_A(f)$ rather than $\mathbf{L}_{(B,n)}$. Sometimes, we shall say, for short, that $\mathbf{L}_A(f)$ interpolates f at A. We shall see later that every Lagrange-Hermite interpolation polynomial is the limit (in an appropriate sense) of Lagrange interpolation polynomials (Theorem 1.16) and this amply justifies the use of the same notation for the interpolation polynomials of Theorems 1.1 and 1.2.

1.3 Linearity and affine invariance

The map $f \to \mathbf{L}_A$ is a linear projector that is, a linear map that coincides with the identity on its range $\mathcal{P}_d$. Of course, there exists a formula corresponding to the Lagrange interpolation formula (1.4). Namely, if A corresponds to (B, n), we can find polynomials ℓ_{ij} such that

$$\mathbf{L}_A(f) = \sum_{i=0}^{k} \sum_{j=0}^{n(i)-1} f^{(j)}(b_i)\ell_{ij} \tag{1.10}$$

but their expression is no longer simple (in general) and it is necessary to look for more useful representations of the interpolation polynomial. (However, see below the discussion leading to (1.41).)

Proposition 1.3. *The interpolation polynomial is invariant under affine mapping — that is, mappings of the form $t : x \to \alpha x + \beta$ with α and β in $\mathbb{K}$ and $\alpha \neq 0$. This means*

$$\mathbf{L}_A(f \circ t) = \mathbf{L}_{t(A)}(f) \circ t. \tag{1.11}$$

Proof. Since both sides are polynomials of degree d, it suffices to prove that they verify the same interpolation conditions. Now, if $a \in A$ with multiplicity n then $t(a) \in t(A)$ with multiplicity n and, for $0 \leq k \leq n-1$, we have

$$\begin{aligned}\left(\mathbf{L}_{t(A)}(f) \circ t\right)^{(k)}(a) &= \alpha^k \left(\mathbf{L}_{t(A)}(f)\right)^{(k)}(t(a)) \\ &= \alpha^k f^{(k)}(t(a)) = (f \circ t)^{(k)}(a) \\ &= \left(\mathbf{L}_A(f \circ t)\right)^{(k)}(a).\end{aligned}$$

□

Proposition 1.4. *If $A^1 \subset A^2$ then $\mathbf{L}_{A^1} \circ \mathbf{L}_{A^2} = \mathbf{L}_{A^1}$.*

Here it is intended that equality holds true on the space of functions for which both sides are well defined. In case of sets with repeated points $A^1 \subset A^2$ means that if A^1 corresponds to (B^1, n^1) and A^2 corresponds to (B^2, n^2) then $B^1 \subset B^2$ and $n^1(b) \leq n^2(b)$ for every $b \in B^1$. This being said, the proof of the proposition goes along the same lines as the previous one. We just need to check that, for every f for which both sides are defined, the polynomials $\mathbf{L}_{A^1}\big(\mathbf{L}_{A^2}(f)\big) = \mathbf{L}_{A^1}(f)$ satisfy the same interpolation conditions.

1.4 A ring theoretic approach

In Theorem 1.2, f and p are actually required to share, for $i = 0, 1, \ldots, k$, the same Taylor polynomials at a_i to the order $n(i)$, that is,

$$p = \mathbf{L}_{(B,n)}(f) \iff \begin{cases} p \in \mathcal{P}_d(\mathbb{K}) \\ p \equiv \mathbf{T}_i(f) \qquad [w_i] \quad (i = 0, 1, \ldots, k) \end{cases} \tag{1.12}$$

where

$$\mathbf{T}_i(f)(x) = \sum_{j=0}^{n(i)-1} \frac{f^{(j)}(a_i)}{j!}(x - a_i)^j.$$

This observation suggests a ring theoretic proof of Theorem 1.2. Indeed, since the points a_i are pairwise distinct, the k polynomials w_i are co-prime. Hence, the chinese remainder theorem in the principal ring $\mathcal{P}(\mathbb{K})$ ($\simeq \mathbb{K}[X]$) ensures the existence of a unique polynomial p satisfying the conditions on the right hand side of (1.12). Note that if we know that $p \equiv T_i(f)$ $[w_i]$ $(i = 0, 1, \ldots, k)$ then one but can say that $p \equiv \mathbf{L}_{(B,n)}(f) \quad [w_0 \ldots w_k]$. If it is further known that $\deg p \leq d$ then we may conclude that $p = \mathbf{L}_{(B,n)}(f)$. This idea is used in the proof of the next theorem.

1.5 The Neville-Aitken formula

It is a useful relation between interpolation polynomials at d nodes and $d-1$ nodes. A popular algorithm for computing the interpolation polynomial, the so-called Neville algorithm is based on this relation.

Theorem 1.5.

$$(a_0 - a_d)\mathbf{L}[A; f] = \mathbf{L}[A \setminus \{a_d\}; f](\cdot - a_d) - \mathbf{L}[A \setminus \{a_0\}; f](\cdot - a_0) \tag{1.13}$$

Proof. Let $w(x) = (x - a_0) \ldots (x - a_d)$. It follows from the definition that

$$\begin{aligned} \mathbf{L}[A; f] &\equiv \mathbf{L}[A \setminus \{a_d\}; f] \quad [w(X).(X - a_d)^{-1}] \\ \mathbf{L}[A; f] &\equiv \mathbf{L}[A \setminus \{a_0\}; f] \quad [w(X).(X - a_0)^{-1}]. \end{aligned}$$

Multiplying the first formula by $(X - a_d)$, the second by $(X - a_0)$ and subtracting the first to the second, we obtain

$$(a_0 - a_d)\mathbf{L}[A; f] \equiv (x - a_d)\mathbf{L}[A \setminus \{a_d\}; f] - (x - a_0)\mathbf{L}[A \setminus \{a_0\}; f] \qquad [w].$$

In view of the discussion in the previous paragraph, the Neville-Aitken formula (1.13) follows for both sides are polynomials of degree not greater than d. □

[1.13]

1.6 A first remainder formula

The following theorem shows that Lagrange-Hermite interpolation always provides good local approximation of sufficiently smooth functions.

Theorem 1.6. *Let I be an interval (in $\mathbb{R}$) containing $A = \{a_0, \ldots, a_d\}$ and let $f \in \mathsf{C}^{d+1}(I)$. For every $x \in I$ there exists $\xi \in I$ such that*

$$f(x) - \mathbf{L}[A; f](x) = \frac{f^{(d+1)}(\xi)}{(d+1)!}(x - a_0) \ldots (x - a_d). \tag{1.14}$$

Consequently

$$||f - \mathbf{L}[A; f]||_I \leq \frac{||f^{(d+1)}||_I}{(d+1)!}.||(\cdot - a_0) \ldots (\cdot - a_d)||_I. \tag{1.15}$$

Proof. We shall prove the theorem under the weaker assumption that the $(d+1)$-th derivative exists on I. For a fixed x in I, we may choose a constant K (depending on x) such that the function

$$g : t \to f(t) - \mathbf{L}[a_0, \ldots, a_d; f](t) - K(t - a_0) \ldots (t - a_d)$$

vanishes at the $d+2$ points $x, a_0, \ldots, a_d$ taking multiplicity into account. Then, using sufficiently many times Rolle's theorem, we get that for some ξ, $g^{(d+1)}(\xi) = 0$ which gives $(d+1)! \cdot K = f^{d+1}(\xi)$ by differentiating the expression of g. One concludes the proof by using $g(x) = 0$. □

1.7 Chebyshev's nodes

Theorem 1.6 provides first insight into the way the choice of the interpolation points influences the quality of approximation furnished by Lagrange-Hermite interpolation. Indeed, in view of (1.15), if we are free to choose the interpolation points and if we do not know any further information on the function f to be interpolated, then the most natural strategy is to take $a_0, \ldots, a_d$ in order that $||(\cdot - a_0)(\cdot - a_1) \ldots (\cdot - a_d)||_I$ be as small as possible. In other words, we must find a monic polynomial of smallest sup-norm on I and choose its roots as interpolation points. It turns out that these points are uniquely determined. Namely they are the roots of the $d+1$-th Chebyshev polynomial the definition of which we now briefly recall.

We restrict ourselves to the case of $I = [-1, 1]$. The general case $I = [a, b]$ is reduced to the case $I = [-1, 1]$ by using an affine mapping $x \to ax + b$. Now, recall that for every $d \in \mathbb{N}$, the d-th **Chebyshev polynomial** T_d is the (unique) polynomial of degree d for which $T_d(\cos\theta) = \cos d\theta$ for $\theta \in \mathbb{R}$. The sequence (T_d) verifies the terms recurrence relation $T_{d+1}(x) = 2xT_d(x) - T_{d-1}(x)$ $(d = 1, 2, \ldots)$. It follows that the leading term of T_d is 2^{d-1}. The roots — hereafter called (d-th)-**Chebyshev's points** — are easy to compute,

$$T_d^{-1}(0) = \left\{\cos\left(\frac{2k+1}{d}\frac{\pi}{2}\right),\ k = 0, 1, \ldots d-1\right\}. \tag{1.16}$$

In particular the root are simple and belong to $[-1, 1]$. Moreover $||T_d||_I = 1$ and the sup-norm is attained at the $d+1$ points $m_k = \cos\frac{k\pi}{d}$, $k = 0, 1, \ldots, d$. Precisely, we have $T_d(m_k) = (-1)^k$. From this, using the Chebyshef equioscillation theorem[3], we deduce that $x^d - 2^{1-d}T_d(x)$ is the best polynomial approximation of degree $d-1$ of the monomial x^d on I and this is equivalent to Theorem 1.7 below. We shall give an alternate direct proof.

Theorem 1.7. *Let $I = [-1, 1]$. The number $||(.-a_0)(.-a_1)\dots(.-a_d)||_I$ is made minimal when and only when $\{a_0, \dots, a_d\} = T_{d+1}^{-1}(0)$. In other words, the (unique) monic polynomial of degree $d+1$ of minimal norm on I is $t_{d+1} := 2^{-d}T_{d+1}$.*

Proof. For simplicity of notation, we shall write $b_k = \cos\frac{2k+1}{d+1}\frac{\pi}{2}$ $(0 \le k \le d)$ and $t := t_{d+1}$, thus $t(x) = (x-b_0)(x-b_1)\dots(x-b_d)$. Now, supposing that p is a monic polynomial of degree $d+1$ such that $||p||_I \le ||t||_I$, we want to prove that p must be equal to t. Since $|t|$ reaches its maximum at the $d+2$ points m_k $(0 \le k \le d+1)$[4] we have $|p(m_k)| \le ||p||_I \le ||t||_I = |t(m_k)|$. Hence, since $2^d t(m_k) = (-1)^k$, we have

$$(-1)^k(p-t)(m_k) \le 0. \qquad (0 \le k \le d+1) \tag{1.17}$$

Now, since $(-1)^k(p-t)(m_k) \le 0$ and $(-1)^k(p-t)(m_{k+1}) \ge 0$, we have only the following two (mutually exclusive) possibilities

(a) For $k = 0, \dots, d$, there exists $\xi_k \in]m_k, m_{k+1}[$ such that $(-1)^k(p-t)'(\xi_k) > 0$ ($(-1)^k(p-t)$ must increase somewhere between m_k and m_{k+1})

(b) There exists k_0 such that $p-t$ is constant on $[m_{k_0}, m_{k_0+1}]$ (which forces in particular $(p-t)(m_{k_0}) = (p-t)(m_{k_0+1}) = 0$).

(b) immediately gives $p = t$. We shall show that (a) leads to a contradiction and must therefore be excluded. Indeed, if (a) holds true then we can find a root of $(p-t)'$ in every interval $]\xi_k, \xi_{k+1}[$ for $k = 0, \dots, d-1$. This gives at least d roots for $(p-t)'$ which is impossible since, again by (a), $(p-t)'$ is a nonzero polynomial of degree at most $d-1$ (recall that $p-t$ itself is of degree at most d for the leading terms cancel). □

§ 2 Divided differences

From now on, unless otherwise stated, we work with sets of interpolation points for which repetition is allowed. Thus the interpolation points are not necessarily distinct and every interpolation polynomial must be understood as a Lagrange-Hermite interpolation polynomial.

2.1 Definition as leading coefficients

Let $A = \{a_0, \dots, a_d\}$ and f a function for which $\mathbf{L}_A(f)$ is well defined. The coefficient of x^d in $\mathbf{L}_A(f)$ is called the $d-th$ **divided difference** of f at A and is denoted by $f[a_0, \dots, a_d]$ or simply $f[A]$. If A corresponds to (B, n), $B = \{b_0, \dots, b_k\}$, then it follows from (1.10) that

$$f[a_0, \dots, a_d] = \sum_{i=0}^{k} \sum_{j=0}^{n(j)-1} f^{(j)}(b_i) \times \left(\text{coef. of } x^d \text{ in } l_{ij}\right).$$

[3] See e.g. (Davis 1975, P. 152) or (Nurnberger 1989, p. 32)

[4] Since we work with t_{d+1}, we have $m_k = \cos\frac{k\pi}{d+1}$, $k = 0, \dots, d+1$.

In particular, if the points are pairwise distinct, the Lagrange interpolation formula (1.4) gives

$$f[a_0, \ldots, a_d] = \sum_{i=1}^{d} f(a_i) \prod_{j \neq i} \frac{1}{a_i - a_j}. \tag{1.18}$$

Thus, $f[a_0] = f(a_0)$ and if $a_0 \neq a_1$,

$$f[a_0, a_1] = \frac{f(a_1) - f(a_0)}{a_1 - a_0}.$$

Notice that the map $f \to f[a_0, \ldots, a_d]$ is linear.

2.2 Newton's Formula

Theorem 1.8.

$$\mathbf{L}[a_0, \ldots, a_d; f](x) = \sum_{i=0}^{d} f[a_0, \ldots, a_i](x - a_0) \ldots (x - a_{i-1}) \tag{1.19}$$

(An empty product is taken to be 1.*)*

Proof. The formula is clearly true for $d = 0$. Assume it is also true for $d = k$ and take $d = k + 1$. The polynomial

$$p(x) = \mathbf{L}[a_0, \ldots, a_d; f] - f[a_0, \ldots, a_d](x - a_0) \ldots (x - a_{d-1})$$

is of degree $d - 1$ for the leading terms cancel. Moreover it interpolates f (in the Hermite sense) at $\{a_0, \ldots, a_{d-1}\}$. Hence, in virtue of the uniqueness of the interpolation polynomial, we have $p = \mathbf{L}[a_0, \ldots, a_{d-1}; f]$ and

$$\mathbf{L}[a_0, \ldots, a_d; f] = \mathbf{L}[a_0, \ldots, a_{d-1}; f] + f[a_0, \ldots, a_d](x - a_0) \ldots (x - a_{d-1}).$$

The theorem follows using the induction hypothesis. □

2.3 A second remainder formula

Theorem 1.9.

$$f(x) - \mathbf{L}[a_0, \ldots, a_d; f](x) = f[a_0, \ldots, a_d, x](x - a_0) \ldots (x - a_d).$$

Proof. Put $A' = \{a_0, \ldots, a_d, x\}$. Applying Newton's formula (1.19), we have

$$\mathbf{L}[A'; f](t) = \mathbf{L}[A; f](t) + f[a_0, \ldots, a_d, x](t - a_0) \ldots (t - a_d).$$

We get the result by taking $t = x$ for $\mathbf{L}[A'; f](x) = f(x)$. □

Comparing this result with Theorem 1.6 we obtain the following corollary.

Corollary. *If I is a (real) interval containing $A = \{a_0, \ldots, a_d\}$ and $f \in \mathsf{C}^{d+1}(I)$ then there exists $\xi \in I$ such that*

$$f[A] = \frac{f^{(d+1)}(\xi)}{d!}.$$

A more precise relation connecting the divided difference and the derivatives of a function, the so-called Hermite-Genocchi formula, will be presented in § 3.

2.4 Three key properties

Conveniently combined, the following three properties enable to recursively compute every divided difference once we know the values of the function at the points of A and some of its derivatives in case of repeated points.

Theorem 1.10. *Let $d \geq 0$ and f a sufficiently smooth function.*

(i) $f[a_0, \dots, a_d]$ *is a symmetric function of the* a_i*'s.*

(ii) If $a_0 = a_1 = \cdots = a_d$ *then* $f[a_0, \dots, a_d] = \frac{f^{(d)}(a_0)}{d!}$.

(iii) If $a_0 \neq a_d$ *then*

$$f[a_0, \dots a_d] = \frac{f[a_0, \dots, a_{d-1}] - f[a_1, \dots, a_d]}{a_0 - a_d}. \tag{1.20}$$

Proof. (i) By definition, $f[a_0, \dots, a_d]$ depends only on $A = \{a_0, \dots, a_d\}$ and the way the points are ordered is therefore irrelevant.

(ii) In this case, $\mathbf{L}[A; f]$ is the Taylor polynomial of f at a_0 to the order d the leading coefficient of which is $f^{(d)}(a_0)/d!$.

(iii) This is a consequence of the Neville-Aitken formula (1.13) in which we compare the leading terms of both sides. □

For. (1.20) can be written as

$$f[A] = \frac{f[A \setminus \{a_d\}] - f[A \setminus \{a_0\}]}{a_0 - a_d}$$

and, in view of the symmetry proved in the first point, we have for $i \neq j$,

$$f[A] = \frac{f[A \setminus \{a_i\}] - f[A \setminus \{a_j\}]}{a_j - a_i}. \tag{1.21}$$

Here, $A \setminus \{a_j\}$ denotes the set obtained by deleting *one* (and only one) occurrence of a_j. Thus, if the multiplicity of a_j is greater than 1, one still has $a_j \in A \setminus \{a_j\}$.

2.5 The Leibniz formula

This is a very elegant formula for the divided differences of a product. We shall have no occasion to apply it[5] but something similar to the main idea of the proof will be used later in a different context.

Theorem 1.11.

$$(f \cdot g)[a_0, \dots, a_d] = \sum_{j=0}^{d} f[a_0, \dots, a_j]\, g[a_j, \dots, a_d]. \tag{1.22}$$

[5]The Leibniz formula has been used for obtaining recurrence relations for B-splines functions, see (Nurnberger 1989, p. 101).

Let $A = \{a_0, a_1, \ldots, a_d\}$. We define

$$A_i = \{a_0, a_1, \ldots, a_i\} \text{ and } A^j = \{a_{d-j}, a_{d-j+1}, \ldots, a_d\}.$$

The formula to be proved is

$$(f \cdot g)[A] = \sum_{j=0}^{d} f[A_j]\, g[A^{d-j}]. \tag{1.23}$$

Proof. We shall use Newton's formula for $p = \mathbf{L}_A(f)$ and $\tilde{p} := \mathbf{L}_A(g)$. However, for $\tilde{p}$, the interpolation points will be taken in the reverse order. Thus the divided differences to be used are $f[A_i]$ for p and $g[A^j]$ for $\tilde{p}$. Now, we have

$$p(x) = \sum_{i=0}^{d} f[A_i]\, N_i(x) \quad \text{and} \quad \tilde{p}(x) = \sum_{j=0}^{d} g[A^j]\, N^j(x)$$

where

$$N_i(x) = \prod_{l=0}^{i-1}(x - a_l) \quad \text{and} \quad N^j(x) = \prod_{l=d-j+1}^{d}(x - a_l).$$

We claim that $p \cdot \tilde{p}$ satisfies the same interpolation conditions as $\mathbf{L}_A(fg)$. In case of repeated points, this follows from the classical Leibniz formula for the computation of the derivatives of a product. The details are left to the reader. At this point, we may conclude that

$$p \cdot \tilde{p} \equiv \mathbf{L}_A(f \cdot g) \quad [w] \tag{1.24}$$

where $w(x) = (x - a_0) \ldots (x - a_d)$. On the other hand, we have

$$p \cdot \tilde{p} = \sum_{i,j=0}^{d} f[A_i] g[A^j]\, N_i N^j.$$

However, for $i + j > d$, $N_i N^j \equiv 0\ [w]$ so that

$$p \cdot \tilde{p} = \sum_{i+j \le d} f[A_i] g[A^j]\, N_i N^j \quad [w]. \tag{1.25}$$

We conclude from (1.24) and (1.25) that

$$\mathbf{L}_A(f \cdot g) \equiv \sum_{i+j \le d} f[A_i] g[A^j]\, N_i N^j \quad [w].$$

Now, both sides being of degree $\le d$ they must be equal and the theorem follows by comparing their leading terms. $\square$

§ 3 THE SIMPLEX FUNCTIONAL AND THE HERMITE-GENOCCHI FORMULA

3.1 Simplices

For every $d \ge 1$ the (convex compact) subset of $\mathbb{R}^{d+1}$

$$\Delta^d = \left\{ (t_0, t_1, \ldots, t_d) \in \mathbb{R}_+^{d+1} \,:\, \sum_{i=0}^{d} t_i = 1 \right\}$$

is called the d-th (standard) **simplex**. Let S_d denote the convex hull of the canonical basis in $\mathbb{R}^d$, that is,

$$S^d = \left\{ (u_1, \ldots, u_d) \in \mathbb{R}_+^d \ : \ \sum_{i=1}^{d} u_i \leq 1 \right\},$$

then Δ_d is the image of S_d under the affine mapping s_0 defined by

$$s_0 : (t_1, t_2, \ldots, t_d) \in S_d \longrightarrow \left(1 - \sum_{i=1}^{d} t_i, t_1, \ldots, t_d \right) \in \Delta_d.$$

Note that if $(v_0, v_1, \ldots, v_d)$ denotes the canonical basis of $\mathbb{R}^{d+1}$ then

$$s_0(t_1, \ldots, t_d) = v_0 + \sum_{i=1}^{d} t_i (v_i - v_0).$$

Recall that the Lebesgue measure dm on Δ_d is the unique (Borel) positive measure on Δ_d satisfying the relation

$$\int_{\Delta_d} F(t) dm(t) := \int_{S_d} (F \circ s_0)(t) dm(t) \quad (F \in \mathsf{C}(\Delta^d)), \tag{1.26}$$

where the measure dm on the right hand side refers to the usual d-dimensional Lebesgue measure. The following proposition shows that the Lebesgue measure on Δ_d inherits the symmetry properties of the Lebesgue measure on S_d.

Proposition 1.12. *The Lebesgue measure dm on Δ^d is invariant under any permutation of the variables.*

This means that if σ is any permutation on the indices $0, 1, \ldots d$ and u_σ is the linear map on $\mathbb{R}^{d+1}$ defined by $u_\sigma(v_i) = v_{\sigma(i)}$ $(i = 0, \ldots, d)$ then

$$\int_{\Delta_d} F(t) dm(t) = \int_{\Delta_d} F(u_\sigma t) dm(t) \quad (F \in \mathsf{C}(\Delta^d)). \tag{1.27}$$

Proof. We shall prove (1.27) when σ is the transposition $(0, j)$, $j = 1, 2, \ldots, d$, that is, the permutation which exchanges 0 with j leaving the others indices unchanged. This is sufficient for the transpositions $(0, j)$ $(j = 1, \ldots, d)$ form a set of generators for the group of permutations. Since the proof is similar for every value of j, we shall only treat the case $j = 1$. Iin view of (1.26), when $\sigma = (0, 1)$, (1.27) reduces to

$$\begin{aligned} &\textstyle\int_{S_d} F\big(1 - \sum_{i=1}^{d} t_i, t_1, t_2, \ldots, t_n)\big) dm(t) \\ &\qquad = \textstyle\int_{S_d} F\big(u_1, 1 - \sum_{i=1}^{d} u_i, u_2, u_3, \ldots, u_n)\big) dm(u). \end{aligned} \tag{1.28}$$

Now, one readily verifies that (1.28) is but the change of variable formula used with $t = \psi(u)$,

$$\psi : \begin{cases} t_j = u_j & j = 2, \ldots, d \\ t_1 = 1 - \sum_{i=1}^{d} u_i \end{cases},$$

since $\psi(S_d) = S_d$ and $|J_\psi| = 1$. $\square$

[1.29]

The next proposition gives a useful way of computing integrals on S_d and hence on Δ_d.

Proposition 1.13. *For $F \in \mathsf{C}(S_d)$, we have*

$$\int_{S_d} F(t)dm(t) = \int_0^1 \int_0^{u_1} \cdots \int_0^{u_{d-1}} F(u_1 - u_2, \dots, u_{d-1} - u_d, u_d) du_1 \dots du_d. \quad (1.29)$$

Proof. This is given by (Fubini's Theorem and) the change of variable $t = \xi(u)$ with

$$\xi : \begin{cases} t_i = u_i - u_{i+1} & i = 1, \dots, d-1 \\ t_d = u_d \end{cases}.$$

Indeed $t \in \Delta^d$ if and only if $u \in \{0 \leq u_d \leq u_{d-1} \leq \cdots \leq u_2 \leq u_1 \leq 1\}$ and $J_\xi = 1$. □

Applying this result with $F = 1$, we get

Corollary. $m(\Delta_d) = m(S_d) = \frac{1}{d!}$.

3.2 The simplex functional

Let $a_0, \dots, a_d$ be $d+1$ (not necessarily distinct) points in $\mathbb{K}$, Ω the convex hull of these points and $f \in \mathsf{C}(\Omega)$. We define

$$\int_{[a_0,\dots,a_d]} f := \int_{\Delta_d} f(t_0 a_0 + t_1 a_1 + \cdots + t_d a_d) dm(t) \quad (1.30)$$

with the convention that, when $d = 1$,

$$\int_{[a_0]} f := f(a_0).$$

In view of Proposition 1.12, the result does not depend on the particular ordering of the points, hence it is a symmetric function of the a_i's or equivalently, a function of $A := \{a_0, a_1, \dots, a_d\}$ and we may write

$$\int_{[a_0,\dots,a_d]} f = \int_{[A]} f.$$

The (linear continuous) map

$$f \in \mathsf{C}(\Omega) \longrightarrow \int_{[A]} f \in \mathbb{K}$$

is called the **simplex functional** (attached to A) and is to play a fundamental role in this text.

3.3 The recurrence formula for the simplex functional

Theorem 1.14. *Let $A = \{a_0, a_1, \dots, a_d\} \subset \mathbb{K}$. For $f \in \mathsf{C}^1(\mathrm{cv}(A))$ we have*

$$(a_i - a_j) \int_{[A]} f' = \int_{[A \setminus \{a_j\}]} f - \int_{[A \setminus \{a_i\}]} f. \quad (1.31)$$

Proof. If $a_i = a_j$ there is nothing to prove so we shall assume that $a_i \neq a_j$. Moreover, since the simplex function is a symmetric function of the points, we may restrict ourselves to the case $i = d$ and $j = d-1$. The case $d = 1$ is easy and illuminating. Indeed

$$\begin{aligned}(a_1 - a_0)\int_{[a_0,a_1]} f' &= \int_0^1 (a_1 - a_0) f'(a_0 + t(a_1 - a_0))dt \\ &= f(a_1) - f(a_0) = \int_{[a_1]} f - \int_{[a_0]} f.\end{aligned}$$

We shall now assume $d \geq 2$. By definition, we have

$$\int_{[A]} f' = \int_{S_d} f'\big(a_0 + \sum_{i=1}^{d} t_i(a_i - a_0)\big) dm(t). \tag{1.32}$$

We claim that

$$\int_{[A]} f' = \int_0^1\int_0^{u_1} \cdots \int_0^{u_{d-1}} f'\big(a_0 + \sum_{j=1}^{d} u_j(a_j - a_{j-1})\big) du_1 \dots du_d. \tag{1.33}$$

This follows from Proposition 1.13 once we observed that

$$\begin{aligned}&F(t_1, \dots, t_d) = f'\big(a_0 + t_1(a_1 - a_0) + \dots + t_d(a_d - a_0)\big)\\ &\quad\Longrightarrow F(u_1 - u_2, \dots, u_{d-1} - u_d, u_d) = f'\big(a_0 + \textstyle\sum_{j=1}^{d} u_j(a_j - a_{j-1})\big).\end{aligned}$$

Now, setting $K = a_0 + \sum_{j=1}^{d-2} u_j(a_j - a_{j-1})$, we have

$$\begin{aligned}&(a_d - a_{d-1}) \int_0^{u_{d-1}} f'\big(K + u_{d-1}(a_{d-1} - a_{d-2}) + u_d(a_d - a_{d-1})\big)du_d\\ &\quad= \big[f\big(K + u_{d-1}(a_{d-1} - a_{d-2}) + u_d(a_d - a_{d-1})\big)\big]_0^{u_{d-1}}\\ &\quad= f\big(K + u_{d-1}(a_d - a_{d-2})\big) - f\big(K + u_{d-1}(a_{d-1} - a_{d-2})\big).\end{aligned}$$

Reporting this in the left hand side of (1.33) we may eliminate the integral term in the variable u_d and, using again Proposition 1.13, go back to the simplex functional thereby obtaining the required formula. □

3.4 The Hermite-Genocchi formula

All the substance of the general Hermite-Genocchi formula (1.34) below is already contained in the recurrence relation for the simplex functional.

Theorem 1.15. *Let $A = \{a_0, a_1, \dots, a_d\} \subset \mathbb{K}$. For $f \in \mathsf{C}^d(\mathrm{cv}(A))$ we have*

$$f[A] = \int_{[A]} f^{(d)}. \tag{1.34}$$

Proof. We prove the formula by induction on d. If $d = 0$, the formula is trivially true. So we assume that it is true for $d-1$ and prove it for d. If all the points coincide then, since $m(\Delta_d) = 1/d!$, we have using Theorem 1.10 (ii)

$$\int_{[A]} f^{(d)} = f^d(a_0)m(\Delta_d) = \frac{f^d(a_0)}{d!} = f[A].$$

[1.34]

Let us assume that, for some i and j, we have $a_i \neq a_j$. Using successively the recurrence relation for the simplex functional, the induction hypothesis and Theorem 1.10 (iii) — in the form (1.21), we obtain

$$\begin{aligned}\int_{[A]} f^{(d)} &= \frac{\int_{[A\setminus\{a_i\}]} f^{(d-1)} - \int_{[A\setminus\{a_j\}]} f^{(d-1)}}{a_j - a_i} \\ &= \frac{f[A\setminus\{a_i\}] - f[A\setminus\{a_j\}]}{a_j - a_i} \\ &= f[A].\end{aligned}$$

□

3.5 The interpolation polynomials as functions of the nodes

Thanks to the Hermite-Genocchi formula, Newton's formula (1.19) can be rewritten as

$$\mathbf{L}[a_0, \ldots, a_d; f](x) = \sum_{k=0}^{d} \int_{[a_0,\ldots,a_k]} f \;\; (x - a_0)\ldots(x - a_{k-1}). \tag{1.35}$$

The next theorem follows immediately with the help of standard theorems on the continuity and differentiability of functions defined by an integral. We only state it in the real case.

Theorem 1.16. *Let I be an interval and $f \in \mathsf{C}^d(I)$, the map*

$$(a_0, \ldots, a_d) \in I^{d+1} \longrightarrow \mathbf{L}[a_0, \ldots, a_d; f] \in \Pi_d(\subset \mathsf{C}(I))$$

is continuous and even of class C^k if $f \in \mathsf{C}^{d+k}(I)$.

In particular, if $(a_0^s, \ldots, a_d^s) \to 0$ as $s \to \infty$ then

$$\mathbf{L}[a_0^s, \ldots, a_d^s; f] \to \mathbf{T}_0^d(f) \qquad (s \to \infty)$$

where $\mathbf{T}_0^d(f)$ is the Taylor polynomial of f at the origin and to the order d. More generally any Lagrange-Hermite interpolation polynomial of a sufficiently smooth function is the limit of a sequence of Lagrange interpolation polynomials.

§ 4 HERMITE'S FORMULAS FOR HOLOMORPHIC FUNCTIONS.

4.1 Cauchy's formulas

Let Ω be a simply connected domain in the complex plane and $f \in \mathsf{H}(\Omega)$ that is, f is holomorphic (analytic) on Ω. Recall that if γ is a piecewise C^1 curve in Ω that turns once (in the positive sense) around $a \in \Omega$, that is to say $\mathrm{Ind}_\gamma(a) = 1$ where

$$\operatorname*{Ind}_{\gamma}(a) := \frac{1}{2i\pi} \int_\gamma du/(u - a),$$

the Cauchy formula is

$$f(a) = \frac{1}{2i\pi} \int_\gamma \frac{f(u)}{u - a} du \tag{1.36}$$

and, more generally, for $n \in \mathbb{N}^\star$,

$$\frac{f^{(n)}(a)}{n!} = \frac{1}{2i\pi}\int_\gamma \frac{f(u)}{(u-a)^{n+1}}du. \tag{1.37}$$

4.2 Hermite's formula for divided differences

Theorem 1.17. *Let Ω, γ and f as above. If $\{a_0, \dots, a_d\} \subset \Omega$ and $\mathrm{Ind}_\gamma(a_i) = 1$, $i = 0, \dots, d$ then one has*

$$f[a_0, \dots, a_d] = \frac{1}{2i\pi}\int_\gamma \frac{f(u)}{(u-a_0)\dots(u-a_d)}du. \tag{1.38}$$

Proof. The proof is by induction, in the same manner as we did in the proof of Theorem 1.15. By Cauchy's formula (1.36), since $f[a_0] = f(a_0)$, (1.38) is true for $d = 0$. Let us assume that (1.38) is true for $d = k$ and let us prove it for $d = k+1$. If all the points coincide (1.37) yields the result, otherwise, the formula to be proved being symmetric, we may assume that $a_0 \neq a_d$. An elementary calculation then shows

$$\frac{a_0 - a_d}{(u-a_0)\dots(u-a_d)} = \frac{1}{(u-a_0)\dots(u-a_{d-1})} - \frac{1}{(u-a_1)\dots(u-a_d)}.$$

Multiplying by $f(u)$, integrating over γ and using both the induction hypothesis and the recurrence relation (1.20), we get

$$\begin{aligned}\frac{1}{2i\pi}\int_\gamma \frac{(a_0-a_d)f(u)du}{(u-a_0)\dots(u-a_d)} &= f[a_0, \dots, a_{d-1}] - f[a_1, \dots, a_d] \\ &= (a_0 - a_d)f[a_0, \dots, a_d].\end{aligned}$$

□

4.3 Hermite's Remainder Formula

Theorem 1.18 (Hermite). *Under the same assumption as in Theorem 1.17 and if moreover $\mathrm{Ind}_\gamma(z) = 1$ then one has*

$$f(z) - \mathbf{L}[a_0, \dots, a_d; f](z) = \frac{1}{2i\pi}\int_\gamma \frac{(z-a_0)\dots(z-a_d)}{(u-a_0)\dots(u-a_d)}\frac{f(u)du}{u-z}. \tag{1.39}$$

Proof. It follows immediately from the conjunction of Theorems 1.9 and 1.17. □

As a corollary, we obtain an integral formula for the interpolation polynomial.

Corollary. *Under the same assumptions, setting $w(u) = (u-a_0)\cdots(u-a_d)$ we have*

$$\mathbf{L}[a_0, \dots, a_d; f](z) = \frac{1}{2i\pi}\int_\gamma \frac{w(u)-w(z)}{w(u)(u-z)}f(u)du. \tag{1.40}$$

Proof. In (1.39), replace $f(z)$ by the right-hand side of (1.36) (in which we take $a = z$). □

[1.40]

This result gives a way to find a Lagrange-like interpolation formula for $\mathbf{L}_A$ as in (1.10). Indeed, if A corresponds to (B, n) with say, $B = \{b_0, \ldots, b_k\}$ (see 1.2 for the definition) then the rational fraction

$$R_z(u) := \frac{w(u) - w(z)}{w(u)(u-z)}$$

can be written as

$$R_z(u) = \sum_{i=0}^{k} \sum_{j=1}^{n(j)} \frac{l_{ij}(z)}{(u-b_i)^j}$$

where the l_{ij} are polynomials in z of degree at most d (the details of the proof are left to the reader). Now, plugging this expression of R_z in (1.40) and using Cauchy's formula, we obtain

$$\mathbf{L}[a_0, \ldots, a_d; f](z) = \sum_{i=0}^{k} \sum_{j=0}^{n(j)-1} \frac{f^{(j)}(b_i)}{j!} l_{i,j+1}(z) \tag{1.41}$$

so that the polynomial ℓ_{ij} of (1.10) are given by $\ell_{ij} = l_{i,j+1}/j!$.

4.4 A glance to complex approximation

Hermite's remainder formula is the starting point of the rich and beautiful theory of Lagrange interpolation of holomorphic functions which uses deep connections between complex approximation theory, plane potential theory and geometric function theory. The following result is a very poor illustration but is easy to prove and will be sufficient in the sequel to illustrate a strong contrast between univariate and multivariate Lagrange interpolation.

Theorem 1.19. *Let f be an entire function ($f \in H(\mathbb{C})$) and let, for every $d \in \mathbb{N}$, $A^d := \{a_0^d, \ldots, a_d^d\}$ be a set of $d+1$ not necessarily distinct points. If $A := \cup_{d=0}^{\infty} A^d$ is bounded then $\mathbf{L}_{A^d}(f)$ converges uniformly to f on every compact subset of the plane.*

Proof. Let K be a plane compact set. We choose $R > 0$ such that

$$A \cup K \subset \{|z| \le R\}.$$

We shall apply Hermite's remainder formula (1.39) on taking $\gamma : t \in [0, 2\pi] \to 4Re^{it}$. Thus $\gamma(t)$ moves along the circle centered at the origin and of radius $4R$. Of course, for every $y \in A \cup K$ we have $\mathrm{Ind}_\gamma(y) = 1$ and the use of (1.39) is legitimate. Now, observe that if $z \in K$ and $a \in A$ then $|z - a| \le 2R$ and if $u \in \gamma([0, 2\pi])$ and $a \in A$ or $a = z$ then $|u - a| \ge |u| - |a| \ge 3R$. Hence, for a fixed $z \in K$, Hermite's remainder formula gives

$$|f(z) - \mathbf{L}_{A^d}(f, z)| \le \frac{1}{2\pi} \cdot \mathrm{length}\,(\gamma) \cdot \left|\left| \frac{(z-a_0)\ldots(z-a_d)}{(u-a_0)\ldots(u-a_d)} \frac{f(u)}{u-z} \right|\right|_{\gamma([0,2\pi])}$$

which implies

$$||f(z) - \mathbf{L}_{A^d}(f, z)||_K \le 4R \cdot \left(\frac{2R}{3R}\right)^{d+1} \cdot \frac{||f||_{\gamma([0,2\pi])}}{3R} \longrightarrow 0 \quad (d \to \infty).$$

This proves the uniform convergence on K. □

§ 5 Lebesgue's function and constant

In this section we return to Lagrange interpolation. Thus the interpolation points a_i are pairwise distinct and $\mathbf{L}_A(f)$ is well defined for every function defined on $A = \{a_0, a_1, \ldots, a_d\}$. In particular, if I is a compact interval, $I = [a, b]$, $\mathbf{L}_A$ defines a linear operator on the Banach space $(\mathsf{C}(I), \|\cdot\|_I)$. We shall study the continuity of this operator. When there is no danger of confusion, we shall write $||f||$ instead of $||f||_I$.

5.1 $\mathbf{L}_A$ as a continuous operator

Let $A = \{a_0, \ldots, a_d\} \subset I$ be a set of $d+1$ distinct points. It follows from Lagrange's interpolation formula (1.4) that

$$|\mathbf{L}_A(f, x) - \mathbf{L}_A(\tilde{f}, x)| \leq \max_{i=0,\ldots,d} |f(a_i) - \tilde{f}(a_i)|. \sum_{i=0}^{d} |\ell_i(x)|. \tag{1.42}$$

The function $\delta_A := \sum_{i=0}^{d} |\ell_i|$ is called the **Lebesgue function** of A and its sup-norm $||\delta_A||_I$ denoted by $\Delta(A, I)$ (— or $\Delta(A)$ or Δ — is the **Lebesgue constant** of A (and I).

It follows from (1.42) that for every continuous functions f and $\tilde{f}$, one has

$$||\mathbf{L}_A(f) - \mathbf{L}_A(\tilde{f})||_I \leq \Delta(A, I)||f - \tilde{f}||_I, \tag{1.43}$$

This shows that $\mathbf{L}_A$ is a continuous linear operator. From a numerical point of view, we may say that δ_A and $\Delta(A)$ measure the **stability** of the interpolation process at A. In practice, if f is the *ideal function*, usually one merely knows some approximation $\tilde{f}$ of f. Inequality (1.43) shows how the interpolating polynomial is modified when f is replaced by $\tilde{f}$. Consequently, if we are free to choose A, we have interest to choose it in order that Δ_A be as small as possible.

5.2 The norm of $\mathbf{L}_A$

Recall that the norm of a continuous linear operator U on $\mathsf{C}(I)$ is defined by

$$|||U||| := \sup_{f \neq 0} \frac{||Uf||}{||f||} = \sup_{||f||=1} ||Uf|| \,.$$

Inequality (1.43) shows that $|||\mathbf{L}_A||| \leq \Delta(A, I)$. Equality actually holds true.

Theorem 1.20. *$\Delta(A, I)$ is the norm of the continuous linear operator $\mathbf{L}_A$: $f \in \mathsf{C}(I) \to \mathbf{L}_A(f) \in \mathsf{C}(I)$, that is $\Delta(A, I) = |||\mathbf{L}_A|||$.*

Proof. To prove the remaining inequality, it suffices to find a function f_0 such that $||f_0||_I = 1$ and $\|\mathbf{L}_A(f_0)\| \geq \delta(A, I)$. Take $y_0 \in I$ such that $\delta_A(y_0) = \Delta(A, I)$. We shall choose as f_0 any function satisfying

(i) $||f_0||_I = 1$ and

(ii) $f_0(a_i) = \mathrm{sign}(\ell_i(y_0))$, $i = 0, 1, \ldots, d$.

[1.43]

Such a function is easily seen to exist. Now,

$$\begin{aligned}\Delta(A,I) &= \sum_{i=0}^{d} \mid \ell_i(y_0) \mid = \sum_{i=0}^{d} \operatorname{sign}(\ell_i(y_0)) \cdot \ell_i(y_0) = \sum_{i=0}^{d} f_0(a_i)\ell_i(y_0) \\ &= \mathbf{L}_A(f_0)(y_0) \leq \mid \mathbf{L}_A(f_0)(y_0) \mid \leq ||\mathbf{L}_A(f_0)||_I,\end{aligned}$$

and this concludes the proof. □

5.3 The affine invariance

The following theorem shows that the Lebesgue constant is invariant (in an appropriate sense) under affine transformations.

Theorem 1.21. *If $t : x \to \alpha x + \beta$ with $\alpha \neq 0$, $J = t(I)$ and $B = t(A)$ then*

$$\Delta(B,J) = \Delta(A,I). \tag{1.44}$$

Proof. This is a consequence of the previous invariance formula (1.11) (and of Theorem 1.20). Indeed, since the map $\phi : f \in \mathsf{C}(I) \to f \circ t \in \mathsf{C}(J)$ is bijective and isometric, we have

$$\begin{aligned}\|\mathbf{L}_B\| &= \sup_{\|f\|_J=1} \|\mathbf{L}_B(f)\|_J = \sup_{\|h\|_I=1} \left\|\mathbf{L}_B(h \circ t^{-1})\right\|_J \\ &= \sup_{\|h\|_I=1} \left\|\mathbf{L}_A(h) \circ t^{-1})\right\|_J = \sup_{\|h\|_I=1} \|\mathbf{L}_A(h)\|_I \\ &= \|\mathbf{L}_A\|.\end{aligned}$$

□

5.4 Lebesgue's inequality

Theorem 1.22. *Let $f \in \mathsf{C}(I)$ and A a set of $d+1$ distinct points in I, we have*

$$||f - \mathbf{L}_A(f)||_I \leq (1 + \Delta(A,I))\operatorname{dist}(f, \mathcal{P}_d), \tag{1.45}$$

where $\operatorname{dist}(f, \mathcal{P}_d) = \inf\{||f-p||_I, p \in \mathcal{P}_d\}$.

This means that if, for example, $\Delta(A,I)$ is ≤ 9 then in using $\mathbf{L}_A(f)$ to approximate f rather than its polynomial of best approximation we loose only a precision of one decimal.

Proof. It uses the fact that $p \in \mathcal{P}_d \implies \mathbf{L}_A(p) = p$. Take $p_{opt} \in \mathcal{P}_d$ such that $||f - p_{opt}|| = dist(f, \mathcal{P}_d)$ then, since $f - \mathbf{L}_A(f) = (f - p_{opt}) - \mathbf{L}_A(f - p_{opt})$, we have using (1.43),

$$\begin{aligned}||f - \mathbf{L}_A(f)|| &\leq ||f - p_{opt}|| + ||\mathbf{L}_A(f - p_{opt})|| \\ &\leq ||f - p_{opt}|| + \Delta(A,I)||f - p_{opt}|| \\ &\leq (1 + \Delta(A,I))||f - p_{opt}|| \\ &\leq (1 + \Delta(A,I))\operatorname{dist}(f, \mathcal{P}_d),\end{aligned}$$

and the inequality is proved. □

5.5 Lebesgue's constants and sequences of interpolation polynomials

We shall limit ourselves to recall without proofs a few classical results.

For $f \in \mathsf{C}^k(I)$, the classical Weierstrass theorem says that $\operatorname{dist}(f, \mathcal{P}_d)$ tends to 0 as d goes to ∞. The first Jackson theorem is a precise version of this result.

Theorem 1.23 (Jackson). *For every continuous function f on $I = [a, b]$, one has*

$$\operatorname{dist}(f, \mathcal{P}_d) \leq M\omega\Big(f; \frac{b-a}{d}\Big)$$

where $\omega\big(f; \delta\big)$ denotes the modulus of continuity of f, that is

$$\omega(f; \delta) := \sup\{|f(t) - f(t')| : t, t' \in I \text{ and } |t - t'| \leq \delta\},$$

and M is a constant independent of f, d and I.

Suitably repeated applications of this results leads to the second Jackson theorem .

Theorem 1.24 (Jackson). *Let $f \in \mathsf{C}^k(I)$ with $I = [a, b]$ and $k \geq 0$, then one has*

$$\operatorname{dist}(f, \mathcal{P}_d) \leq M^{k+1} \frac{(b-a)^k}{d(d-1)\ldots(d-k+1)} \omega\Big(f^{(k)}; \frac{b-a}{d-k}\Big)$$

where M is the same constant as the previous theorem.

It follows from Theorem 1.24 that for $f \in \mathsf{C}^k(I)$, $\operatorname{dist}(f, \mathcal{P}_d) = o(d^{-k})$ as $d \to \infty$). Hence if $A^d = \{a_0^d, \ldots, a_d^d\}$ is such that $\Delta(A^d, I) = 0(d^k)$ then, in view of Lebesgue's inequality (1.45), we have

$$\lim_{d \to \infty} \mathbf{L}_{A^d}(f) = f \quad \text{uniformly on } I$$

for every $f \in \mathsf{C}^k(I)$. However, unfortunately, sequences of nodes (A^d) with a Lebesgue constant which grows as slowly as a polynomial are not easy to find. From this point of view, as shown by the following theorem, the equidistributed points, a very natural choice, are not good points (at least if we want to use interpolants of high degree).

Theorem 1.25. *If $E^d = \{k/d : k = 0, \ldots, d\}$, $d = 1, 2, \ldots$, is the array of equidistributed points in $[0, 1]$ then*

$$\Delta(E^d, [0, 1]) \sim \frac{2^{d+1}}{ed \log d} \quad (d \to \infty).$$

The following theorem together with the Banach-Steinhaus Theorem (and Theorem 1.20) shows that whatever (A^d) is, there always exists a continuous function f (and, in fact, plenty of them) for which $\mathbf{L}_{A^d}(f)$ does not converge uniformly to f. It is even possible to construct continuous functions f for which $\mathbf{L}_{A^d}(f)$ has a very chaotic behavior. Many ingenious mathematicians in the last decades have devoted a great part of their energy to constructing such functions.

[1.45]

Theorem 1.26 (Bernstein). *Let, for $d \in \mathbb{N}^\star$, $A^d = \{a_0^d, a_1^d, \dots a_d^d\} \subset I$. One has*

$$\liminf_{d\to\infty} \frac{\Delta(A^d)}{\ln d} > 0.$$

This result is optimal in the sense that there exists sequences (A^d) the Lebesgue constant of which grows as $\ln d$. The most important of such sequences is undoubtedly given by $A^d = T_{d+1}^{-1}(0)$ where T_{d+1} is the $(d+1)$-th Chebyshev polynomial. The classical result (nowadays much precised) is the following theorem of Bernstein.

Theorem 1.27 (Bernstein). *If $C^d = \{\cos(\frac{2k+1}{d+1}\frac{\pi}{2}) : k = 0, 1, \dots d\} \subset [-1, 1]$ then*

$$\Delta(C^d, [-1, 1]) \sim \frac{2}{\pi} \ln d \quad (d \to \infty).$$

As a corollary, in view of Theorem (1.23), we obtain that if $\lim_{d\to\infty} \ln d \cdot \omega(f, 2/d)) = 0$ then $\mathbf{L}[C^d; f]$ converges to f uniformly on $[-1, 1]$. The condition holds true, for instance, when $f \in \mathsf{Lip}_1$, that is, satisfies the **Lipschitz condition** of order 1, $|f(x) - f(x')| \le M\,|x - x'|$.

Note

Everything in this chapter is very classical and, with perhaps the exception of § 3 and § 4, is part of any introductory course on numerical analysis or approximation theory.

The interpolation polynomial was used as early as 1686 by Newton (in what is now called the Newton form). Strangely enough, the Lagrange formula does not seem to have appeared in print before Waring in 1776. Lagrange himself used it in 1795. The classical error estimate of Theorem 1.6 seems to be due to Cauchy (1840). The modern proof was established independently by Genocchi in 1878, Schwarz and Stieltjes (both in 1882), see (Peano 1957, p. 441). Hermite interpolation was first considered in the case of holomorphic functions in the seminal paper (Hermite 1878). The author established the remainder formula for holomorphic functions as well as its corollaries. The main motivation of Hermite was to unify the theories of Taylor approximation on one side and Lagrange interpolation on the other side by presenting them as particular cases of a more general procedure. He readily understood that something similar could be done in the real variable case and, in a addenda to his work, he included a study of the real case which he concluded with a form of the Hermite-Genocchi formula.

For both an original presentation of divided differences and many interesting historical information we refer to the recent survey paper of Carl de Boor (2005).

The modern study of univariate polynomial interpolation may be divided in three chapters all of them very ancient but, it seems, still far from being exhausted. The first one is concerned with the effective computation of the interpolation polynomials that is, studies the (numerical) properties of the various available algorithms. For recents examples (in defense of the barycentric Lagrange interpolation formula) the reader might consult (Berrut & Trefethen 2005) and (Higham 2004). The second chapter studies interpolation in the realm of real analysis. It consists essentially in investigating the properties of the Lebesgue functions and related objects and in studying the approximation properties of the Lagrange-Hermite interpolants at specific sets of nodes, typically the roots of well-studied orthogonal polynomials. We refer the reader to (Brutman 1997) and to (Szabados & Vertesi 1990) to know more on this field. A recent and rather elementary treatment can be found in (Phillips 2003). The last of the three main chapters deals with constructive complex approximation. For a classical account of Lagrange interpolation of holomorphic functions, we refer to (Walsh 1969) and (Lebedev & Smirnov 1968). A more recent, but less complete, treatment can be found in (Gaier 1987). A remarkable study of polynomial interpolation, focusing on approximation of entire functions and applications to number theory, is (Guelfond 1963).

INDEX

BIBLIOGRAPHY

Berrut, Jean-Paul & Lloyd N. Trefethen (2005), 'Barycentric Lagrange interpolation', *Siam review* **?**, ?

Brutman, Lev (1997), 'Lebesgue functions for polynomial interpolation - a survey', *Annals of numerical mathematics* **4**, 111–127.

Davis, Philip J. (1975), *Interpolation and Approximation*, Dover, New York.

de Boor, Carl (2005), 'Divided differences', *Surveys in Approximation Theory* **1**, 46–69. [Online article at] http://www.math.technion.ac.il/sat.

Gaier, Dieter (1987), *Lectures on complex approximation*, Birkhauser, Boston.

Guelfond, Aleksandr Osipovich (1963), *Calcul des différences finies*, Dunod, Paris.

Hermite, Charles (1878), 'Sur la formule d'interpolation de Lagrange', *Jurnal fur reine und andgewandke mathematik* **84**, 70–79.

Higham, Nicholas J (2004), 'The numerical stability of barycentric Lagrange interpolation', *IMA Journal of Numerical Analysis* **24**, 547–556.

Lebedev, N. A. & V. I. Smirnov (1968), *Functions of a complex variable (Constructive Theory)*, M.I.T Press.

Nurnberger, Gunther (1989), *Approximation by spline functions*, Springer-Verlag, Berlin.

Peano, Giuseppe (1957), *Opere scelte, vol.1*, Edizione Cremonese, Roma.

Phillips, George M. (2003), *Interpolation and Approximation by Polynomials*, CMS books in mathematics, 14, Springer-Verlag, New-York.

Szabados, J. & P. Vertesi (1990), *Interpolation of functions*, World scientific, Singapore.

Walsh, Joseph L. (1969), *Interpolation and approximation by rational functions in the complex domain*, number XX *in* 'Colloquium publications', 5th edn, American Mathematical Society, Providence, Rhode Island.

CONTENTS