The use of derivatives in the optimisation of the truncation error of polynomial approximations

F. Oliveira-Pinto

Computer Laboratory, University of Cambridge, Corn Exchange Street, Cambridge CB2 3QG

The truncation error of polynomial interpolations can be written as a form expressible as the product of two functions so that one factor depends only on the points we use for the interpolations, while the other depends essentially on the function we aim to approximate. To optimise this error term in a min-max sense, we concentrate on the first part and introduce extra degrees of freedom by use of extra information in the form of values of derivatives at some of the interpolation points. This paper gives a discussion of the 'best' choice of the extra information and the implications of such 'extended' approximations.

(Received April 1973)

1. Introduction

Let us consider a function $Z(x)$ continuous in $[-1, +1]$. We suppose $Z(x)$ to be given numerically on a set of prescribed points $x_{s\to 0,1,\ldots q} \in [-1, +1]$ and, for explanation purposes, we further suppose that no other information concerning Z is available, beyond the belief that it may be approximated by polynomials of moderate degree. (This latter assumption is somewhat relaxed in the last sections of this paper.)

We aim to approximate $Z(x)$ by a polynomial $P_N(b, x)$ which coincides with $Z(x)$ at the points $x_{s\to 0,1,\ldots q}$ and whose degree $N(N \geq q)$ and coefficients $b_{k\to 0,1,\ldots N}$ are chosen so as to optimise the approximation in a certain sense.

Due to the absence of a definition of $Z(x)$ in $[-1, +1]$ we cannot minimise the modulus of

$$\varepsilon_N(x) = Z(x) - P_N(b, x) \tag{1.1}$$

for all x. We can, however, split $\varepsilon_N(x)$ into two factors, one of which contains the known (x_s dependent) component, and the other the unknown (Z dependent) component; of the two factors the former may indeed be minimised.

To explain the splitting of $\varepsilon_N(x)$ we first consider the remainder formula for a polynomial interpolation of degree q which can be written as (Davis, 1963, p. 64),

$$\varepsilon_q(x) = \rho_q(x) \,.\, \mathscr{R}_{q+1}(x) \tag{1.2}$$

with

$$\mathscr{R}_{q+1}(x) = \prod_{s=0}^{q} (x - x_s) \tag{1.3}$$

where $\rho_q(x)$ contains the information which is function dependent.

In our present problem we have $N \geq q$ and so choose to write

$$\varepsilon_N(x) = \bar{\rho}_N(q, x)\, \bar{u}_{N+1}(m, x) \tag{1.4}$$

where

$$\bar{u}_{N+1}(m, x) = \prod_{s=0}^{q} (x - x_s)^{m_s} \,, \tag{1.5}$$

with $m_0 + m_1 + \ldots + m_q = N + 1$.

For convenience, we insert a normalising constant and choose to examine the polynomial given by

$$u_{N+1}(m, x) = E \prod_{s=0}^{q} (x - x_s)^{m_s} \tag{1.6}$$

where m_s are positive integers.

We start by studying some of the limitations of this approach in Section 2. Afterwards we obtain for sets of equally-spaced points $\bar{x}_{s\to 0,1,\ldots q}$ explicit numerical values for $m_{s\to 0,1,\ldots q}$.

We leave for a second paper (Oliveira-Pinto, 1974) the presentation of a well-conditioned algorithm (based on a modified Gram-Schmidt process) which has enabled us to experiment extensively with these approximating forms.

2. The optimisation of the truncation error

Our aim is to reduce all the maxima and minima in x of the polynomial (1.6) to the same height which is one of the known requirements of min-max approximations (Rice, 1964, p. 56). However, the reduction of all the extrema of (1.6) to the same height is not usually possible for fixed x_s, and so we are constrained to define a 'near-optimisation' scheme for (1.6) where we make the maximum linear distance of $q + 2$ of its extrema from their average value (disregarding signs) a minimum.

These $q+2$ extrema are the maxima and minima of $u_{N+1}(m, x)$ in between zeros and for $x \in [-1, \min(x_s)]$, $[\max(x_s), +1]$. We denote by $x'_{s\to 0,1,\ldots q+1}$ the corresponding abscissae in x.

Since the height of maxima and minima of $u_{N+1}(m, x)$ on its zeros $x_{s\to 0,1,\ldots q}$ cannot be numerically controlled, we must not allow these type of extrema to occur. This is done by restricting the set of possible solutions for $m_{s\to 0,1,\ldots q}$ to the set of *odd* integers 1, 3, ... with the possible exception of m_0 and m_q in the case of $x_0, x_q = \mp 1$.

To simplify the formulation of this approximation process we start by defining a reference function $U_{q+1}(x - x_s)$ in the following terms:

(*a*) $U_{q+1}(x - x_s)$ is continuous in $[-1, +1]$ with its zeros and extrema in x on the same set of abscissae as those of $u_{N+1}(m, x)$;

(*b*) The $q + 2$ extrema of $U_{q+1}(x - x_s)$ have the same amplitude and their signs agree with the corresponding signs of $u_{N+1}(m, x'_s)$;

(*c*) Moreover, the difference $u_{N+1}(m, x) - U_{q+1}(x - x_s)$ cannot have extrema in x other than those corresponding to the extrema of $u_{N+1}(m, x)$.

Since it may be difficult to visualise such a function we give here one possible way of constructing $U_{q+1}(x - x_s)$.

Let us consider a non-negative step function $w(x)$ in $[-1, +1]$ which for $[-1, x_0)$, $[x_0, x_1)$, ... $[x_q, +1]$ with $x_0 < x_1 < \ldots < x_q$ has the values

$$w(x) \equiv \frac{1}{|h'_s|} \text{ where } h'_s = u_{N+1}(m, x'_s),\ s \to 0, 1, \ldots q + 1 \,. \tag{2.1}$$

The product $w(x) \,.\, u_{N+1}(m, x)$, satisfies (*a*), (*b*), (*c*). It can therefore be taken as a reference function.

We are now in a position to state that we require

$$\max_x \left| E \prod_{s=0}^{q} (x - x_s)^{m_s} - U_{q+1}(x - x_s) \right| , x \in [-1, +1] \tag{2.2}$$

to be a minimum in the odd integer powers $m_{s\to 0,1,\ldots q} > 0$. We shall denote by $m^*_{s\to 0,1,\ldots q}$ the coordinates of such a minimum. The factor E is a normalisation factor to $u_{N+1}(m, x)$

and it is such that the *mean* of the absolute value of the $q + 2$ extrema of $u_{N+1}(m, x)$ is equal to

$$\max_x \left| U_{q+1}(x - x_s) \right|$$

The reason why we cannot use an ordinary 'absolute' norm

$$\max_x \left| \prod_{s=0}^{q} (x - x_s)^{m_s} \right| , \; x \in [-1, +1] \tag{2.3}$$

which corresponds to $U_{q+1}(x - x_s) \equiv 0$ and $E = 1$ in (2.2) needs an explanation.

Let us consider the norm (2.3) when we take $x_{s \to 0, 1, \ldots q}$ as the zeros of the Chebyshev polynomial $T_{q+1}(x) = \cos((q + 1) \arccos x)$, that we represent by $x^*_{s \to 0, 1, \ldots q}$ and A the coefficient of its highest power of x. Making $m_s \equiv 1$ we have

$$T_{q+1}(x) = A \prod_{s=0}^{q} (x - x_s^*) \tag{2.4}$$

which fulfills the requirement of all its $q + 2$ extrema having the same amplitude. We can prove however that $m_s \equiv 1$ are not the coordinates of a minimum. In fact, defining $u_{N+1}(m, x)$ by

$$u_{N+1}(m, x) = (x - x_0^*)^2 (x - x_q^*)^2 T_{q+1}(x) \tag{2.5}$$

we have

$$[(x)^2 - (x_0^*)^2]^2 |T_{q+1}(x)| < |T_{q+1}(x)| , \; x \in [-1, +1] \tag{2.6}$$

because $x_q^* = -x_0^*$ and so our statement is proved.

We are then going to solve (2.2) subject to the condition that

$$1 \leq m_{s \to 0, 1, \ldots q} \leq M + 1 \tag{2.7}$$

for M fixed (preferably an even integer) and independent of the relative position of $x_{s \to 0, 1, \ldots q}$ in $[-1, +1]$. M defines the highest order derivative of the data function $Z(x)$ that may be required at $x = x_s$.

Unfortunately, for a given set of points $x_{s \to 0, 1, \ldots q}$ and a fixed M, no simple method is at present known for the computation of $m_{s \to 0, 1, \ldots q}$ iteratively. But for reasonably small values of q and M, say $q \leq 24$ and $M = 2$, the total number of possible polynomials $u_{N+1}(m, x)$ is from(2.7)—2^{q+1}—which is a number small enough to make feasible, with modern automatic computers, the direct evaluation of each polynomial $u_{N+1}(m, x)$, so as to choose from amongst them the required $u_{N+1}(m^*, x)$.

This direct verification method has the advantage that we may obtain without extra cost a '2nd best' solution which may be easier to use and, in practice, give comparable polynomial approximations. This was the method used when computing the values of **Table 1** for equidistant values of $x_{s \to 0, 1, \ldots q}$ described in the next section.

3. The distribution of powers for equidistant data

We shall now study the error-function $u_{N+1}(m, x)$ for a special set of sampling points: the equidistant set.

The reason for investigation of this case is the fact that it occurs often in practice. For example, sampling of observational functions by clock mechanisms leads to equidistant sampling. Further, with numerical problems of the step-by-step type, we often finish with a set of numbers defined at equal step-size intervals. But it is well known how disastrous Langrangian polynomial approximations may be for equally-spaced $x_{s \to 0, 1, \ldots q}$ and large q (e.g. **Fig. 1**). In Lanczos (1961) p. 13 we read 'We thus come to the conclusion that interpolation in the large by means of high order polynomials is not obtainable by Langrangian interpolation of equidistant data. If we fit our data exactly by a Langrangian polynomial of high order we shall generally encounter exceedingly large error oscillations around the end (sic) of the range'. To reduce these large oscillations near the boundaries of the normalised interval for x we have computed, for equidistant $x_{s \to 0, 1, \ldots q}$, the set of integer powers $m^*_{s \to 0, 1, \ldots q} \leq M + 1$ which makes the function $u_{N+1}(m, x)$ to have its extremes of almost the same height.

The result is condensed in Table 1 where only half of the powers $m^*_{s \to 0, 1, \ldots q}$ due to their symmetry are represented. In

Table 1 The 'odd' powers of the Error-functions $E_{q+1} \prod_s (x - \bar{x}_s)^{m_s^*}$, $\bar{x}_s = \frac{2s - q}{q}$, $m_s^* \leq 3$

Δ	$2^{-q} E_{q+1}$	max. deviat.	$m_0^* m_1^* m_2^* \ldots$	$q + 1$	$2^{-q} E_{q+1}$	max. deviat.	$m'_0 m'_1 m'_2 \ldots$
	1·17	+0·04	2 1	4	0·74	−0·34	1 1
0·44	1·45	−0·04	2 1 1	5	0·79	+0·44	1 1 1
0·78	1·70	+0·22	2 1 1	6	0·80	+0·78	1 1 1
1·19	3·15	−0·22	3 1 1 1	7	1·91	+0·39	2 1 1 1
1·64	3·74	+0·18	3 1 1 1	8	2·08	+0·65	2 1 1 1
2·13	4·33	+0·25	3 1 1 1 1	9	2·18	+0·97	2 1 1 1 1
2·64	4·88	+0·36	3 1 1 1 1	10	2·21	+1·3	2 1 1 1 1
3·16	5·35	+0·49	3 1 1 1 1 1	11	2·16	+1·8	2 1 1 1 1 1
3·68	47·29	−0·64	3 3 1 1 1 1	12	5·71	+0·65	3 1 1 1 1 1
4·20	53·56	−0·62	3 3 1 1 1 1 1	13	5·91	+0·93	3 1 1 1 1 1 1
4·72	60·59	−0·59	3 3 1 1 1 1 1	14	5·96	+1·3	3 1 1 1 1 1 1
5·24	68·17	−0·54	3 3 1 1 1 1 1 1	15	5·84	+1·6	3 1 1 1 1 1 1 1
5·75	76·02	−0·49	3 3 1 1 1 1 1 1	16	5·57	+2·0	3 1 1 1 1 1 1 1
6·27	83·75	+0·46	3 3 1 1 1 1 1 1 1	17	5·19	+2·4	3 1 1 1 1 1 1 1 1
6·78	90·92	+0·56	3 3 1 1 1 1 1 1 1	18	4·74	+2·8	3 1 1 1 1 1 1 1 1
7·29	97·04	+0·71	3 3 1 1 1 1 1 1 1 1	19	4·24	+3·2	3 1 1 1 1 1 1 1 1 1
7·80	101·68	+0·87	3 3 1 1 1 1 1 1 1 1	20	3·72	+3·6	3 1 1 1 1 1 1 1 1 1
8·31	975·01	−0·95	3 3 3 1 1 1 1 1 1 1 1	21	104·47	+1·0	3 3 1 1 1 1 1 1 1 1 1
8·81	1075·58	−0·94	3 3 3 1 1 1 1 1 1 1 1	22	105·21	+1·2	3 3 1 1 1 1 1 1 1 1 1
9·32	1176·06	−0·93	3 3 3 1 1 1 1 1 1 1 1 1	23	103·85	+1·4	3 3 1 1 1 1 1 1 1 1 1 1
9·82	1271·42	+1·2	3 3 3 1 1 1 1 1 1 1 1 1	24	100·52	+1·6	3 3 1 1 1 1 1 1 1 1 1 1
10·3	1355·98	+1·6	3 3 3 1 1 1 1 1 1 1 1 1 1	25	95·51	+2·0	3 3 1 1 1 1 1 1 1 1 1 1 1

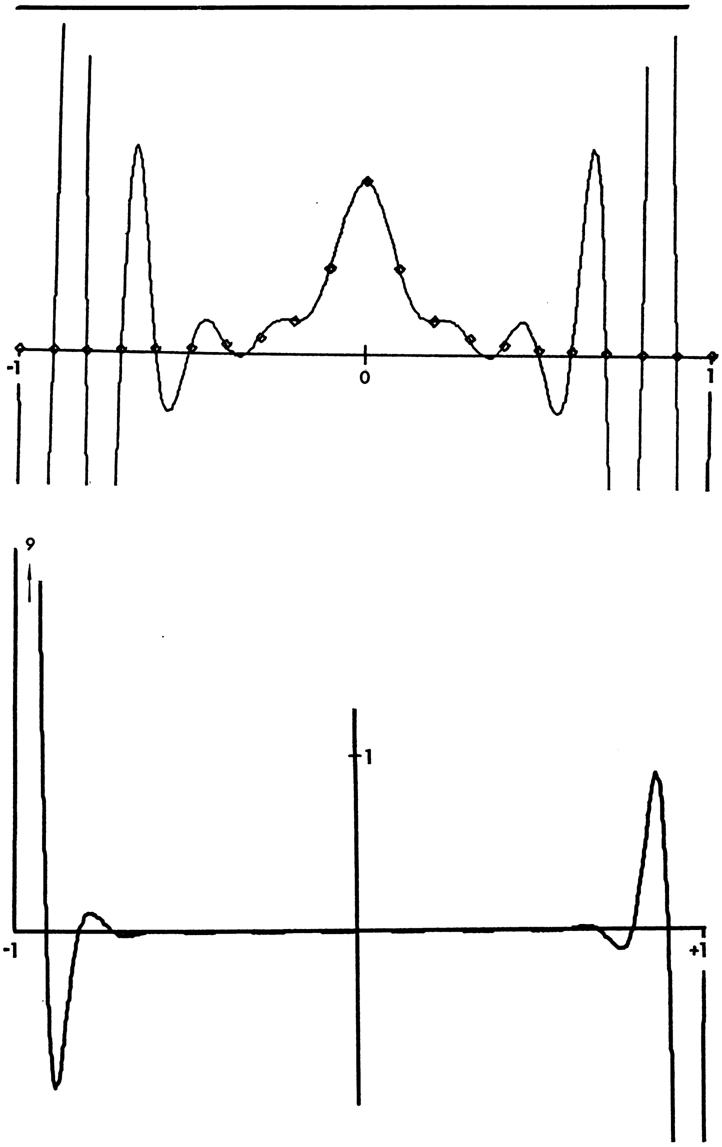

Fig. 1 Ordinary polynomial interpolation and corresponding $u_{q+1}(m, x)$

the second half of that Table we give a 'second best' set that we call here $m'_{s\to 0,1,\dots q}$ subject to the condition $m'_s \leq m^*_s$ since we believe it to be pointless to have, as a second choice, a $u_{N+1}(m', x)$ which is more expensive to construct. Finally, in the column of the maximum deviations, we have respectively

$$\max_x \left| u_{N+1}(m^*, x) \right| - 1 \text{ or } \max_x \left| u_{N+1}(m', x) \right| - 1$$

and thus an upper bound for these errors is always available. For easy reference the first column of the table contains the maximum deviations to one of the $m_s \equiv 1$ Lagrange interpolator case. The lower plot of **Fig. 2** shows the plot of $u_{N+1}(m^*, x)$ of Table 1 for $q = 20$.

We now comment on Table 1.

1. To balance the 'exceedingly large error oscillations around the end(s) of the range' the tabulated polynomials show, as expected, larger powers in the zeros around the ends of $[-1, +1]$ than in the middle. Let us for example take the polynomial

$$u_{N+1}(m, x) = E_{q+1}(x - \bar{x}_0)^3 (x - \bar{x}_1) \dots (x - \bar{x}_{q-1})(x - \bar{x}_q)^3 \quad . \quad (3.1)$$

Because $\bar{x}_q = -\bar{x}_0$ we have

$$u_{N+1}(m, x) = [(x)^2 - (\bar{x}_0)^2]^2 R_{q+1}(x) \quad (3.2)$$

$$R_{q+1}(x) = E_{q+1}(x - \bar{x}_0)(x - \bar{x}_1) \dots (x - \bar{x}_{q-1})(x - \bar{x}_q) \quad . \quad (3.3)$$

The first factor in (3.2) may be considered as a 'damping' function to the corresponding Lagrange polynomial $R_{q+1}(x)$ and the polynomial $u_{N+1}(m, x)$ displays, as desired, much smaller oscillations around the ends of the interval $[-1, +1]$ than the corresponding Langrangian one $R_{q+1}(x)$.

2. In the first half of the Table the same pattern of powers extends over consecutive values of q (e.g. for $q \to 11, 12, \dots 19$) and for such sub-sets the absolute value of the maximum deviation decreases first with increasing q, reaches a minimum and often increases again until a different pattern arises. The same is not, however, true for the second half of the Table where the maximum deviation within the same pattern always increases.
3. Looking again at the columns of the maximum deviations, we note that the general tendency of its values is to increase with increasing q. This shows how poor polynomial approximations can be for large sets of equidistant data points, even though non-Langrangian ones. This is due to the lack of flexibility in the choice of powers $m^*_{s\to 0,1,\dots q}$ since they can only have integer values. This also limits for a given M, the maximum number of equidistant sampling points that one can safely use.

4. Numerical experiments and conclusions

Let us consider the classical example introduced by Runge around 1900 (Lanczos, 1961, p. 12). The data of this example are defined by the values $Z_{s\to 0,1,\dots 20}$ obtained by direct sampling at the equidistant points $X_s \to \pm 10, \pm 9, \pm 8, \dots 0$ of

$$Z\,(X) = \frac{1}{1 + X^2}$$

or, which is equivalent, from

$$Z(x) = \frac{1}{1 + 100x^2}$$

now at the points $x_s \to \pm 1{\cdot}0, \pm 0{\cdot}9, \dots 0$. These 21 sampled values $Z_{s\to 0,1,\dots 20}$ define the 'basic' table of values which we shall now approximate using different polynomial techniques.

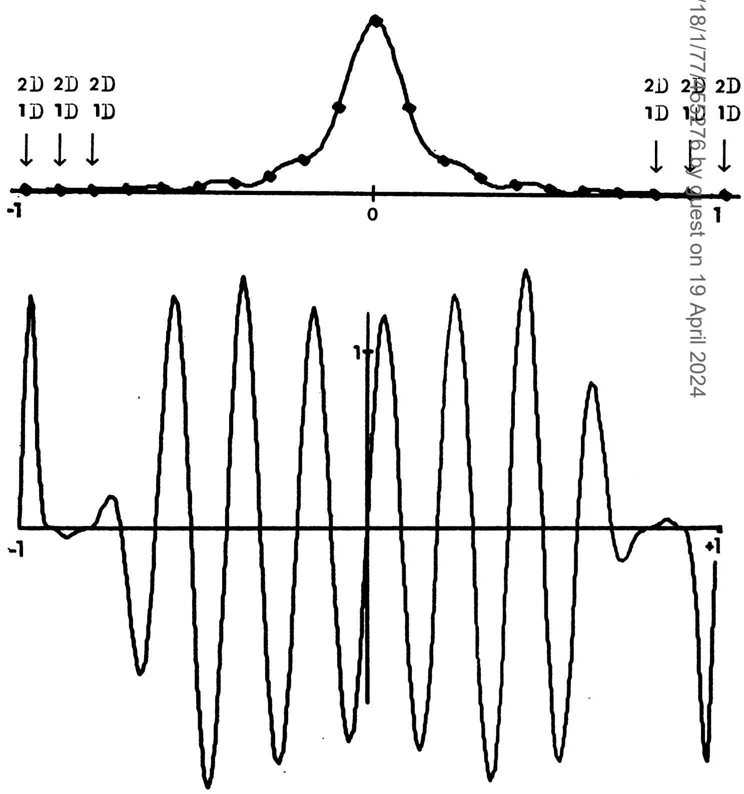

Fig. 2 Extended polynomial interpolation with the corresponding $u_{N+1}(m, x)$

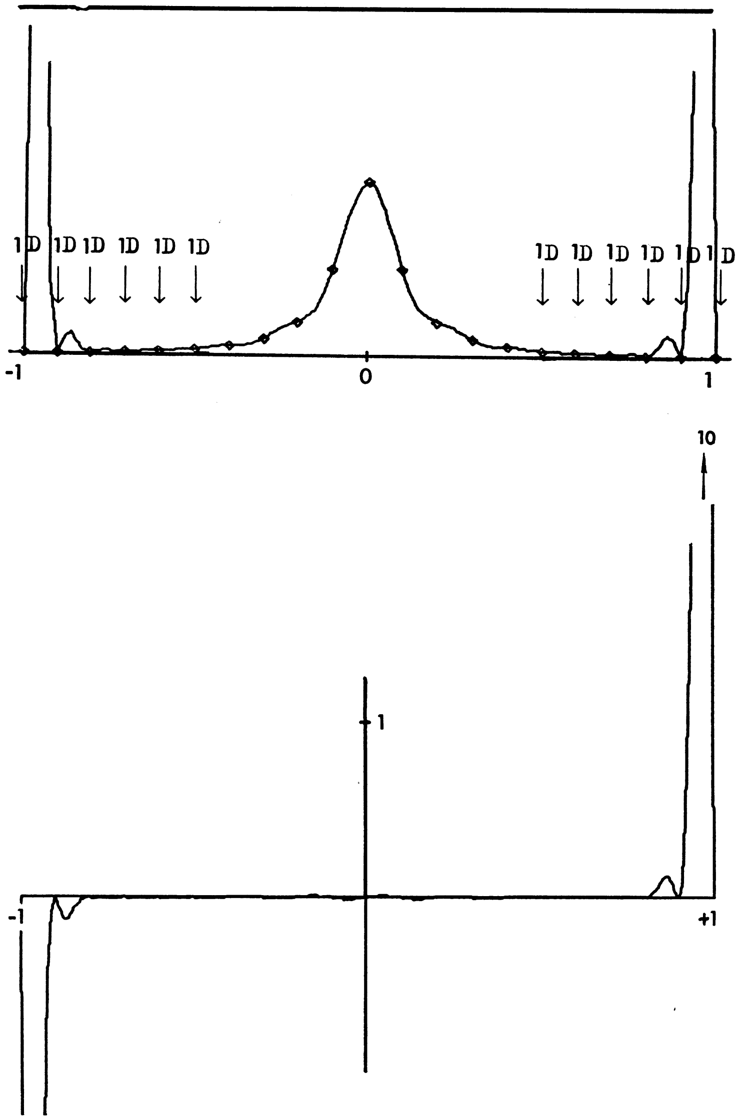

Fig. 3 **Extended polynomial interpolation with non-optimising derivative information**

Experiment 1:
For the Langrangian type of approximation we have obtained what is plotted in the upper part of Fig. 1. The strong oscillations in between data values near the boundaries of the normalised interval for x, $[-1, +1]$, are disappointing. The function $u_{N+1}(m, x)$ corresponding to this polynomial approximation is shown in the lower plot of the same figure, and the correlation between the exceedingly large oscillations in both plots (apart from a scaling factor) is overwhelming.

Experiment 2:
Including now derivative information such that the error $u_{N+1}(m, x)$, $q = 20$ is 'best' according to Table 1 we obtained the upper plot of Fig. 2 where we used in the first and last 3 tabular points, 1st and 2nd derivatives of $Z(x)$. The oscillations in between data values at the end of the normalised interval $[-1, +1]$ are much smaller (by a factor 2×10^4 as can be shown for $x = \pm 0{\cdot}95$) than the corresponding ones in the Langrangian experiment. Due to the presence of derivative values this type of polynomial approximation converges to $Z(x)$ everywhere in $[-1, +1]$. The lower plot illustrates the function $u_{33}(m^*, x)$.

Experiment 3:
To emphasise the importance in the proper selection of the derivative information we prepared an example where we have taken the same number of derivative values as used in the previous experiment but in different tabular points. Thus, we considered 1st derivative values on the first and last six tabular points and the result is given in the upper plot of **Fig. 3.** This approximation has improved around $x = \pm 0{\cdot}3$ when compared with the two previous experiments but near ± 1 is still very poor. Note again the strong similarity between the shape of the approximation obtained and the corresponding $u_{33}(m, x)$ in the lower part of the same figure.

Acknowledgements
I would like to mention Dr. D. Barton and Dr. J. Fitch for their valuable discussions during the preparation of the manuscript.

References

Davis, P. J. (1963). *Interpolation and approximation*, Blaisdell, New York.
Lanczos, C. (1961). *Linear differential operators*, van Nostrand, London.
Oliveira-Pinto, F. (1974). *Extended orthonormal polynomials and their use in numerical approximation* (to appear).
Rice, J. R. (1964). *The approximation of functions*, Vol. 1, Addison-Wesley, Reading, Mass.