BOL. SOC. BRAS. MAT., VOL. 10 Nº 2 (1979), 75-86

Polynomial interpolation

*Thomas Bloom**

Introduction.

What is today usually referred to as the Lagrange interpolating polynomial has also been credited by various authors to Newton, Waring or Euler (see [1], page 23 for a discussion of the conflicting opinions). It is certain that a version of the Lagrange interpolating procedure appeared in the work of Newton (references to the early original papers may be found in [12], Chapter 1). The case of multiple points was handled by Hermite in 1878.

The question of polynomial interpolation for functions of several variables is quite natural and important. Results however, have not been entirely satisfactory (see the discussion in [6], for example). In fact, a good theory of divided differences for functions of several variables is still lacking (see §9).

In this paper we will present the result of Paul Kergin [4], [7], [8] which generalizes Lagrange interpolation to the case of functions of several variables (Theorem 5.2). This is a surprising result and one might at first wonder that such a generalization had not been discovered long ago. There are however a number of subtle points (see the discussion in 5.3).

The paper is organized as follows: §1 to §4 review the classical Lagrange and Hermite interpolation and the difficulties faced if one tries to generalize these results to the case of several variables. In §5 Kergin's interpolation is presented and a brief sketch of the proof is given in §7. In §6 methods of calculating the interpolating polynomial are given – in particular, a formula due to Micchelli-Milman [10].

The successful generalization of Lagrange interpolation to several variables emboldens one to look at various other interpolation and approximation questions in several variables. This is done in §8 and §9. These two sections are independent of the results in the previous sections of this paper.

---

*This paper is based on a lecture given by the author at I.M.P.A., Rio de Janeiro in November 1978. The author is supported by the National Research Council of Canada.

Recebido em abril de 1979.

1. Lagrange interpolation.

1.1 We will first review the classical Lagrange interpolation formula. It gives a polynomial (of one variable) of degree $\leq m$ which assumes given values at $(m+1)$ distinct points on the real line.

1.2. Theorem. *Let $A_0, \ldots, A_m$ be $(m+1)$ distinct points on the real line and let $b_0, \ldots, b_m$ be $(m+1)$ real numbers. There is a unique polynomial $p(x)$ of degree $\leq m$ such that*

$$(1.3) \qquad p(A_j) = b_j \quad \text{for} \quad j = 0, \ldots, m.$$

Proofs. There are several simple proofs available. In fact one can give a formula for p as follows.

Let $\delta_j(x)$ $(j = 0, \ldots, m)$ be the unique polynomial of degree m such that

$$(1.4) \qquad \delta_j(A_i) = \begin{cases} 0 & \text{for} \quad i \neq j \\ 1 & \text{for} \quad i = j. \end{cases}$$

Of course $\delta_j(x)$ is given by

$$(1.5) \qquad \delta_j(x) = \frac{\prod_{i \neq j} (x - A_i)}{\prod_{i \neq j} (A_j - A_i)}.$$

Then we have

$$(1.6) \qquad p(x) = \sum_{j=0}^{m} b_j \delta_j(x).$$

The uniqueness of p is clear since any polynomial of degree $\leq m$ vanishing at $(m+1)$ distinct points is identically zero.

A second method of proof is to consider the coefficients of $p(x)$ as unknowns and use (1.1) to solve for them. That is, we write

$$p(x) = c_m x^m + c_{m-1} x^{m-1} + \ldots + c_1 x + c_0.$$

Then (1.1) becomes

$$(1.7) \qquad c_m (A_j)^m + c_{m-1} (A_j)^{m-1} + \ldots + c_1 A_j + c_0 = b_j \text{ for } j = 0, \ldots, m.$$

We regard (1.5) as a set of linear equations in unknowns $c_0, c_1, \ldots, c_m$. Since the Van der Monde determinant

$$\det \begin{vmatrix} A_0^m \ldots A_0 & 1 \\ A_m^m \ldots A_m & 1 \end{vmatrix} \neq 0$$

the equations (1.7) have a unique solution.

2. Hermite interpolation.

2.1 We need not assume the points $A_0, \ldots, A_m$ distinct. If some of the points coincide there is still a unique polynomial interpolation procedure.

Suppose $A_0, \ldots, A_\mu$ are distinct points with multiplicities $m_0, \ldots, m_\mu$. The multiplicity is just a positive integer associated with a point which counts the number of times it occurs. We have $m_0 + \ldots + m_\mu = m$. We also regroup the values b_j as follows. We assume we are given real numbers $b_{j1}, \ldots, b_{jmj}$ for $j = 0, \ldots, \mu$.

2.2Theorem. *There is a unique polynomial $p(x)$ of degree $\leq m$ such that*

$$(2.3) \qquad \begin{aligned} P(A_j) &= b_{j1} & &\text{for} \quad j = 0, \ldots, \mu \\ \frac{d^{s-1}p}{dx^{s-1}}(A_j) &= b_{js} & &\text{for} \quad s = 1, \ldots, m_j \\ & & & \qquad j = 0, \ldots, \mu. \end{aligned}$$

Proof: The theorem may be proved by either of the two methods used in §1 to prove Lagrange interpolation [14, Chapter 1].

3. The interpolation problem in several variables.

3.1 We will use the following notation and terminology.

We let $x = (x_1, \ldots, x_n)$ be coordinates for $\mathbb{R}^n$. We will use the usual multi-index notation. Thus, for $I = (i_1, \ldots, i_n)$ a multi-index notation. We will use the usual multi-index notation. Thus for $I = (i_1, \ldots, i_n)$ a multi-index, $x^I = x^{i_1} \ldots x^{i_n}$, $I! = i_1! \ldots i_n!$ and $|I| = i_1 + \ldots + i_n$.

A polynomial of (total) degree $\leq m$ may be written uniquely in the form

$$(3.2) \qquad p(x) = \sum_{|I| < m} c_I x^I$$

where the c_I are real numbers.

3.3 We will consider the following interpolation problem in $\mathbb{R}^n$. Let $A_0, \ldots, A_m$ be $(m+1)$ distinct points of $\mathbb{R}^n$ and let $b_0, \ldots, b_m$ be $(m+1)$ real numbers. Find a polynomial $p(x)$ of (total) degree $\leq m$ such that

$$(3.4) \qquad p(A_j) = b_j \quad \text{for} \quad j = 0, \ldots, m.$$

The equations (3.4) are, of course, the generalization to $\mathbb{R}^n$ of the equations (1.3).

If one substitutes (3.2) into (3.4) and attempts to solve for the unknown coefficients C_I one has $(m+1)$ linear equations in $\binom{m+n}{m}$ unknowns. A solution always exists but if $n > 1$ it is never unique.

(3.5) Example. Consider $A_0 = (0,0)$ and $A_1 = (1,0)$ in $\mathbb{R}^2$ (with (x, y) as coordinates) and $b_0 = 1$, $b_1 = 3$. Then $m = 1$ and we must look for a polynomial $p(x, y) = c_{10}x + c_{01}y + c_{00}$ such that

$$(3.6) \qquad \begin{aligned} p(0,0) &= 1 \\ p(1,0) &= 3. \end{aligned}$$

Now, any linear function of the form $2x + c_{01}y + 1$ satisfies (3.6). The coefficient c_{01} may be chosen arbitrarily.

3.7 This dilemma and two remedies are discussed in Glaeser's article [6].

One approach is to interpolate $\binom{m+n}{m}$ values at $\binom{m+n}{m}$ points. For the corresponding equations to have a unique solution, i.e. to interpolate with a polynomial of (total) degree $\leq m$, the points must not be located on any algebraic variety of degree $\leq m$. This is a cumbersome hypothesis.

A second approach (Glaeser's schémas d'interpolation) consists of requiring the coefficients of p to satisfy certain linear relations in addition to (3.4). That is, one requires that p lie in a certain $(m+1)$ dimensional subspace of the space of all polynomials of degree $\leq m$. The subspace, and hence p, is not canonical but is chosen arbitrarily subject to the condition that it ensure uniqueness.

We will present Kergin's result which provides a canonical interpolating polynomial in the case of several variables – not exactly to the above interpolation problem but to a slight variation on that problem. To introduce it, we will return to the classical (one variable) case and once again look at Lagrange interpolation.

4. Lagrange interpolation of a function.

4.1 Again we begin with distinct points $A_0, \ldots, A_m$ on the real line. Instead of interpolating values $b_0, \ldots, b_m$ we try to interpolate values of a function $f(x)$ at $A_0, \ldots, A_m$. That is, we want the unique polynomial $p(x)$ of degree $\leq m$ such that

$$p(A_j) = f(A_j) \quad \text{for} \quad j = 0, \ldots, m. \tag{4.2}$$

These equations are identical to the equations (1.3). We have merely replaced b_j by $f(A_j)$. What then is the advantage of (4.2)?

In fact, by considering $p(x)$ to be interpolating the values of a function $f(x)$ one has a slightly different perspective which we will now exploit.

Assume $f(x)$ is defined for all x and is of class C^m (m-times continuously differentiable). For $A = (A_0, \ldots, A_m)$ we let $\mathscr{L}(A, f)$ denote the polynomial of degree $\leq m$ which satisfies (4.2), i.e. the Lagrange interpolating polynomial. Then the difference $\mathscr{L}(A, f) - f$ vanishes at the points $A_0, \ldots, A_m$. Applying the mean value theorem we have

$$\frac{d}{dx}(\mathscr{L}(A, f) - f)(\xi) = 0 \tag{4.3}$$

at (at least) m distinct points ξ. There is one such ξ in each of the intervals $[A_j, A_{j+1}]$ (we assume $A_0 < A_1 < \ldots < A_m$).

For $J \subset \{0, \ldots, m\}$ we denote by $[A_j]_{j\in J}$ the interval spanned by $\{A_j\}_{j\in J}$. Then, by repeated applications of the mean value theorem we have:

For every integer r, $1 \leq r \leq m$ and for every subset $J \subset \{0, \ldots, m\}$ with $\operatorname{card}(J) = r+1$ there exists a point $\xi \in [A_j]_{j\in J}$ such that

$$\frac{d^r}{dx^r}(\mathscr{L}(A, f) - f)(\xi) = 0. \tag{4.4}$$

Thus, in the case f is differentiable the Lagrange interpolating polynomial has, "for free", certain additional properties which may be summarized as follows. (We use the notation $\mathscr{C}^m(\mathbb{R})$ for the space of m-times continuously differentiable functions on $\mathbb{R}$ and $\mathscr{P}^m(\mathbb{R})$ for the space of polynomials of degree $\leq m$.)

4.5 Theorem. *Let $A_0, \ldots, A_m$ be distinct points of $\mathbb{R}$. There is a unique linear map $\mathscr{L} : \mathscr{C}^m(\mathbb{R}) \to \mathscr{P}^m$ such that*

(a) *$\mathscr{L}(f)(A_j) = f(A_j)$ for $j = 0, \ldots, m$. That is, $\mathscr{L}(f)$ interpolates f at the points $A_0, \ldots, A_m$.*

(b) *For every integer r, $1 \leq r \leq m$ and for every $J \subset \{0, \ldots, m\}$ with $\operatorname{card}(J) = r+1$ there exists a point $\xi \in [A_j]_{j\in J}$ such that*

$$\frac{d^r}{dx^r}\mathscr{L}(f)(\xi) = \frac{d^r}{dx^r}f(\xi).$$

Of course, $\mathscr{L}(f)$ is just the Lagrange interpolating polynomial, which, to emphasize its dependence on A, has also been denoted by $\mathscr{L}(A, f)$.

4.6 Remark. Theorem (4.5) is valid without assuming $A_0, \ldots, A_m$ distinct. In that case $\mathscr{L}(A, f)$ is merely the Hermite interpolating polynomial (see §2). If A_j has multiplicity m_j then (b) implies that

$$\frac{d^r}{dx^r}\mathscr{L}(A, f)(A_j) = \frac{d^r}{dx^r}f(A_j) \tag{4.7}$$

$$\text{for } r = 0, \ldots, m_j - 1.$$

5. Kergin interpolation.

5.1 The result of Kergin is a direct generalization of Theorem 4.5 and Remark 4.6 to the case of $\mathbb{R}^n$.

To state it we will use the following notation: $\mathscr{C}^m(\mathbb{R}^n)$ denotes the space of m-times continuously differentiable functions on $\mathbb{R}^n$; $\mathscr{P}^m(\mathbb{R}^n)$ denotes the space of polynomials of degree $\leq m$; for $\{A_j\}_{j\in J}$ a collection of points in $\mathbb{R}^n$ we denote by $[A_j]_{j\in J}$ the convex hull spanned by those points. For k an integer ≥ 1 we denote $Q^k(\mathbb{R}^n)$ the space of constant coefficient homogeneous linear differential operators of order k. Thus, any $q \in Q^k(\mathbb{R}^n)$ may be written uniquely in the form

(5.2) $q=\sum_{|I|=k} d_I \frac{\partial}{\partial x^I}$ where $\frac{\partial}{\partial x^I}=\left(\frac{\partial}{\partial x_1}\right)^{i_1}\cdots\left(\frac{\partial}{\partial x_n}\right)^{i_n}$

and the d_I are real numbers.

We note that in the case of one variable $Q^k(\mathbb{R})$ is spanned by the single operator $\left(\frac{d}{dx}\right)^k$.

5.2 Theorem. *Let $A_0, \ldots, A_m$ be $(m+1)$ points in $\mathbb{R}^n$. There is a unique linear map*

$$\chi : \mathscr{C}^m(\mathbb{R}^n) \to \mathscr{P}^m(\mathbb{R}^n)$$

such that

(a) *$\chi(f)(A_j)=f(A_j)$ for $j=0, \ldots, m$. That is, $\chi(f)$ interpolates f at the points $A_0, \ldots, A_m$.*

(b) *For every integer r, $1 \le r \le m$ and for every $J \subset \{0, \ldots, m\}$ with card$(J)=r+1$ and for every $q \in Q^r(\mathbb{R}^n)$ there exists a point $\xi \in [A_j]_{j\in J}$ such that*

$$q(\chi(f))(\xi) = q(f)(\xi).$$

Briefly then one has the same results as Theorem 4.5 and Remark 4.6 on replacing the interval spanned by points by the convex hull spanned by points and by replacing the differential operator $\left(\frac{d}{dx}\right)^r$ by the space of all constant coefficient linear homogeneous differential operators of that order.

We remark that to have uniqueness in Theorem 5.2 it is not sufficient merely to use the operators $\left(\frac{\partial}{\partial x}\right)^I$ with $|I|=r$ in statement (b).

5.3 Remarks. We have formulated Theorems 3.2 and 4.5 to emphasize the similarity between the cases $n=1$ and $n>1$. There are however certain important differences.

First, unlike the one variable case, there is no formula for Kergin's interpolating polynomial depending only on the values of f at the interpolating points. In fact, the coefficients of the polynomial $\chi(A, f)$ depend on integrals of derivatives of f over faces in the convex hull of the points $A_0, \ldots, A_m$ (see 6.5). The polynomial cannot, in general, be constructed unless f is differentiable. This makes the calculation of $\chi(A, f)$ quite difficult. In §6 we will give two methods for calculating $\chi(A, f)$. In §7 we will sketch a proof of Theorem 5.2.

6. Calculation of $\chi(A, f)$.

6.1 We will outline two methods of obtaining $\chi(A, f)$. The first one (see 6.2) is valid when f is itself a polynomial. The coefficients of $\chi(A,f)$ are then polynomials in the coefficients of f and the components of the points $\{A_j\}$. The second method (see 6.5) is valid for any $f \in \mathscr{C}^m(\mathbb{R}^n)$ and is a formula due to *P.* Milman (which was also discovered independently by C. Micchelli [10]). This formula may be used to establish the existence part of Theorem 5.2.

6.2 Let $A_0, \ldots, A_m$ be $(m+1)$ points (not necessarily distinct) in $\mathbb{R}^n$. Let $\lambda : \mathbb{R}^n \to \mathbb{R}$ be a linear function.

For s an integer ≥ 1 the calculation of the polynomial $\chi(A, \lambda^s)$ has a simple reduction to a one variable interpolation problem. It merely depends on Hermite interpolation of the function x^s at the points $\lambda(A_0), \ldots, \lambda(A_m)$. To be specific, letting $\lambda(A)$ denote $\{\lambda(A_j)\}_{j=0,\ldots,m}$ we have

(6.3) $$\chi(A, \lambda^s) = \mathscr{L}(\lambda(A), x^s) \circ \lambda.$$

Since any polynomial on $\mathbb{R}^n$ is a linear combination of functions of the form λ^s, the linearity of χ enables one to compute $\chi(A,f)$ for any polynomial.

In particular, if the degree of f is $\le m$, then $\chi(A, f) = f$.

6.4 As an illustration we return to the Example 3.5. Here $A_0=(0,0)$; $A_1=(1,0)$ and $b_0=1$; $b_1=3$. We consider the problem of interpolating the function $f(x, y)=x^2y+2x+1$ rather than just the values 1 and 3. Of course $f(A_0)=1$ and $f(A_1)=3$. Since $\deg(2x+1)\le 1$ $\chi(A, 2x+1)=2x+1$ and we need merely determine $\chi(A, x^2y)$.

We consider the three linear functions $\mathbb{R}^2 \to \mathbb{R}$ given by $\lambda_1(x, y)=x+y$; $\lambda_2(x, y)=x-y$ and $\lambda_3(x, y)=y$. Then $x^2y=\frac{1}{6}\{\lambda_1^3-\lambda_2^3-2\lambda_3^3\}$. Using (6.3) we see that $\chi(A, \lambda_1^3)=\lambda_1$; $\chi(A, \lambda_2^3)=\lambda_2$ and $\chi(A, \lambda_3^3)=0$. Thus, by linearity of χ, $\chi(A, f)=2x+\frac{y}{3}+1$.

6.5 In general, the coefficients of $\chi(A,f)$ depend on integrals of derivatives of f over faces in the convex hull of $A_0, \ldots, A_m$. An explicit formula may be given as follows:

Let $$S_k=\left\{(t_0, \ldots, t_k)\in\mathbb{R}^{k+1} \;\middle|\; \sum_{j=0}^{k} t_j=1,\ t_j\ge 0\right\}.$$

Recall that if f is of class $\mathscr{C}^k$ on $\mathbb{R}^n$ then its total k^{th} order derivative $D^k(f)$ at $x \in \mathbb{R}^n$ is a *k-linear* symmetric map from $\mathbb{R}^n$ to $\mathbb{R}$. That is, given k vectors $v_1, \dots, v_k \in \mathbb{R}^n$,

$$D^k f(x)(v_1, \dots, v_k) \in \mathbb{R}$$

and the value is symmetric in $v_1, \dots, v_k$. We let $d\mu_k(t)$ denote Lebesgue measure on S_k. We set

$$\text{(6.6)} \qquad \chi_k(f) = \int_{S_k} D^k(f)\left(\sum_{j=0}^{k} t_j A_j\right)(x - A_0, \dots, x - A_{k-1})\, d\mu_k(t)$$

and

$$\text{(6.7)} \qquad \chi(A, f) = \sum_{k=0}^{m} \chi_k(f).$$

For $n = 1$ formula (6.6) is essentially the standard Hermite-Genocchi formula for the k^{th} order divided differences of a differentiable function and (6.7) reduces to the standard formula for Lagrange interpolation in terms of the divided differences of a function (see [5], section 2.6 or [12], Chapter 1).

7. Sketch of the proof of theorem 5.2.

7.1 We will first sketch the uniqueness statement in Theorem 5.2.

By applying (b) of Theorem 5.2 to operators of order m it is clear that if f is small in the $\mathscr{C}^m$ norm then the coefficients of $\chi(A, f)$ are small. Hence χ is continuous.

Since polynomials are dense in $\mathscr{C}^m$, it suffices, by the linearity and continuity of χ to show that it is uniquely determined on functions of the form λ^s where s is an integer ≥ 1 and λ is a linear function on $\mathbb{R}^n$. Again, by applying (b) of Theorem 5.2 to operators of order 1 which annihilate λ, formula (6.3) and hence the uniqueness of χ follows.

7.2 For the existence part of Theorem 5.2 one may use the formula (6.6). Kergin's proof [7], [8] does not use such a formula.) From (6.6) it is immediate that χ is linear (and continuous). To prove (b) of Theorem 5.2 one establishes the formula

$$\text{(7.3)} \qquad \int_{[A_j]_{j\in J}} q(f) = \int_{[A_j]_{j\in J}} q(\chi(A, f))$$

for all $q \in Q^k(\mathbb{R}^n)$ and $\text{card}(J) = k + 1$. Since both integrands are continuous functions, they must be equal at some points of $[A_j]_{j\in J}$ and hence (b) of Theorem 5.2 follows.

8. Approximation problems.

8.1 The Weierstrass theorem which states that polynomials are dense in the space of continuous functions on a compact set is valid in one or several variables. It is one of the rare results on one variable approximation which has a satisfactory generalization to several variables.

The theory of polynomial interpolation and approximation in one variable is quite rich and extensive. The one variable results often lead, quite naturally, to problems in several variables and it is reasonable to suppose that many one-variable results have interesting generalizations to the case of several variables.

As an illustration, we will review the basic onevariable theory of Tchebyshev approximation and briefly discuss the several variable case.

8.2 Let $f(x)$ be a continuous function on the interval $[-1, 1]$. What is the polynomial of degree $\leq m$ which best approximates f? Alternatively, where should one locate the points $A_0, \dots, A_m$ so that $\mathscr{L}(A, f)$ is the best approximation to f?

Of course, it is necessary to be precise about what one means by best approximation. Several interpretations are possible but we will use the uniform norm

$$\|f\|_\infty = \sup_{x\in[-1,1]} |f(x)|.$$

We quote the following well-known results due to Tchebyshev.

8.3 Theorem. *There is a unique polynomial $P \in \mathscr{P}^m(\mathbb{R})$ such that*

$$\|f - P\|_\infty = \min_{Q\in\mathscr{P}^m(\mathbb{R})} \|f - Q\|_\infty.$$

In other words, there is a unique polynomial of degree $\leq m$ which provides the best approximation.

The next theorem characterizes that best approximating polynomial.

8.4 Theorem. *$P(x)$ is the best approximating polynomial of degree $\leq m$ to f if and only if there are points $-1 \leq x_1 < \dots < x_{m+2} \leq 1$ such that $|(f - P)(x_j)| = \|f - P\|_\infty$ for $j = 1, \dots, m + 2$ and the sign of $f - P$ alternates at successive points in the sequence $x_1, \dots, x_{m+2}$.*

This theorem has, in fact, been developed into an algorithm to obtain the best approximation by computer.

8.5 In the case $f = x^m$ a precise formula for the best approximating $P \in \mathscr{P}^{m-1}(\mathbb{R})$ is known. In fact, the Tchebyshev polynomial $T_m(x) = \frac{1}{2^{m-1}} \cos(m \arc\cos x)$ has leading term x^m and alternately the values $\pm \frac{1}{2^{m-1}}$ at $m+1$ distinct points. Thus $x^m - T_m(x)$ is of degree $\leq m-1$ and has the required properties.

8.6 To what extent do the results 8.3, 8.4 and 8.5 generalize to several variables?

Theorem 8.3 is not valid in two variables – the uniqueness part does not hold. (The existence part is valid because of general Banach space arguments.) The following example (which arose from discussions with L. Bos) illustrates this point.

We consider approximation on the set $K = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1$ and $0 \leq y \leq 1\}$. Let $h(x) = 4T_3(x)$ be the Tchebyshev polynomial of degree 3. Consider $f(x, y) = 4h(x)y(1-y)$. We will show that the polynomial $P \in \mathscr{P}^2(\mathbb{R}^2)$ which provides the best approximation to f (in the uniform norm on K) is not unique.

First, note that $f\left(x, \frac{1}{2}\right) = h(x)$. If $P \in \mathscr{P}^2(\mathbb{R}^2)$ then $\| f - P \|_K \geq \sup_{0 \leq x \leq 1} \left| f\left(x, \frac{1}{2}\right) - P\left(x, \frac{1}{2}\right) \right| \geq 1$. If we consider $P = 0$ we have $|f(x, y)| \leq |h(x)| \, |4y(1-y)| \leq 1$ for $(x, y) \in K$. Hence $P = 0$ provides a best approximation. However $\left(y - \frac{1}{2}\right)^2$ also provides a best approximation since

$$\left| 4h(x)y(1-y) - \left(y - \frac{1}{2}\right)^2 \right| \leq 4y(1-y) + \left(y - \frac{1}{2}\right)^2 \leq 1$$

for $0 \leq y \leq 1$.

Theorem 8.4 clearly depends on the zero set of a polynomial in one variable being finite.

These simple observations do not, of course, preclude a generalization of Tchebyshev theory to several variables. One specific problem in that direction is the following. Find a polynomial of degree $\leq m-1$ which gives the best approximation (in the uniform norm) to the function $x^m + y^m$ on the unit disc $= \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$.

9. Differentiable functions.

9.1 A good theory of higher order divided differences for functions of several variables is lacking. Such a theory would have important applications to numerical solutions of partial differential equations.

In one variable, the Lagrange interpolation formula is connected with higher order differences. In fact, we have the formula

$$\text{(9.2)} \qquad \mathscr{L}(A, f) = \sum_{i=0}^{m} \Delta(A_0, \ldots, A_i, f) \prod_{j=0}^{i-1} (x - A_j)$$

where $A_0, \ldots, A_m$ are distinct. Here $\Delta(A_0, \ldots, A_i, f)$ denotes the i^{th} order divided difference of f. The leading coefficient of $\mathscr{L}(A, f)$ is $\Delta(A_0, \ldots, A_m, f)$ and this may be taken as the definition of the m^{th} order divided difference.

Kergin's interpolation may thus be considered as giving one answer to the question of higher order divided differences. However it is not adequate to answer the problem (9.4).

9.3 We will discuss a general problem in the theory of differentiable functions. The one variable version of this problem was solved by Whitney [15] and the solution is expressed in terms of divided differences. The several variable version of this problem is presumably related to the problem of higher order divided differences.

9.4 Problem. Let X be a closed subset of $\mathbb{R}^n$ and f a function defined on X. Does there exist a function $F \in \mathscr{C}^m(\mathbb{R}^n)$ (for $m = 1, 2, \ldots, +\infty$) such that F/X (the restriction of F to X) is equal to f. In other words, can one decide from the values of f on X if it has an extension to a function in $\mathscr{C}^m(\mathbb{R}^n)$? This is, of course, a somewhat vague question.

For example, if X is a closed submanifold (of class $\mathscr{C}^\infty$) it suffices that f be of class $\mathscr{C}^m$ on the manifold X.

More interesting, is the result of Whitney [15] which solved this problem for arbitrary closed subsets of $\mathbb{R}$. He proved.

9.5 Theorem. (Whitney). *Let X be a closed subset of $\mathbb{R}$ and f a function defined on X. Consider $\Delta(A_0, \ldots, A_m; f)$ for $A_0, \ldots, A_m \in X$. Suppose that $\Delta(A_0, \ldots, A_m, f)$ always has a limit as the points $A_0, \ldots, A_m$ converge to a point $x \in X$. Then f has an extension of class $\mathscr{C}^m$.*

Merrien [11] proved that, in the above circumstances, if f has an extension of class $\mathscr{C}^m$ for all m then it has an extension of class $\mathscr{C}^\infty$.

Whitney's proof of (9.5) uses the fact that the complement of a closed set in $\mathbb{R}$ is a union of intervals.

In $\mathbb{R}^n$ things are, of course, much more complicated. The case when X is a closed rectangle has been solved by Glaeser [6]. The $\mathscr{C}^1$ problem has been studied recenty by S. Birnbaum [3]. There is, of course, related work of Whitney [9].

One particular case would be when X is algebraic. This is related to generalizations of a theorem of Glaeser [2]. See also the result of G. Schwarz [13].

References

[1] A. A. Bennett, W. E. Milne and H. Bateman, *Numerical Integration of Differential Equations*, Dover, New York (1956).

[2] E. Bierstone and T. Bloom (to appear).

[3] S. Bromberg. Thesis, Centro de Investigacion *y* de Estudio Avanzados, Instituto Politecnico Nacional, Mexico.

[4] C. DeBoor, *Polynomial Interpolation*, Proc. Int. Cong. Helsinki (1978).

[5] P. J. Davis, *Interpolation and Approximation*, Blaisdell, New York (1963).

[6] G. Glaeser, *L'interpolation des Fonctions Differentiables de Plusieurs Variables*, Proc. Liverpool Singularities Symposium II. Lecture Notes in Math. 209, Springer 1-33.

[7] P. Kergin, *Interpolation of C^k Functions*, Thesis, University of Toronto (1978).

[8] P. Kergin, *A Natural Interpolation of C^k Functions* (to appear in J. Approximation Theory).

[9] B. Malgrange, *Ideals of Differentiable Functions*, Oxford (1966).

[10] C. Micchelli and P. Milman (to appear).

[11] J. Merrien, *Prolongateurs des Fonctions Differentiables D'une Variable Reélle*, J. Math. Pures Appl. 45 (1966) 291-309.

[12] N. E. Nφrlund. *Vorlesungen Uber Differenzenrechnung*, Berlin, Julius Springer (1924).

[13] G. Schwarz, *Smooth Functions Invariant Under the Action of a Compact Lie Group*, Topology 14 (1975) 63-68.

[14] B. Wendroff, *Theoretical Numerical Analysis*, Academic Press, New York (1966).

[15] H. Whitney, *Differentiable Functions Defined in Closed Sets*, I. Trans. Amer. Math. Soc. 36 (1934) 369-387.

Department of Mathematics
University of Toronto
Toronto, Canada
M5S 1A1