OPTIMUM QUADRATURE FORMULAE AND NATURAL SPLINES

BY

D. Kershaw

thesis submitted for the degree of Doctor of Philosophy of the
iversity of Edinburgh, 1972.

To my wife Carol and my sons

Samuel and Edward

Contents

Introduction and Summary

Introduction

This thesis is concerned mainly with the investigation of two types of quadrature formulae. These are Sard's best second order formula [12] and the third order optimum quadrature which was introduced by Schoenberg [15]. The formulae are closely related to certain types of natural spline, consequently we shall examine the convergence properties of these in order to obtain results on the convergence of the quadrature formulae.

The quadrature formulae arise out of the following considerations (c.f. Handscomb [5]).

It is well known that the error in Simpson's rule,

$$\int_0^1 x(t)dt = \frac{1}{6}\,[x(0) + 4x(\tfrac{1}{2}) + x(1)],$$

can be written $-x^{(4)}(t')/2880$ when $x \in C^4[0,1]$. The number t', which satisfies $0 < t' < 1$, is usually indeterminate and the error has to be estimated from $||x^{(4)}||/2880$ (the uniform norm on $[0,1]$). However if x does not possess a bounded fourth derivative on the interval then this estimate will be useless even when the error is in fact bounded. An alternative approach to the problem of finding a realistic estimate is to construct an expression for the error when x has only a bounded derivative of lower order. This can be done with the aid of Peano's method, and since we shall make frequent use of it in this thesis we present it here in a general framework (see [3])

Peano's method for remainders

Let L be a linear functional on $C^n[a,b]$, and let L_n be an approximation to L in the sense that $(L-L_n)s^r = 0$, $r=0,1,\ldots,n-1$. If $x \in C^n[a,b]$ we have the expansion

$$x(s) = \sum_{r=0}^{n-1} \frac{(s-a)^r}{r!} x^{(r)}(a) + \int_a^b K_{n-1}^+(s-t)x^{(n)}(t)dt$$

where $K_{n-1}^+(s-t) = \frac{(s-t)^{n-1}}{(n-1)!}$ for $s > t$ and zero otherwise.

Since $L-L_n$ is linear

$$(L-L_n)x(s) = (L-L_n)\int_a^b K_{n-1}^+(s-t)x^{(n)}(t)dt$$
$$= \int_a^b x^{(n)}(t)k(t)dt$$

where $k(t) = (L-L_n)K_{n-1}^+(s-t)$. The function k is known as the Peano kernel of $L-L_n$, and it follows that if $x \in C^n[a,b]$ we may write

$$Lx = L_n x + \int_a^b k(t)x^{(n)}(t)dt. \qquad (0.1)$$

For example with Simpson's rule and $x \in C^2[0,1]$ the remainder will be given by

$$\int_0^1 x^{(2)}(t)k(t)dt \text{ where } k(t) = \frac{(1-t)^2}{2} - [4K_1^+(\tfrac{1}{2}-t)+(1-t)]/6,\ 0 \leq t \leq 1, \qquad (0.2)$$

and the error can be estimated from this.

Sard's best quadrature and optimum quadrature

It will be noticed that although Simpson's rule is exact for cubics we used only the property that it is exact for linear functions in the determination of the remainder in (0.2). Thus in a sense we had a spare parameter which in this case was chosen to produce Simpson's rule. This suggests the possibility of choosing this parameter to minimize the error in some way. These rather vague statements will now be made more precise; let $t_o, t_1, \ldots, t_N$ be given numbers which satisfy $0 = t_o < t_1 < \ldots < t_N = 1$, and let

$$R(x) = \int_0^1 w(t)x(t)dt - \sum_{i=o}^{N} H_i x(t_i) \tag{0.3}$$

where w is some integrable weight function. We shall call the equation (0.3) a quadrature formula with remainder R, and suppose that R vanishes identically if $x(s) = 1, s, \ldots, s^{n-1}$. A more convenient expression can be found for R by means of Peano's method, namely that

$$\text{if } x \in C^n[0,1] \text{ then } R(x) = \int_0^1 x^{(n)}(t)k(t)dt \tag{0.4}$$

$$\text{where } k(t) = \int_0^1 w(s)K^+_{n-1}(s-t)ds - \sum_{i=o}^{N} H_i K^+_{n-1}(t_i-t). \tag{0.5}$$

Clearly if $n = N+1$ the quadrature weights can be calculated by the use of Lagrange's interpolation formula, however for $n \leq N$ they are not uniquely defined. The suggestion made by Sard in [12] was that they should be calculated so as to minimize $\int_0^1 [k(t)]^2 dt$, and since the formula is to be exact for all polynomials of degree n-1 the minimization is to be performed subject to the following conditions being true.

$$\int_0^1 w(t)t^r dt = \sum_{i=o}^{N} H_i t_i^r, \quad r=0,1,\ldots,n-1. \tag{0.6}$$

Sard calls the resulting quadrature formula a <u>best quadrature</u>

formula of order n. It is not difficult to show that there is a unique formula of order n when the quadrature points are given. There are obviously many other criteria which could be chosen, however Sard's leads to an attractive theory and it is the one which will be considered here.

One need not be satisfied with this minimum and the further problem can be posed of finding the quadrature points as well as weights so as to decrease $\int_0^1 [k(t)]^2 dt$ still further. This type of formula is called an optimum or optimal quadrature formula of order n and seems to have been suggested first by Schoenberg [15]. The analogy between this and the usual Gaussian quadrature is clear.

The main aim of this thesis is to present some results on the simplest non trivial examples of the formulae, the best of second order and the third order optimum. We shall investigate their convergence and show for the optimum formula that it has properties similar to that of a Gaussian quadrature formula with positive weight function.

Natural Splines

Invaluable tools in these investigations have been certain natural splines which are intimately connected with the quadrature formulae. Splines were introduced in 1946 by Schoenberg in [13], however it was not until 1964 that Schoenberg showed in [14] how they are related to Sard's formulae. We present here a derivation of his result in the case of optimum formulae, and since for this formula we do not know beforehand where the quadrature points will lie we shall set $a = \min [0, t_0]$, $b = \max [1, t_N]$ and let $x \in C^n[a,b]$. The Peano kernel is as before but we now require $t_o, t_1, \ldots, t_N$, $H_o, H_1, \ldots, H_N$ to satisfy (0.6) and to minimize $\int_a^b [k(t)]^2 dt$. This problem can be reformulated by Lagrange's method of undetermined multipliers as that of finding a minimum of

$$\tfrac{1}{2}\int_a^b [k(t)]^2 dt + \sum_{r=o}^{n-1} \lambda_r \left\{ \int_0^1 w(t) t^r dt - \sum_{i=o}^{N} H_i t_i^r \right\},$$

this leads easily to the following sets of equations.

$$\begin{aligned} &\int_a^b k(t) K^+_{n-1}(t_i - t) dt + \sum_{r=o}^{n-1} \lambda_r t_i^r = 0, \quad i=0,1,\ldots,N, \\ &\int_a^b k(t) K^+_{n-2}(t_i - t) dt + \sum_{r=1}^{n-1} r\lambda_r t_i^{r-1} = 0, \quad i=0,1,\ldots,N. \end{aligned} \tag{0.7}$$

(We shall assume that none of the weights vanish and so we have removed the non zero multipliers in the second set of equations.)

Let $c_o, c_1, \ldots, c_N$, $d_o, d_1, \ldots, d_N$ be scalars which satisfy

$$\sum_{i=o}^{N} c_i t_i^r + r \sum_{i=o}^{N} d_i t_i^{r-1} = 0, \quad r=0,1,\ldots,n-1, \tag{0.8}$$

then we see from (0.7) that if $y \in C^n(-\infty,\infty)$, where

$$y^{(n)}(t) = \sum_{i=o}^{N} \{ c_i K^+_{n-1}(t_i - t) + d_i K^+_{n-2}(t_i - t) \}, \tag{0.9}$$

it will follow that $\int_a^b k(t) y^{(n)}(t) dt = 0$. Consequently this function is integrated exactly by the optimum formula (if it exists). Clearly $y^{(n)}$ is identically zero for $t \geq t_N$, and because of (0.8) it is easily seen that $y^{(n)}$ is identically zero for $t \leq t_o$. We shall call such a function a <u>natural quintic Hermitian spline of degree 2n-1</u> with the knots $t_o, t_1, \ldots, t_N$. The integration of (0.9) leads to the representation of y in $(-\infty,\infty)$ as

$$p_{n-1}(t) + (-1)^n \sum_{i=o}^{N} \{c_i K^+_{2n-1}(t_i - t) + d_i K^+_{2n-2}(t_i - t)\}, \qquad (0.10)$$

where p_{n-1} is an arbitrary polynomial of degree n-1.

The result which corresponds to this for best quadrature (see Schoenberg [14]) is that the best quadrature formula of order n integrates exactly any natural spline of degree 2n-1 with the knots $t_o, t_1, \ldots, t_N$. This spline has the representation

$$p_{n-1}(t) + (-1)^n \sum_{i=o}^{N} c_i K^+_{2n-1}(t_i - t) \qquad (0.11)$$

where $$\sum_{i=o}^{N} c_i t_i^r = 0, \; r=0,1,\ldots,n-1. \qquad (0.12)$$

It is easily seen that these definitions can be replaced by the following.

Definition 0.1

A natural spline of degree 2n-1 with the knots $t_o, t_1, \ldots, t_N$ is in $C^{2n-2}(-\infty,\infty)$, is a polynomial of degree at most 2n-1 in each interval $[t_i, t_{i+1}]$, i=0,1,...,N-1 and for $t < t_o$ and $t > t_N$ is a polynomial of degree at most n-1.

Definition 0.2

A natural Hermitian spline of degree 2n-1 with the knots $t_o, t_1, \ldots, t_N$ is in $C^{2n-3}(-\infty,\infty)$, is a polynomial of degree at most 2n-1 in each interval $[t_i, t_{i+1}]$, i=0,1,...,N-1 and for $t < t_o$, $t > t_N$ is a polynomial of degree at most n-1.

These splines can be used for interpolation purposes, the first in a manner analogous to Lagrangian interpolation and the second analogous to Hermitian interpolation. The convergence of these interpolation processes will clearly induce results on the convergence of the respective quadrature formulae. Consequently we shall devote some space to investigations of this convergence.

Finally, we state the next results. We do not make use of them in the sequel, however they are interesting and not difficult to prove.

Theorem A

If y is a natural spline of degree 2n-1 with the knots $t_o, t_1, \ldots, t_N$ then

$$\int_{t_o}^{t_N} [y^{(n)}(t)]^2 dt \leq \int_{t_o}^{t_N} [z^{(n)}(t)]^2 dt$$

for any $z \in C^n[t_o, t_N]$ which satisfies $y(t_i) = z(t_i)$, $i=0,1,\ldots,N$. Furthermore equality holds if and only if $y = z$.

Theorem B

If y is a natural Hermitian spline of degree 2n-1 with the knots $t_o, t_1, \ldots, t_N$

$$\int_{t_o}^{t_N} [y^{(n)}(t)]^2 dt \leq \int_{t_o}^{t_N} [z^{(n)}(t)]^2 dt$$

for any $z \in C^n[t_o, t_N]$ which satisfies $y(t_i) = z(t_i)$, $y^{(1)}(t_i) = z^{(1)}(t_i)$, $i=0,1,\ldots,N$. Furthermore equality holds if and only if $y = z$.

Summary

Chapter two contains results on the convergence of natural cubic spline interpolation when the function with which the spline agrees at the knots is continuous. An analogue of Weierstrass' theorem on uniform approximation by polynomials is proved for approximation by natural cubic splines. However this does not imply the convergence of the interpolation process and we present two theorems which show that with certain restrictions on the spacing of the knots convergence is assured. Since the norm of the natural spline operator is of interest we find upper bounds on it with these same restrictions on the spacing. On the other hand a distribution of knots is found for which the norm is unbounded. The chapter ends with two theorems on the convergence of the natural spline interpolation process when the function which is being interpolated is in C^4. The proof of one of these theorems appeared in S.I.A.M. Journal of Numerical Analysis, and a copy of the paper will be found at the end of the thesis.

The work which comes closest to that of this chapter is that of Cheney and Schurer [2], Nord [11] and Meir and Sharma [10]. These authors discuss similar problems for the periodic cubic spline.

Chapter three is devoted to Sard's second order best quadrature, this is the simplest non trivial formula of its type. Sard has tabulated weights for this when the intervals are equal and for constant weight function. We present here explicit formulae for both the weights and the L_2 norm of the Peano kernel. The convergence theorems of chapter 2 are used to prove convergence

theorems for the quadrature formula. The contents of this chapter have been accepted by I. J. Schoenberg for publication in the Journal of Approximation Theory.

The results of chapter four are preliminary to the investigations in the later chapters and are probably not of great interest in themselves. The natural quintic spline and Sard's third order best quadrature formula are defined here, and a theorem on the convergence of interpolation by this spline is proved. This is required for the subsequent estimation of the Peano kernel for the best quadrature formula. The chapter closes with a theorem which is perhaps of more intrinsic interest than the others. This states roughly that the addition of an extra point in a quadrature formula gives rise to a best quadrature formula with a smaller Peano kernel (measured in the L_2-norm). The result is not surprising but it does not seem to have been stated explicitly before; however a special case of it, when the weight function is constant, is implicit in Karlin [6]. An examination of the natural quintic Hermitian spline occupies chapter five and a proof of its convergence to the function it interpolates is given. The bulk of the chapter consists of an investigation of the qualitative properties of the fundamental Hermitian splines. These are of some interest in themselves since it will be seen that they have the same qualitative behaviour as the fundamental polynomials in Hermite's interpolation formula. Their main use however is in the final chapter where they are used to deduce properties of third order optimum quadrature formulae. Karlin [6] (see also Schoenberg [15]) has shown that such optimum formulae exist for the optimum quadrature of $\int_0^1 w(t)x(t)dt$ when $w=1$. The proof was given in outline only and it is not clear if it is applicable when w is not constant. We present here an existence proof in the more general situation. It is shown that if the weight function is positive then the quadrature weights are also positive. The chapter ends with a result on the distribution of quadrature points when $w=1$ and two theorems on the convergence of the formula.

We conclude the thesis with an appendix in which two types of interpolatory cubic spline are investigated. These are shown to have more favourable convergence properties than the interpolatory natural cubic spline even though they use only the same information for their construction.

Acknowledgements

I wish to express my warmest thanks to my supervisor Professor Michaelson of the University of Edinburgh for his constant encouragement and invaluable criticisms. My thanks go also to Mrs. S. A. Brennan and Mrs. J. K. Howard who expertly typed this thesis, and also to the department of mathematics of the University of Lancaster for the use of their facilities.

CHAPTER 1

Notation and Preliminary Results

The numbers $t_o, t_1, \ldots, t_N$ which satisfy $a \leq t_o < t_1 < \ldots < t_N \leq b$ will be called knots or quadrature points.

$$h_i = t_{i+1} - t_i, \quad i=0,1,\ldots,N-1,$$

$$h = \max_i h_i, \; k = \min_i h_i,$$

$$\alpha_i = h_i/(h_{i-1}+h_i), \quad i=1,2,\ldots,N-1.$$

y will invariably denote a natural interpolating spline, that is to say it will be a natural spline (cubic, quintic or Hermitian quintic) which takes preassigned values (in the Hermitian spline the first derivative also takes preassigned values) at the knots $t_o, t_1, \ldots, t_N$.

s will denote a natural spline of whichever type is being considered. An element of $C^n[a,b]$ will be denoted by x or z, the value of n to be taken appropriate to the context. (x will usually be the function with which y agrees at the knots.)

The function norm will be the uniform norm;
$||x|| = \max |x(t)|, \; t_o \leq t \leq t_N$.

The vector norm will be the uniform one; if $\underset{\sim}{z} = [z_1, z_2, \ldots, z_N]^T$ then $||\underset{\sim}{z}|| = \max_i |z_i|$.

This induces the matrix norm, $||A|| = \max_i \sum_{j=1}^{n} |a_{ij}|$.

We shall make frequent use of the following result.

If A is strictly diagonally dominant, i.e. $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$,

$i=1,2,\ldots,n$ then A is invertible and $||A^{-1}|| \leqslant \max\limits_i \frac{1}{|a_{ii}| - \sum\limits_{j\neq i} |a_{ij}|}$.

The following results for certain tridiagonal matrices can be easily deduced from [7] and [8].

(i) If $A^{-1} = \begin{bmatrix} 2 & 1 & 0 & \ldots & 0 & 0 \\ 1 & 4 & 1 & \ldots & 0 & 0 \\ \ldots & & & & & \\ 0 & 0 & 0 & \ldots & 4 & 1 \\ 0 & 0 & 0 & \ldots & 1 & 2 \end{bmatrix}$, an $n \times n$ matrix (1.1)

then $a_{ij} = -T_{i-1}T_{n-j}/(3U_{n-2})$, $1 \leqslant i < j \leqslant n$, where T_r and U_r are the r-th Chebyshev polynomials of the first and second kinds respectively on $[-1,1]$ each with argument -2.

(ii) If $B^{-1} = \begin{bmatrix} \lambda_1 & \alpha_1 & 0 & \ldots & 0 \\ 1-\alpha_2 & \lambda_1 & \alpha_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ 0 & 0 & 0 & \ldots & \lambda_n \end{bmatrix}$

where $0 < \alpha_j < 1$, $j=1,2,\ldots,n$, $\lambda_j\lambda_{j+1} > 1$, $j=1,2,\ldots,n-1$, then

$$\lambda_i b_{ii} > 1, \quad (-1)^{i-j} b_{ij} > 0, \qquad i,j = 1,2,\ldots,n. \tag{1.2}$$

(iii) If $B^{-1} = \begin{bmatrix} 2 & \alpha_1 & 0 & \ldots & 0 \\ 1-\alpha_2 & 2 & \alpha_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ 0 & 0 & 0 & & 2 \end{bmatrix}$

and $\alpha,\beta \geqslant 0$, $0 < \alpha_j < 1$, $j=1,2,\ldots,n$, then

$$-\beta.2^{j-i} < 3(-1)^{j-i}[\alpha b_{ij}+\beta b_{i+1j}] < \alpha.2^{j-i+1}, \quad 1 \le j \le i-1,$$

$$3\alpha-2\beta < 6[\alpha b_{ii}+\beta b_{i+1i}] < 4\alpha, \tag{1.3}$$

$$3\beta-2\alpha < 6[\alpha b_{ii+1}+\beta b_{i+1i+1}] < 4\beta,$$

$$-\beta.2^{i-j+2} < 3(-1)^{i-j}[\alpha b_{ij}+\beta b_{i+1j}] < \alpha.2^{i-j+1}, \quad i+2 \le j \le n.$$

(iv) If
$$C^{-1} = \begin{bmatrix} -3 & 1-\alpha_1 & 0 & \cdots & 0 \\ \alpha_2 & -3 & 1-\alpha_2 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{1.4}$$

where $0 < \alpha_j < 1$, $j=1,2,\ldots,n$, then

$$|c_{i1}| < 9.2^{-i}/8, \quad |c_{in}| < 9.2^{i+1-n}/8, \quad i=1,2,\ldots,n.$$

(v) If
$$D^{-1} = \begin{bmatrix} \frac{1}{2} & 1 & 0 & \cdots & 0 & 0 \\ 1 & 4 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 4 & 1 \\ 0 & 0 & 0 & \cdots & 1 & \frac{1}{2} \end{bmatrix}, \; n \times n, \; n \ge 4 \tag{1.5}$$

then $d_{ij} = \frac{-1}{3} T_{i-2}T_{n-j-1}/U_{n-4}$, $1 \le i \le j \le n$,

where the argument of the Chebyshev polynomials is -2.

Chapter 2

Convergence Properties of the Natural Cubic Spline

We shall consider in this chapter the problem of estimating $||y-x||$ when y is the natural cubic spline which interpolates to x at the knots which satisfy $0 = t_0<t_1<\ldots<t_N = 1$. In addition we shall find conditions on the spacing of the knots which will ensure convergence as the maximum interval decreases to zero. It is of interest to determine if there are situations under which there may be divergence and it is shown that for a certain knot spacing the natural cubic spline operator is unbounded.

Results similar to these have been found in the simpler situation when the spline is assumed to be periodic. In particular Nord in [11] has exhibited a periodic cubic spline which diverges.

For the purposes of this chapter it will be sufficient to consider the spline on the interval $0 = t_0\leq t\leq t_N = 1$, and we take the following as the definition of the spline.

<u>Definition</u>

A <u>natural cubic spline</u> with the knots $t_0,t_1,\ldots,t_N$ is in $C^2[0,1]$ and is such that

(a) it is a polynomial of degree at most three in each interval

(b) $y_0^{(2)} = y_N^{(2)} = 0$.

An important property of the natural cubic spline is that it is determined uniquely by its values $y_0,y_1,\ldots,y_N$. This is well known, but in order to introduce some relations which will be required later a proof will be outlined here.

Existence and Uniqueness

Since $y^{(2)}$ is linear in each interval and is in $C[0,1]$ we can set $y_j^{(2)} = \kappa_j$, $j = 0,1,\ldots,N$ and write

$$y^{(2)}(t) = \frac{(t_{j+1}-t)}{h_j}\kappa_j + \frac{(t-t_j)}{h_j}\kappa_{j+1}, \quad t_j \leq t \leq t_{j+1}, \; j=0,1,\ldots,N-1.$$

Integrate this twice and impose the interpolation conditions $(y_j=x_j,\ j=0,1,\ldots,N)$; then

$$y(t) = \frac{1}{6h_j}\left\{[(t_{j+1}-t)^3-h_j^2(t_{j+1}-t)]\kappa_j + [(t-t_j)^3-h_j^2(t-t_j)]\kappa_{j+1}\right\} +$$

$$+\frac{1}{h_j}\left\{(t_{j+1}-t)x_j + (t-t_j)x_{j+1}\right\} \tag{2.1}$$

for $t_j \leq t \leq t_{j+1}$, $j = 0,1,\ldots,N-1$.

The imposition of the condition of continuity of $y^{(1)}$ at the interior knots, $t_1,t_2,\ldots,t_{N-1}$ leads to well known relations which will be written, (see [1]p.11)

$$(1-\alpha_j)\kappa_{j-1} + 2\kappa_j+\alpha_j\,\kappa_{j+1} = 6[t_{j-1},t_j,t_{j+1}]x, \quad j=1,2,\ldots,N-1, \tag{2.2}$$

where $\alpha_j = h_j/(h_{j-1}+h_j)$, $h_j = t_{j+1}-t_j$, and $[t_{j-1},t_j,t_{j+1}]$ is the second divided difference operator at the points t_{j-1},t_j,t_{j+1}.

Now $\kappa_0=y_0^{(2)}=y_N^{(2)}=\kappa_N=0$, and so (2.2) is a set of N-1 linear algebraic equations in the unknowns $\kappa_1,\kappa_2,\ldots,\kappa_{N-1}$. The matrix of this set of equations, B^{-1}, is strictly diagonally dominant (since $0<\alpha_j<1$) and so there is a unique solution. Consequently the spline is unique and can be constructed in each interval with the aid of (2.1). Further $||B|| \leq 1$ and so we have the following inequalities,

$$\max_j |\kappa_j| \leq 6 \max \left| [t_{j-1},t_j,t_{j+1}]x \right| \leq 3||x^{(2)}||, \quad 1 \leq j \leq N-1. \tag{2.3}$$

Another set of equations can be found in terms of $\lambda_0,\lambda_1,\ldots,\lambda_N$, where $\lambda_j = y_j^{(1)}$, by the use of Hermite's two point interpolation formula instead of (2.1). These can also be found in [1] and are, with the same notation as above,

$$2\lambda_0 + \lambda_1 = 3[t_0,t_1]x$$

$$\alpha_j\lambda_{j-1}+2\lambda_j+(1-\alpha_j)\lambda_{j+1} = 3\alpha_j[t_{j-1},t_j]x+3(1-\alpha_j)[t_j,t_{j+1}]x, \quad j=1,2,\ldots,N-1, \tag{2.4}$$

$$\lambda_{N-1}+2\lambda_N = 3[t_{N-1},t_N]x.$$

Convergence theorems

In the following it will be assumed that $y_i=x_i$, $i = 0,1,\ldots,N$.

Theorem 2.1

If $x \in C^2[0,1]$ then $||y-x|| \leq \frac{1}{2}h^2||x^{(2)}||$

Proof

Since $y_j-x_j = 0$, $j = 0,1,\ldots,N$, we have, from Lagrange's linear interpolation formula with remainder,

$$y(t)-x(t) = \tfrac{1}{2}(t-t_j)(t-t_{j+1})[y^{(2)}(t')-x^{(2)}(t')], t_j \leq t \leq t_{j+1}, t_j<t'<t_{j+1}, \tag{2.5}$$

and so, for $t_j \leq t \leq t_{j+1}$,

$$\max|y(t)-x(t)| \le \frac{1}{8} h_j^2 \,|\max y^{(2)}(t)-x^{(2)}(t)|$$
$$\le \frac{1}{8} h^2[||x^{(2)}||+\max(|\kappa_j|,|\kappa_{j+1}|)]$$
$$\le \frac{1}{8} h^2[4||x^{(2)}||] \quad \text{(from (2.3))}.$$

The right hand side of this inequality is independent of j and so the result follows. This theorem allows us to prove a result for natural cubic spline approximation which is analogous to Weierstrass' theorem for polynomial approximation.

Let $\sum_N(h)$ denote the space of natural cubic splines with N+1 knots* where $t_0=0$, $t_N=1$, and maximum interval length h; $\sum_N(h)$ is clearly a subspace of C[0,1].

Theorem 2.2.

Given $\varepsilon>0$, $x \in C[0,1]$ then there exists, for sufficiently small h, an element $s \in \sum_N(h)$ such that $||x-s||<\varepsilon$.

Proof

From Weierstrass' theorem we can find a polynomial p such that $||x-p||<\frac{1}{2}\varepsilon$.

Let $\pi = ||p^{(2)}||$, and choose a set of knots so that $h^2<\varepsilon/\pi$. Then if $s_j=p_j$, $j=0,1,\ldots,N$, it follows from theorem 2.1 that $||s-p||\le h^2\pi/2<\varepsilon/2$. Since $||x-s||\le||x-p||+||p-s||$ the theorem is proved.

Clearly we cannot conclude from this that the natural cubic spline always converges to any continuous function with which it agrees at the knots. However the next theorem leads to a simple criterion for such convergence. It is convenient to

* the knots are $t_0, t_1, \ldots, t_N$.

define here the interpolatory spline operator S_N, this is the projection operator from $C[0,1]$ to $\sum_N(h)$ such that if $y = S_N x$ then $y_i = x_i$, $i=0,1,\dots,N$. (c.f. [2] where a similar operator is defined for periodic cubic splines.)

Theorem 2.3

$$||S_N x - x|| < (1+0.75h^2/k^2)\omega(x;h)$$

where $k = \min_j h_j$, and ω denotes the modulus of continuity of x with interval length $h = \max_j h_j$.

Proof

Let $y = S_N x$, then for $t_i \leq t \leq t_{i+1}$ we may write

$$6h_i[y(t)-x(t)] = [(t_{i+1}-t)^3-h_i^2(t_{i+1}-t)]\kappa_i + [(t-t_i)^3-h_i^2(t-t_i)]\kappa_{i+1} + 6\left\{(t_{i+1}-t)(x_i-x(t))+(t-t_i)(x_{i+1}-x(t))\right\}. \tag{2.6}$$

Hence, for $t_i \leq t \leq t_{i+1}$, after some simple manipulation and estimation, it will follow that

$$\max|y(t)-x(t)| \leq \omega(x;h)+\frac{1}{24} h_i^2 \max\left\{|2\kappa_i+\kappa_{i+1}|, |\kappa_i+2\kappa_{i+1}|\right\}. \tag{2.7}$$

But, from (2.3), $\max_i|\kappa_i| \leq 6 \max_j|[t_{j-1},t_j,t_{j+1}]x|$

$$\leq 6\omega(x;h)\max_j \frac{1}{h_{j-1}h_j} \leq 6\omega(x;h)/k^2,$$

and so $\quad h_i^2 \max_i|\kappa_i| \leq 6\omega(x;h)h^2/k^2$,

which, when used in (2.7), will give the result.

The following is immediate.

Corollary

$$||S_N x - x|| \to 0 \text{ as } h \to 0 \text{ if } \max h_j/\min h_j \text{ is bounded.}$$

Another criterion for convergence will be found in terms of

$\lambda = \max_j h_{j+1}/h_j$, and $\mu = \min_j h_{j+1}/h_j$.

<u>Theorem</u> 2.4

If $\frac{1}{2} < \mu^2, \lambda^2 < 2$ then

$$||S_N x - x|| < [1 + \frac{1}{3}\max(2P+Q, P+2Q)]\omega(x;h)$$

where

$$P = 2\lambda/(2-\lambda^2), \quad Q = \mu/(2\mu^2-1).$$

<u>Proof</u>

It is clear from ~~eqn.~~ (2.7) that bounds are required for each of $|2\kappa_i+\kappa_{i+1}|, |\kappa_i+2\kappa_{i+1}|$.

Now, with the notation that the matrix of the equations (2.2) is B^{-1} we have

$$\kappa_i = 6\sum_{j=1}^{N-1} b_{ij}[t_{j-1},t_j,t_{j+1}]x, \kappa_{i+1} = 6\sum_{j=1}^{N-1} b_{i+1j}[t_{j-1},t_j,t_{j+1}]x;$$

and since $|[t_{j-1},t_j,t_{j+1}]x| \leq \dfrac{\omega(x;h)}{h_{j-1}h_j}$, it follows that

$$\frac{1}{6}\,\frac{|2\kappa_i+\kappa_{i+1}|}{\omega(x;h)} \leq \sum_{j=1}^{N-1} |2b_{ij} + b_{i+1j}|\cdot\frac{1}{h_{j-1}h_j}. \tag{2.8}$$

There is a similar expression for $\dfrac{1}{6}\,\dfrac{|\kappa_i+2\kappa_{i+1}|}{\omega(x;h)}$.

The use of the inequalities (1.3) will lead to the following inequality,

$$h_i^2\frac{|2\kappa_i+\kappa_{i+1}|}{4\omega(x;h)} < 2\left[\sum_{j=1}^{i-1}\frac{1}{2^{i-j}}\,\frac{h_i^2}{h_{j-1}h_j} + \frac{h_i}{h_{i-1}}\right] + \left[\frac{h_i}{h_{i+1}} + \sum_{j=i+2}^{N-1}\frac{1}{2^{j-i-1}}\,\frac{h_i^2}{h_{j-1}h_j}\right].$$

Now, $$\frac{h_i^2}{h_{j-1}h_j} = \frac{h_i^2}{h_{i-1}^2}\cdot\frac{h_{i-1}^2}{h_{i-2}^2}\cdots\cdot\frac{h_{j+1}^2}{h_j^2}\cdot\frac{h_j}{h_{j-1}} \leq (\lambda^2)^{i-j}.\lambda, \quad 1\leq j\leq i-1,$$

with a similar result when $i+2 \leq j \leq N-1$.

It follows that

$$h_i^2\,\frac{|2\kappa_i+\kappa_{i+1}|}{4\omega(x;h)} < 2\lambda\left[1+\sum_{j=1}^{i-1}\left(\frac{\lambda^2}{2}\right)^{i-j}\right]+\frac{1}{\mu}\left[1+\sum_{j=i+2}^{N-1}\left(\frac{1}{2\mu^2}\right)^{j-i-1}\right].$$

If $\frac{1}{2}<\mu^2$, $\lambda^2<2$ each of these geometric series can be bounded by its sum to infinity and the result will be that

$$h_i^2\,\frac{|2\kappa_i+\kappa_{i+1}|}{4\omega(x;h)} < \frac{4\lambda}{2-\lambda^2}+\frac{2\mu}{2\mu^2-1} = 2(2P+Q).$$

Similarly it can be proved that if $\frac{1}{2}<\mu^2$, $\lambda^2<2$,

$$h_i^2\,\frac{|\kappa_i+2\kappa_{i+1}|}{4\omega(x;h)} < \frac{2\lambda}{2-\lambda^2}+\frac{4\mu}{2\mu^2-1} = 2(P+2Q).$$

The result of the theorem follows when these results are used in (2.7).

Corollary

$||S_N x-x||\to 0$ as $h\to 0$ if $\frac{1}{2}+\varepsilon<\mu^2$, $\lambda^2<\frac{1}{2}-\eta$, where $\varepsilon,\eta>0$.

The proof is immediate.

Bounds for $||S_N||$

Upper bounds for $||S_N||$ can easily be calculated from the results of the last two theorems with the use of the inequality $\omega(x;h)\leq 2||x||$. However $||S_N||$ has practical implications and more precise bounds than these will be established.

Definition

The fundamental natural splines $L_0, L_1, \ldots, L_N$ are natural cubic splines which satisfy $L_j(t_i)=\delta_{ij}$, $0\leq i,\ j\leq N$.

With this definition we can write $S_N x=\sum_{j=0}^{N} x_j L_j$, consequently

$$||S_N|| = \max_{[0,1]}\sum_{j=0}^{N}|L_j(t)|. \tag{2.9}$$

Theorem 2.5

L_i changes sign in $[0,1]$ only when $t=t_j$, $j=0,1,\ldots,N$, $j\neq i$.

Proof

The proof is in two parts. First it will be shown that L_i vanishes in $[0,1]$ only at the knots $t_0,t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_N$. The proof will be completed by showing that $L_i^{(1)}$ is not zero at any of the knots except perhaps t_i.

Let $L_i(t')=0$, $t'\neq t_j$, $t_0<t'<t_N$. Then L_i has at least N+1 distinct zeros in $[0,1]$. By Rolle's theorem $L_i^{(1)}$ vanishes at least N times in $(0,1)$, and, again by Rolle's theorem, $L_i^{(2)}$ vanishes at least N-1 times in $(0,1)$. Consequently as $L_i^{(2)}(t_0) = L_i^{(2)}(t_N) = 0$ it follows that $L_i^{(2)}$ vanishes at not less than N+1 points in $[0,1]$. But $L_i^{(2)}$ is linear in each of the N intervals and so $L_i^{(2)} \equiv 0$ in at least one of them. Let this interval be $[t_k,t_{k+1}]$, then, as $L_i^{(2)}(t_k) = L_i^{(2)}(t_{k+1}) = 0$, L_i decomposes into two natural cubic splines, one with the knots $t_0,t_1,\ldots,t_k$, the other with the knots $t_{k+1},\ldots,t_N$. From the uniqueness property one of these splines is identically zero.

There are three cases to consider.

$i = 0,N$.

Consider L_0. Then for $t\geq t_{k+1}$, $L_0(t) = 0$. Further $k\neq 0$, for otherwise L_0 would be a non zero linear polynomial in $[t_0,t_1]$ and identically zero in $[t_1,t_2]$. This would contradict the continuity of $L_0^{(1)}$ at $t=t_1$. Therefore $L_0^{(r)}(t_k)=0$, $r=0,1,2$, and if $k\neq 1$ L_0 will be identically zero in $[t_{k-1},t_k]$. This argument can be repeated until we are lead to the conclusion that $L_0(t)=0$

for $t \geq t_1$. But this means that $L_0(t) = (t_1-t)^3/h_0^3$ for $t_0 \leq t \leq t_1$ which will contradict the constraint $L_0^{(2)}(t_0) = 0$.

A similar argument can be used to prove a contradiction when $i=N$.

$i=1,\ N-1$.

Take L_1; then $k \neq 1$, for otherwise $L_1^{(1)}$ would be discontinuous at t_1. If $k=0$ then L_1 is linear in $[t_0,t_1]$ and so can vanish only at the knot t_0. Therefore $t'>t_1$, and because $L_1^{(2)}(t_1)=0$ the argument above for L_0 can now be used to prove a contradiction as L_1 can be regarded as a fundamental natural cubic spline with the knots $t_1,t_2,\ldots,t_N$ and such that $L_1(t_1)=1$.

When $k \geq 2$ we have $L_1^{(r)}(t_k)=0$, $r = 0,1,2$; therefore $L_1(t)=0$ for $t \geq t_2$. That is to say L_1 is a natural cubic spline which satisfies $L_1(t_0) = L_1^{(2)}(t_0) = 0$; $L_1(t_1) = 1$; $L_1^{(r)}(t_2) = 0$, $r = 0,1,2$. These last conditions show that $L_1(t) = \dfrac{(t_2-t)^3}{h_1^3}$ in $[t_1,t_2]$, and the first ones that $L_1(t) = a\dfrac{(t-t_0)}{h_0} + b\dfrac{(t-t_0)^3}{h_0^3}$ in $[t_0,t_1]$ where $a+b=1$. When the continuity conditions are imposed at t_1 it will be found that h_0,h_1 have to satisfy $(h_0+h_o)(2h_0+h_1) = 0$, which is impossible

$2 \leq i \leq N-2$

Now $i \neq k,k+1$. For instance if $i=k$ then L_i would be linear in $[t_i,t_{i+1}]$ and zero in $[t_{i+1},t_{i+2}]$ which would contradict the continuity of $L_i^{(1)}$ at t_{i+1}. Similarly $i \neq k+1$.

To be definite let $i \leq k-1$, then $L_i(t)=0$ for $t \geq t_{i+1}$. Therefore L_i is a natural cubic spline with the knots $t_0,t_1,\ldots,t_{i+1}$ and such that $L_i(t_i)=1$, $L_i^{(1)}(t_{i+1})=0$. A count of the zeros of $L_i^{(2)}$ by the repeated use of Rolle's theorem will show that $L_i^{(2)}$ vanishes identically in an interval in $[t_0, t_{i+1}]$. This cannot be $[t_i, t_{i+1}]$ because of the continuity of $L_i^{(1)}$ at t_{i+1}; nor can it be $[t_{i-1}, t_i]$, for then the spline would vanish identically for $t \leq t_{i-1}$ and the continuity of $L_i^{(1)}$ at t_{i-1} would be contradicted. Therefore it is an interval which lie below t_{i-1}; but this means that L_i will be a natural cubic spline which satisfies $L_i^{(r)}(t_{i-1}) = L_i^{(r)}(t_{i+1}) = 0$, $r = 0,1,2$; $L_i(t_i) = 1$. It will be found that it is not possible to construct L_i with these properties. (In fact in $[t_{i-1}, t_i]$ we have $L_i(t) = (t-t_{i-1})^3/h_{i-1}^3$ and in $[t_i, t_{i+1}]$ $L_i(t) = (t_{i+1}-t)^3/h_i^3$. Consequently $L_i^{(1)}(t_i-0) = 3/h_{i-1}$ and $L_i^{(1)}(t_i+0) = -3/h_i$.)

So we have proved that L_i vanishes only at the knots at which it was prescribed to vanish. If it does <u>not</u> change sign at the knot t_j, $j \neq i$ then $L_i^{(1)}(t_j) = 0$. Since L_i vanishes at N points it follows that $L_i^{(1)}$ vanishes at not less than N-1 points which are not knots (one of these may be t_i but this is not important). Hence $L_i^{(1)}$ vanishes at least N times in $[t_0, t_N]$ and so $L_i^{(2)}$ vanishes at least N+1 times in $[t_0, t_N]$; this implies that $L_i^{(2)} \equiv 0$ in at least one interval. The arguments of the first part of the proof can now be adapted to show that $L_i^{(1)}(t_j) \neq 0$, $j \neq i$. Hence the spline changes sign only at the knots at which it was constrained to vanish.

(It can be shown that L_i does not vanish outside $[t_0, t_N]$, however we do not need this result.)

<u>Corollary</u>

Let $L(t) = \sum_{j=0}^{N} |L_j(t)|$

then for $t_i \leq t \leq t_{i+1}$

$$L(t) = \sum_{j=0}^{i} (-1)^{i-j} L_j(t) + \sum_{j=i+1}^{N} (-1)^{j-i-1} L_j(t). \tag{2.10}$$

Proof.

In $[t_i, t_{i+1}]$ L_i and L_{i+1} are each positive, and L_{i-1} and L_{i+2} are each negative etc. The result follows.

If the values $x_0, x_1, \ldots, x_N$ are each subject to an error of $\pm\varepsilon$ then, from (2.9), the maximum error in the natural cubic spline which uses these values instead of the correct ones will not exceed $\varepsilon||S_N||$. We shall now find upper bounds for $||S_N||$.
Let $\Lambda_i = \sum_{j=0}^{i} (-1)^{i-j} L_j + \sum_{j=i+1}^{N} (-1)^{j-i-1} L_j$, then $\Lambda_i(t) > 0$ for $t_i \leq t \leq t_{i+1}$.
Hence $||S_N|| = \max_{[0,1]} |L(t)| = \max_i \max_{[t_i, t_{i+1}]} |L(t)| = \max_i \max_{[t_i, t_{i+1}]} \Lambda_i(t)$.
It follows from the corollary that $\Lambda_i(t_j) = (-1)^{i-j}$, $0 \leq j \leq i$, $\Lambda_i(t_j) = -(-1)^{j-i}$, $i+1 \leq j \leq N$. (2.11)

The second divided differences of Λ_i are easily calculated and are given by

$$[t_{j-1}, t_j, t_{j+1}]\Lambda_i = \begin{cases} -2(-1)^{i-j}/(h_{j-1}h_j) & 1 \leq j \leq i-1 \\ -2/[h_{i-1}(h_{i-1}+h_i)], & j=i \\ -2/[h_{i+1}(h_i+h_{i+1})], & j=i+1 \\ 2(-1)^{j-i}/(h_{j-1}h_j), & i+2 \leq j \leq N-1. \end{cases} \tag{2.12}$$

Let $\Lambda_i^{(2)}(t_i) = p$, $\Lambda_i^{(2)}(t_{i+1}) = q$, then

$$\max_{[t_i, t_{i+1}]} |\Lambda_i(t)| \leq 1 + \frac{h_i^2}{24} \max(|2p+q|, |p+2q|). \tag{2.13}$$

Theorem 2.6

$$||S_N|| \leq 1 + \frac{3}{2} h^2/k^2.$$

Proof

Since Λ_i is a natural cubic spline it follows from (2.3) that

$$|p|,\ |q| \leq 6 \max_j |[t_{j-1},t_j,t_{j+1}]\Lambda_i| \leq 12/k^2.$$

The result follows from this and (2.13).

Theorem 2.7

If $\frac{1}{2} < \mu^2,\ \lambda^2 < 2$ then

$$||S_N|| \leq 1 + \frac{1}{3}\max\{2A+B, A+2B\}$$

where $A = \dfrac{2\lambda}{2-\lambda^2} - \dfrac{\mu}{1+\mu}$, $B = \dfrac{2\mu}{2\mu^2-1} - \dfrac{1}{1+\lambda}$.

Proof

Consider $2p+q = 6\sum_{j=1}^{N-1} [2b_{ij}+b_{i+1j}][t_{j-1},t_j,t_{j+1}]\Lambda_i$

where again the matrix of the equations is B^{-1}.

Then, with the use of (2.12),

$$\frac{2p+q}{12} = -\sum_{j=1}^{i-1} \frac{(-1)^{i-j}}{h_{j-1}h_j}[2b_{ij}+b_{i+1j}] - \frac{1}{h_{i-1}(h_{i-1}+h_i)}[2b_{ii}+b_{i+1i}]$$

$$- \frac{1}{h_{i+1}(h_i+h_{i+1})}[2b_{ii+1}+b_{i+1\,i+1}] + \sum_{j=i+2}^{N-1} \frac{(-1)^{j-i}}{h_{j-1}h_j}[2b_{ij}+b_{i+1j}].$$

The inequalities (1.3) will lead, after some manipulation, (c.f. the proof of theorem 2.4) to the result that

$$\frac{h_i^2}{12}|2p+q| < \frac{2}{3}\left\{2\left[\frac{2\lambda}{2-\lambda^2} - \frac{\mu}{1+\mu}\right] + \left[\frac{2\mu}{2\mu^2-1} - \frac{1}{1+\lambda}\right]\right\} = 2(2A+B)/3.$$

A similar consideration of p+2q leads to

$$\frac{h_i^2}{12}\,|p+2q| < \frac{2}{3}\left\{\left[\frac{2\lambda}{2-\lambda^2} - \frac{\mu}{1+\mu}\right] + 2\left[\frac{2\mu}{2\mu^2-1} - \frac{1}{1+\lambda}\right]\right\} = 2(A+2B)/3$$

The proof of the theorem is completed by the insertion of these results into (2.1

An Unbounded Spline Operator

We prove now a result for natural cubic splines which is similar to the one proved by Nord [11] for periodic cubic splines.

Let $\Lambda_i(t) = \sum_{j=0}^{i} (-1)^{i-j} L_j(t) + \sum_{j=i+1}^{N} (-1)^{j-i-1} L_j(t)$,

then, from (2.9) and (2.10),

$$||S_N|| = \max_i \max_{t_i \leq t \leq t_{i+1}} \Lambda_i(t).$$

However $\max \Lambda_i(t) \geq \Lambda_i(t_i+\frac{1}{2}h_i)$ for $t_i \leq t \leq t_{i+1}$

and so $$||S_N|| \geq \max_i \Lambda_i(t_i+\tfrac{1}{2}h_i).$$

With the aid of (2.1) we obtain easily

$$||S_N|| \geq \max_i \left\{1-h_i^2\left[\Lambda_i^{(2)}(t_i) + \Lambda_i^{(2)}(t_{i+1})\right]/16\right\}.$$

In order to simplify the presentation of the results we shall consider only values of i in the range $2 \leq i \leq N-3$. There will be a slight loss of generality in the results but this will not affect the conclusions

Lemma 2.8

If $2h_0 \leq h_2$, $h_{j-1} \leq h_{j+1}$, $j = 2,3,\ldots,i-2$, $h_{i-2} \leq h_{i-1}+h_i$,

and $h_{i+2} \leq h_i+h_{i+1}$, $h_{j+1} \leq h_{j-1}$, $j = i+2,\ldots,N-3$, $2h_{N-1} \leq h_{N-3}$

then

$$\Lambda_i(t_i+\tfrac{1}{2}h_i)\geq 1+0.75h_i^{\,2}[(h_i+h_{i+1})/h_{i-1}+(h_{i-1}+h_i)/h_{i+1}]/[3h_i^{\,2}+2h_i(h_{i-1}+h_{i+1})+h_{i-1}h_{i+1}].$$

Proof

Let $\Lambda_i^{(2)}(t_j) = \kappa_j$, $j=0,1,\ldots,N$.

Then the equations (2.2) can be rewritten as

$$[2+(2-\alpha_1)(1-\alpha_2)](\kappa_1+\kappa_2)+(2-\alpha_1)\alpha_2(\kappa_2+\kappa_3) = 6[d_1+(2-\alpha_1)d_2],$$

$$(1-\alpha_j)(\kappa_{j-1}+\kappa_j)+(2+\alpha_j-\alpha_{j+1})(\kappa_j+\kappa_{j+1})+\alpha_{j+1}(\kappa_{j+1}+\kappa_{j+2})=6[d_j+d_{j+1}],\ j=2,3,\ldots,N-3,$$

$$(1-\alpha_{N-2})(1+\alpha_{N-1})(\kappa_{N-3}+\kappa_{N-2})+[2+\alpha_{N-2}(1+\alpha_{N-1})](\kappa_{N-2}+\kappa_{N-1})=6[(1+\alpha_{N-1})d_{N-2}+d_{N-1}],$$

where we have defined $d_j = [t_{j-1},t_j,t_{j+1}]\Lambda_i$, $j = 1,2,\ldots,N-1$.

These can be rewritten in a form which is suitable for the application of (1.2), namely as

$$\lambda_1(\kappa_1+\kappa_2)+\beta_1(\kappa_2+\kappa_3) = 6[d_1/(2-\alpha_1)+d_2],$$

$$(1-\beta_j)(\kappa_{j-1}+\kappa_j)+\lambda_j(\kappa_j+\kappa_{j+1})+\beta_j(\kappa_{j+1}+\kappa_{j+2})=6(d_j+d_{j+1})/(1-\alpha_j+\alpha_{j+1}),\ j=2,3,\ldots,N-3, \tag{2.14}$$

$$(1-\beta_{N-2})(\kappa_{N-3}+\kappa_{N-2})+\lambda_{N-2}(\kappa_{N-2}+\kappa_{N-1})=6[d_{N-2}+d_{N-1}]/(1+\alpha_{N-1}),$$

where $\lambda_1 = (1-\alpha_2)+2/(2-\alpha_1)$ $\lambda_{N-2} = \alpha_{N-2}+2/(1+\alpha_{N-1})$,

$\beta_1 = \alpha_2$, $\beta_{N-2} = \alpha_{N-2}$,

$\lambda_j = (2+\alpha_j-\alpha_{j+1})/(1-\alpha_j+\alpha_{j+1})$, $\beta_j = \alpha_{j+1}/(1-\alpha_j+\alpha_{j+1})$, $j=2,3,\ldots,N-3$.

It is easily verified that $\lambda_j\lambda_{j+1}>1$, $j=1,2,\ldots,N-1$ and so the result (1.2) can be used to bound the elements in the inverse of the matrix of the equations (2.14).

If $\{b_{ij}\}$ are the elements in this inverse then

$$\kappa_i+\kappa_{i+1} = \sum_{j=1}^{N-2} b_{ij}\, r_j = \sum_{j=1}^{N-2} [(-1)^{i-j}b_{ij}][(-1)^{i-j}r_j] \tag{2.15}$$

where r_j denotes the right hand side of the j-th equation of (2.14). Now $(-1)^{i-j}b_{ij}>0$ and we shall find conditions on the spacing of the knots to ensure that $-(-1)^{i-j}r_j\geq 0$.

Consider for example $-(-1)^{i-1}r_1 = 12[h_2/(2-\alpha_1)-h_0]/(h_0h_1h_2)$; we see that if $h_2\geq 2h_0$ then $-(-1)^{i-1}r_1\geq 0$. Each of the different forms of r_j are considered in the same fashion and it will be found after some manipulation that for the distribution of of intervals which is given in the statement of the lemma

$$-h_i^2(\kappa_i+\kappa_{i+1})\geq 12h_i^2[(h_i+h_{i+1})/h_{i-1}+(h_{i-1}+h_i)/h_{i+1}]/[3h_i^2+2h_i(h_{i-1}+h_{i+1})+h_{i-1}h_{i+1}].$$

When this is used to replace $-h_i^2(\kappa_i+\kappa_{i+1})$ in $\Lambda_i(t_i+\frac{1}{2}h_i) =$ $= 1-h_i^2(\kappa_i+\kappa_{i+1})/16$ the result follows.

Theorem 2.9

There exists a set of knots for which $||S_N||\to\infty$ as $h\to 0$.

Proof

Let $N = 2M+1$, $2h_0\leq h_2$, $h_{j-1}\leq h_{j+1}$, $j = 2,3,\ldots,M-2$, $h_{M-1} = h_M^2 = h_{M+1}$, $h_{j+1}\leq h_{j-1}$, $j = M+2,\ldots,2M-2$, $2h_{2M}\leq h_{2M-2}$. Then $||S_{2M+1}||\geq\Lambda_M(t_M+\frac{1}{2}h_M)\geq 1+1.5(1+1/h_M)/[3+4h_M+h_M^2]$ which increases without limit as $h_M\to 0$.

If N=2M we consider in a similar fashion the interval $[t_M,t_{M+1}]$.

Convergence for $x \in C^4[0,1]$

The convergence theorems of the earlier part of this chapter were derived under general conditions on the function which was interpolated. We end the chapter by giving two theorems on the approximation by interpolating cubic splines which when x is restricted to being in $C^4[0,1]$ are best possible. A proof of the first theorem can be found in [9]; the second theorem will be proved here.

Theorem 2.10

Let y be a natural cubic spline and $x \in C^4[0,1]$, and let $y_i = x_i$, $i=0,1,\ldots,N$. Then for sufficiently large N there exist knots t_p, t_q, where $0<t_p<t_q<1$ and a constant K such that for $t_p \leq t \leq t_q$,

$$\max|x(t)-y(t)| \leq Kh^4$$
$$\max|x^{(1)}(t)-y^{(1)}(t)| \leq 4Kh^3$$
$$\max|x^{(2)}(t)-y^{(2)}(t)| \leq 8Kh^2.$$

Further, t_p, $1-t_q$ are $O(h \log h)$ as $h \to 0$.

Theorem 2.11

Let x,y be as in theorem 2.10 and in addition let $x^{(2)}(0) = x^{(2)}(1) = 0$, then

$$||x-y|| \leq \frac{3}{64} h^4 ||x^{(4)}||, \quad ||x^{(2)}-y^{(2)}|| \leq \frac{3}{8} h^2 \, ||x^{(4)}||.$$

Proof

We have from (2.5),

$$\max|y(t)-x(t)| \leq \frac{1}{8} h_j^2 \max|y^{(2)}(t)-x^{(2)}(t)|, \text{for } t_j \leq t \leq t_{j+1},\ j=0,1,\ldots,N-1. \quad (2.16)$$

~~Therefore,~~ Since $y^{(2)}$ is linear in $[t_j,t_{j+1}]$, we obtain from Lagrange's linear interpolation formula,

$$\left|y^{(2)}(t)-x^{(2)}(t)\right| \leq \frac{1}{h_j}\left|(t_{j+1}-t)(\kappa_j-x_j^{(2)})+(t-t_j)(\kappa_{j+1}-x_{j+1}^{(2)})\right|+\frac{1}{8}h_j^2\left|x^{(4)}(t')\right|$$

for $t_j \leq t \leq t_{j+1}$, $j=0,1,N-1$, and where $t_j<t'<t_{j+1}$.

It follows that for $t_j \leq t \leq t_{j+1}$,

$$\max|y^{(2)}(t)-x^{(2)}(t)| \leq \max(|\kappa_j-x_j^{(2)}|,|\kappa_{j+1}-x_{j+1}^{(2)}|)+\frac{1}{8}h_j^2||x^{(4)}||, \tag{2.17}$$

consequently we have to estimate $\max\limits_j|\kappa_j-x_j^{(2)}|$. In order to do this rewrite the equations of (2.2) as

$$(1-\alpha_j)e_{j-1}+2e_j+\alpha_j e_{j+1}=6[t_{j-1},t_j,t_{j+1}]x-[(1-\alpha_j)x_{j-1}^{(2)}+2x_j^{(2)}+\alpha_j x_{j+1}^{(2)}] \tag{2.18}$$

$j = 1,2,\ldots,N-1,$

where $e_j = \kappa_j - x_j^{(2)}$, $j = 0,1,\ldots,N$.

With the use of Peano's method the right hand side of (2.18) can be written $\frac{1}{4}\left[(1-\alpha_j)h_{j-1}^2+\alpha_j h_j^2\right] x^{(4)}(t')$, where $t_{j-1} \leq t' \leq t_{j+1}$.

Now y is a natural spline and so with the assumption made in the statement of the theorem it follows that $e_0 = e_N = 0$. Hence, with the uniform matrix norm,

$$||\underset{\sim}{e}|| \leq \frac{1}{4}\max_j|(1-\alpha_j)h_{j-1}^2+\alpha_j h_j^2|\,||x^{(4)}|| \leq \frac{1}{4}h^2||x^{(4)}||.$$

When this is inserted in (2.17) the result is that

$$\max|y^{(2)}(t)-x^{(2)}(t)| \leq \frac{3}{8}h^2||x^{(4)}||, \quad t_j \leq t \leq t_{j+1}, \tag{2.19}$$

and when this is used in (2.16) we obtain

$$\max|y(t)-x(t)| \leq \frac{3}{64}h^4||x^{(4)}||, \quad t_j \leq t \leq t_{j+1}. \tag{2.20}$$

Since the right hand sides of each inequality is independent of j the theorem is proved.

Chapter 3.

Sard's Best Quadrature Formulae of Second Order

We prove here results on the convergence of Sard's quadrature formula by estimating the size of the Peano kernel. When the knots are equally spaced and the integral which is being approximated is $\int_0^1 x(t)dt$ we find explicit equations for the quadrature weights.

We begin by recalling the definition of a quadrature formula and state Schoenberg's theorem on the connection between Sard's second order formula and natural cubic splines.

Note that in this chapter $t_0=0$, $t_N=1$.

Definition

An expression of the form

$$\int_0^1 w(s)x(s)ds = \sum_{j=0}^{N} H_j x(t_j) + R(x) \tag{3.1}$$

is called a quadrature formula with remainder R. If R vanishes when x is any polynomial of degree n-1 then the quadrature formula is said to be of order n.

We shall be concerned in this chapter solely with the case n=2, and in order that the problem of finding the quadrature weights $H_0, H_1, \ldots, H_N$ should not be trivial we shall assume that $N \geq 2$. Then for $x \in C^2[0,1]$, by Peano's method,

$$R(x) = \int_0^1 k(t)x^{(2)}(t)dt \tag{3.2}$$

where k, the Peano kernel of the second order formula, is given by

$$k(t) = \int_0^1 w(s)K_1^{+}(s-t)dt - \sum_{j=0}^{N} H_j\, K_1^{+}(t_j-t) \tag{3.3}$$

The problem which gives rise to Sard's best quadrature of order two is that of finding $H_0, H_1, \ldots, H_N$ to minimize

$$\int_0^1 [k(t)]^2 dt \tag{3.4}$$

subject to the constraints

$$\int_0^1 w(s)s^r ds = \sum_{j=0}^{N} H_j t_j^{\,r}, \quad r = 0,1. \tag{3.5}$$

We note here that, because of (3.5), $k(t_0) = k(t_N) = 0$.

<u>Theorem</u> 3.1 (Schoenberg)

If y is a natural cubic spline with the knots $t_0, t_1, \ldots, t_N$ and if (3.1) is a best quadrature formula of order two then $R(y) = 0$.

The following is an immediate consequence of this theorem and of the definition of the fundamental natural cubic splines.

<u>Corollary</u> 1

The quadrature weights in Sard's second order best quadrature formula are given by

$$H_j = \int_0^1 w(t)\, L_j(t) dt, \quad j = 0,1,N.$$

A less obvious result which will be needed in the discussion of convergence is the next corollary.

<u>Corollary</u> 2

If k is the Peano kernel of Sard's best quadrature formula of order two then

$$\int_0^1 [k(t)]^2 dt = \int_0^1 w(t)[m(t)-y(t)]dt,$$

where

$$m(t) = \int_0^1 w(s)K_3^+(s-t)ds - K_3^+(-t)\int_0^1 (1-s)w(s)ds - K_3^+(1-t)\int_0^1 sw(s)ds, \qquad (3.6)$$

and y is the natural cubic spline which agrees with m at the knots.

<u>Proof</u>

Let $u^{(2)}(t)=k(t)$, $0\leq t\leq 1$,

$$=\int_0^1 w(s)K_1^+(s-t)ds - \sum_{j=0}^N H_j K_1^+(t_j-t).$$

Then

$$\int_0^1 [k(t)]^2 dt = \int_0^1 k(t)u^{(2)}(t)dt = R(u).$$

Now write

$$u^{(2)}(t) = \left[\int_0^1 w(s)K_1^+(s-t)-f^{(2)}(t)\right] - \left[\sum_{j=0}^N H_j K_1^+(t_j-t)-f^{(2)}(t)\right]$$

where $f \in C^2[0,1]$ is to be found so that the second square bracket is the second derivative of a natural cubic spline with the knots $t_0, t_1, \ldots, t_N$. For this to be true we must have

$$\sum_{j=0}^N H_j K_1^+(t_j-t) - f^{(2)}(t) = \sum_{j=0}^N b_j K_1^+(t_j-t)$$

where $\sum_{j=0}^N b_j t_j^r = 0$, $r = 0,1$.

Hence

$$f^{(2)}(t) = \sum_{j=0}^N (H_j-b_j)K_1^+(t_j-t).$$

Since $\sum_{j=0}^N (H_j-b_j)t_j^r = \int_0^1 w(s)s^r ds$, $r = 0,1$,

are the only restrictions which have to be imposed we can choose

$H_j - b_j = 0$, $j = 1,2,\ldots,N-1$ which will leave a pair of equations to be solved for $H_0 - b_0$, $H_N - b_N$. These give

$$H_0 - b_0 = \int_0^1 (1-s)w(s)ds, \quad H_N - b_N = \int_0^1 sw(s)ds.$$

Consequently if

$$f^{(2)}(t) = K_1^+(t_0 - t)\int_0^1 (1-s)w(s)ds + K_1^+(t_N - t)\int_0^1 sw(s)ds$$

then

$$R(u) = R(m)$$

where

$$m^{(2)}(t) = \int_0^1 w(s)K_1^+(s-t)ds - f^{(2)}(t)$$

We note finally that

$$R(m) = \int_0^1 w(s)m(s)ds - \sum_{j=0}^{N} H_j m(t_j),$$

and so, if y is the natural cubic spline such that $y_j = m_j$, $j = 0,1,\ldots,N$,

$$R(m) = R(m-y) = \int_0^1 w(s)[m(s)-y(s)]ds.$$

Notes

1. The function m-y is the Rodriques function for the quadrature formula. This term was introduced by Schoenberg in [15]

2. It is not difficult to verify that
$$m^{(2)}(t_0) = m^{(2)}(t_N) = 0$$
and that $m^{(4)}(t) = w(t) \quad 0<t<1.$

Convergence

Theorem 3.2

If $x \in C^2[0,1]$ then $|R(x)| \leq \frac{1}{2}h^2||x^{(2)}||\int_0^1 |w(t)|dt$

Proof

Let y be the natural cubic spline such that $y_i=x_i$, $i = 0(1)N$. Then from theorem (2.1), $||x-y||\leq\frac{1}{2}h^2||x^{(2)}||$. Since $R(y) = 0$ it follows that

$$|R(x)|=|R(x-y)|=\left|\int_0^1 w(t)[x(t)-y(t)]dt\right|\leq\tfrac{1}{2}h^2||x^{(2)}||\int_0^1 |w(t)|dt,$$

which was to be proved.

The results of theorems 2.3 and 2.4 can be used to furnish proofs of the convergence of Sard's best quadrature for particular spacings of the quadrature points. As these proofs of convergence are similar in principle to that of theorem 3.2 only the statements will be given.

Theorem 3.3

If $x \in C[0,1]$, $h = \max h_j$, $k = \min h_j$, then

$$|R(x)| \leq \left\{1 + \frac{3}{4}\, h^2/k^2\right\}\omega(x;h)\int_0^1 |w(t)|dt.$$

Theorem 3.4

If $x \in C[0,1]$, $\frac{1}{2}<\mu^2$, $\lambda^2<2$, where $\mu = \min \frac{h_{j+1}}{h_j}$, $\lambda = \max \frac{h_{j+1}}{h_j}$

then

$$|R(x)|\leq\left[1+\frac{1}{3}\max(2P+Q,P+2Q)\right]\omega(x;h)\int_0^1 |w(t)|dt.$$

If the conditions on x are strengthened the next theorem follows immediately from theorem 2.11.

Theorem 3.5

If $x \in C^4[0,1]$, $x^{(2)}(0) = x^{(2)}(1) = 0$ then

$$|R(x)| \leq \frac{3}{64} h^4 ||x^{(4)}|| \int_0^1 |w(t)| dt.$$

The last problem which will be considered in this section is that of finding the order of convergence if we assume that $x \in C^4[0,1]$. The result is stated in theorem 3.7, however a preliminary result will be required.

Lemma 3.6

$$\int_0^1 [k(t)]^2 dt = O(h^4) \text{ as } h \to 0 \text{ if } w \in C[0,1].$$

Proof

From corollary 2 of theorem 3.1 we have, with the notation used there,

$$\int_0^1 [k(t)]^2 dt \leq ||m-y|| \int_0^1 |w(t)| dt.$$

Since $m^{(2)}(0) = m^{(2)}(1) = 0$ (as pointed out in the second note to the same corollary) it follows from theorem 2.11 that

$$||m-y|| \leq \frac{3}{64} h^4 ||m^{(4)}|| = \frac{3}{64} h^4 ||w||,$$

and hence

$$\int_0^1 [k(t)]^2 dt \leq \frac{3}{64} h^4 ||w|| . \int_0^1 |w(t)| dt,$$

which is the required result.

Theorem 3.7

If $x \in C^4[0,1]$ and $w \in C[0,1]$ then

$$|R(x)| = O(h^{5/2}).$$

Proof

Let $p(t) = \frac{1}{6}\left[(1-t)^3 x_0^{(2)} + t^3 x_N^{(2)}\right]$

then $x-p \in C^4[0,1]$, and $x^{(2)}-p^{(2)}$ vanishes at $t = 0,1$.

It follows from theorem 3.5 that

$$|R(x-p)| \leq \frac{3}{64} h^4 ||x^{(4)}|| \int_0^1 |w(t)|dt. \tag{3.7}$$

Since $|R(x)| \leq |R(x-p)| + |R(p)|$ we shall now bound $|R(p)|$.

Let z be the natural cubic spline which agrees with p at the knots. The integration by parts twice of $\int_0^1 [p^{(2)}(t)-z^{(2)}(t)]^2 dt$ leads to

$$\int_0^1 \left[p^{(2)}(t)-z^{(2)}(t)\right]^2 dt = x_N^{(2)}\left[p_N^{(1)}-z_N^{(1)}\right] - x_0^{(2)}\left[p_0^{(1)}-z_0^{(1)}\right]. \tag{3.8}$$

Let $z_j^{(1)} = \lambda_j$, $p_j^{(1)} = \pi_j$, $j = 0(1)N$, then the equations (2.4) can be written with this notation as

$$2(\lambda_0-\pi_0)+(\lambda_1-\pi_1) = \tfrac{1}{2}h_0 x_0^{(2)}$$

$$\alpha_j(\lambda_{j-1}-\pi_{j-1})+2(\lambda_j-\pi_j)+(1-\alpha_j)(\lambda_{j+1}-\pi_{j+1}) = 0, \quad j = 1(1)N-1$$

$$(\lambda_{N-1}-\pi_{N-1})+2(\lambda_N-\pi_N) = -\tfrac{1}{2} h_{N-1} x_N^{(2)}.$$

From the inequalities (1.3) we deduce that

$$|\lambda_0-\pi_0| < \frac{1}{3}\left[h_0|x_0^{(2)}| + h_{N-1}|x_N^{(2)}|.2^{-N}\right]$$

$$|\lambda_N-\pi_N| < \frac{1}{3}\left[h_0|x_0^{(2)}|.2^{-N} + h_{N-1}|x_N^{(2)}|\right].$$

Consequently, when these are inserted into (3.8),

$$\int_0^1 \left[p^{(2)}(t)-z^{(2)}(t)\right]^2 dt < \frac{1}{3}\left[h_0|x_0^{(2)}|^2 + h_{N-1}|x_N^{(2)}|^2 + (h_0+h_{N-1})|x_0^{(2)}x_N^{(2)}|.2^{-N}\right]. \quad (3.9)$$

Now $|R(p)|^2 = |R(p-z)|^2 \leqslant \int_0^1 [k(t)]^2 dt \int_0^1 \left[p^{(2)}(t)-z^{(2)}(t)\right]^2 dt,$

and so with the result of lemma 3.7 and inequality (3.9) we see that

$$|R(p)| = O(h^{5/2}).$$

The combination of this with (3.7) gives the result.

Equal interval formulae

Note In this section it will be assumed that $w(t)=1$, $0 \leqslant t \leqslant 1$, and

that $t_j = jh$, $j = 0(1)N$ where $h = 1/N$. (3.10)

The weights for Sard's quadrature formula of order two have been tabulated in [12]*, however they can be expressed in terms of T_r and U_r the Chebyshev polynomials of the first and second kinds respectively, each with argument -2.

Theorem 3.8

(i) $H_0 = H_N = \frac{1}{12} h[3+(1-T_N)/U_{N-1}]$

(ii) $H_i = \frac{1}{2} h[2-(U_{i-1}+U_{N-i-1})/U_{N-1}]$, $i=1(1)N-1$

(iii) $\int_0^1 [k(t)]^2 dt = \frac{1}{144} h^4 \left[\frac{1}{5} + \frac{1}{3} h(1-T_N)/U_{N-1}\right].$

Proof

If y is a natural cubic spline with the knots defined

---

* See also 'Linear Approximation' by A. Sard. A.M.S. Colloquium Pub. 1963.

in (3.10) then, from the Euler-Maclaurin sum formula

$$\int_0^1 y(t)dt = h\left[\tfrac{1}{2}y_0+\sum_{j=1}^{N-1} y_j+\tfrac{1}{2}y_N\right] - \frac{1}{12}h^2\left[y_N^{(1)}-y_0^{(1)}\right].$$

(i) Therefore, with $y=L_0$, it follows that

$$H_0 = \tfrac{1}{2}h - \frac{1}{12}h^2\left[L_o^{(1)}(1)-L_o^{(1)}(0)\right]. \tag{3.11}$$

Let $\lambda_j = L_0^{(1)}(t_j)$, $j = 0(1)N$, then from (2.4) with $\alpha_j = \tfrac{1}{2}$, $j = 1(1)N-1$,

$$2\lambda_0+\lambda_1 = -3/h,$$

$$\lambda_0+4\lambda_1+\lambda_2 = -3/h$$

$$\lambda_{j-1}+4\lambda_j+\lambda_{j+1} = 0, \quad i = 2(1)N-1$$

$$\lambda_{N-1}+2\lambda_N = 0.$$

These can be solved explicitly for λ_0 and λ_N by the use of the result (1.1) to give

$$h\lambda_0 = (T_N+T_{N-1})/U_{N-1},\ h\lambda_N = (T_0+T_1)/U_{N-1}.$$

When these are substituted into (3.11) the stated expression for H_0 will be found after some manipulation. By symmetry $H_N = H_0$.

(ii) The calculation of $H_1,H_2,\ldots,H_{N-1}$ proceeds in the same fashion and will not be given.

(iii) From corollary 2 of theorem 3.1 it is clearly necessary to calculate

$$\int_0^1 [m(t)-y(t)]dt \tag{3.12}$$

where $m(t) = \frac{1}{24}(t^4-2t^3+2t-1)$ (3.13)

and y is the natural cubic spline which agrees with m at the quadrature points as knots.

The Euler-Maclaurin sum formula gives

$$\int_0^1 [m(t)-y(t)]dt = -\frac{1}{12}h^2\left[m^{(1)}(t)-y^{(1)}(t)\right]_0^1 + \frac{1}{720}h^4. \tag{3.14}$$

Let $\lambda_i = y^{(1)}(t_i)$, $\zeta_i = m^{(1)}(t_i)$, $i = 0(1)N$. Then from (2.4) with $\alpha_j = \frac{1}{2}$ we have

$$2(\zeta_0-\lambda_0) + (\zeta_1-\lambda_1) = h^3/24$$

$$(\zeta_{j-1}-\lambda_{j-1})+4(\zeta_j-\lambda_j)+(\zeta_{j+1}-\lambda_{j+1}) = 0,\ j = 1,2,\ldots,N-1.$$

$$(\zeta_{N-1}-\lambda_{N-1})+2(\zeta_N-\lambda_N) = -h^3/24.$$

The use of (1.1) leads to the result that

$$\zeta_0-\lambda_0 = -(\zeta_N-\lambda_N) = \frac{1}{72}h^3(1-T_N)/U_{N-1},$$

and when these are substituted into (3.14) the formula stated in the theorem will be found.

<u>Corollary</u>

(i) $H_j>0$, $j = 0,1,2,\ldots,N$

(ii) $H_{j-1}+4H_j+H_{j+1} = 6h$, $j = 2,3,\ldots,N-2$.

The proofs of these results are straightforward and will be omitted. The second part of the corollary provides an alternative method for the calculation of $H_1,H_2,\ldots,H_{N-1}$, since it is easy to see that we also have

$$4H_1+H_2 = 5.5h = H_{N-2}+4H_{N-1}.$$

When $H_1, H_2, \ldots, H_{N-1}$ have been calculated H_0, H_N can be found from

$$H_0 = \frac{1}{12}\,[7h-2H_1] = H_N.$$

Chapter 4.

The Natural Quintic Spline and Sard's Third Order Best Quadrature

This chapter will be devoted to an examination of the convergence of Sard's third order formula when the knots are equally spaced in $[0,1]$. This will entail, as a result of Schoenberg's theorem, finding the order of approximation of the natural quintic spline. The chapter concludes with a useful theorem which states that in general the addition of an extra knot in a best quadrature formula will decrease the Peano kernel. (This is proved for the third order formula only, however the proof is easily adapted for higher order formulae.)

Note In this chapter except where otherwise stated the knots are given by $t_i = ih = i/N$, $i = 0,1,\ldots,N$, however the definitions which are made here remain valid for a general distribution.

Definition 4.1

A quintic spline with the knots $t_0,t_1,\ldots,t_N$ is in $C^4[0,1]$ and in each interval $[t_j,t_{j+1}]$ is a polynomial of degree at most five. Such a spline has the general form

$$\sum_{i=0}^{N} d_i\, K_5^+(t_i-t)$$

Definition 4.2

A natural quintic spline with the knots $t_0,t_1,\ldots,t_N$ is a quintic spline with these knots such that $y_0^{(3)}=y_0^{(4)}=y_N^{(3)}=y_N^{(4)}=0$. This spline has general form

$$p_2(t) + \sum_{i=0}^{N} d_i\, K_5^+(t_i-t)$$

where $\sum_{i=0}^{N} d_i t_i^r = 0$, $r = 0,1,2$ and p_2 is an arbitrary quadratic,

Lemma 4.1

If y is a quintic spline with the knots $t_i=ih$, $i = 0,1,\ldots,N$ then, with $y_i^{(2)}=\kappa_i$, $y_i^{(4)}=g_i$, $i = 0,1,\ldots,N$, the following equations hold,

$$\begin{aligned} &h^4[59g_0+93g_1+27g_2+g_3] = 120[\Delta^3 y_0-h^3y_0^{(3)}] \\ &h^4[g_{i-2}+26g_{i-1}+66g_i+26g_{i+1}+g_{i+2}] = 120\delta^4 y_i, \quad i = 2,3,\ldots,N-2, \\ &h^4[g_{N-3}+27g_{N-2}+93g_{N-1}+59g_N] = 120[h^3y_N^{(3)}-\nabla^3 y_N]. \end{aligned} \tag{4.1}$$

$$\begin{aligned} h^2[\kappa_0-\kappa_1] &= -h^4[2g_0+g_1]/6 - h^3y_0^{(3)} \\ h^2[\kappa_{i-1}+4\kappa_i+\kappa_{i+1}] &= h^4[7g_{i-1}+16g_i+7g_{i+1}]/60+6\delta^2 y_i, \quad i = 1,2,\ldots,N-1 \\ h^2[-\kappa_{N-1}+\kappa_N] &= -h^4[g_{N-1}+2g_N]/6 + h^3y_N^{(3)}. \end{aligned} \tag{4.2}$$

Proof

The second relations in (4.1) are well known, see for example [1] p.127. The remaining ones can be found by the method of undetermined coefficients. For example to establish the first of (4.1) we require scalars $a_0,a_1,\ldots,a_8$ so that

$$h^4[a_0g_0+a_1g_1+a_2g_2+a_3g_3]-[a_4y_0+a_5y_1+a_6y_2+a_7y_3]-h^3a_8y_0^{(3)}, \quad \sum_{i=0}^{8}|a_i| \neq 0,$$

vanishes identically when y is a quintic spline. Therefore it must vanish when $y(t)=(t-t_0)^r$, $r = 0,1,\ldots,5$ and when $y(t) = K_5^+(t-t_i)$, $i = 1,2$. The solution of the equations which arise from those conditions will lead to the desired result. The others can be verified similarly.

If $y_0, y_1, \ldots, y_N, y_0^{(3)}, y_0^{(4)}, y_N^{(3)}, y_N^{(4)}$ are prescribed we see that (4.1) form a set of N-1 linear algebraic equations which can be solved for $g_1, g_2, \ldots, g_{N-1}$. Furthermore the equations of (4.2) can then be solved for $\kappa_0, \kappa_1, \ldots, \kappa_N$ when these are known. Consequently the quintic spline can be constructed in the interval $[t_i, t_{i+1}]$ from

$$y(t) = \sum_{r=0}^{2} \frac{2^{2r+1}}{(2r+1)!} h^{2r}\left[B_{2r+1}\left(\frac{t-t_i}{2h_i}\right)y_i^{(2r)} + B_{2r+1}\left(\frac{t_{i+1}-t}{2h_i}\right) y_{i+1}^{(2r)}\right], \tag{4.3}$$

where B_s is the s-th Bernoulli polynomial on $0 \leq t \leq 1$.

We note that a natural quintic spline is uniquely determined by its knot values since for this spline we have $y_0^{(3)} = y_0^{(4)} = y_N^{(3)} = y_N^{(4)} = 0$.

<u>Theorem</u> 4.2

Let $z(t) = \sum_{j=0}^{N} d_j K_5^{+}(t_j - t) + x(t)$, $0 \leq t \leq 1$,

where $x \in C^6[0,1]$ and $z^{(3)}(t_0) = z^{(4)}(t_0) = z^{(3)}(t_N) = z^{(4)}(t_N) = 0$. Let y be the natural quintic spline such that $y(t_i) = z(t_i)$, $i = 0,1,\ldots,N$. Then

$$||y-z|| \leq h^6 ||x^{(6)}||/8.$$

<u>Proof</u>

Let $e_i = y_i^{(4)} - z_i^{(4)}$, $i = 0,1,\ldots,N$. Then (4.1) can be rewritten

$$h^4[93e_1+27e_2+e_3] = r_1(z)$$

$$h^4[e_{i-2}+26e_{i-1}+66e_i+26e_{i+1}+e_{i+2}] = r_i(z), \quad i = 2,3,\ldots,N-2 \tag{4.4}$$

$$h^4[e_{N-3}+27e_{N-2}+93e_{N-1}] = r_{N-1}(z),$$

where

$$r_1(z)=120[\Delta^3 z_0-h^3 z_0^{(3)}]-h^4[59z_0^{(4)}+93z_1^{(4)}+27z_2^{(4)}+z_3^{(4)}],$$

$$r_i=120\delta^4 z_i-h^4[z_{i-2}^{(4)}+26z_{i-1}^{(4)}+66z_i^{(4)}+26z_{i+1}^{(4)}+z_{i+2}^{(4)}], \quad i = 2,3,\ldots,N-2,$$

$$r_{N-1}=120[h^3 z_N^{(3)}-\nabla^3 z_N]-h^4[z_{N-3}^{(4)}+27z_{N-2}^{(4)}+93z_{N-1}^{(4)}+59z_N^{(4)}].$$

We have purposely left the components $z_0^{(3)},\ldots,z_N^{(4)}$ in these equations even though they have been assumed to be zero. For consider

$$r_1(z) = \sum_{j=0}^{N} d_j r_1(K_5^+(t_j-t)) + r_1(x);$$

r_1 was constructed to vanish for any quintic spline, consequently $r_1(K_5^+(t_j-t))=0$, $j=0,1,\ldots,N$. Similarly $r_i(K_5^+(t_j-t))=0$, $j=0,1,\ldots,N$. It follows that $r_i(z) = r_i(x)$, $i = 1,2,\ldots,N-1$. We can now use Peano's method for finding a form of $r_i(x)$ which is valid when $x \in C^6[0,1]$. Some rather tedious manipulation leads to the results that

$$r_1(z)=-15h^6 x^{(6)}(t_1'),r_i(z)=-10h^6 x^{(6)}(t_i'),i=2,3,\ldots,N-2,r_{N-1}(z)=-15h^6 x^{(6)}(t_{N-1}').$$

Consequently we see that since the matrix of the equations is strictly diagonally dominant, $\max_i |z_i^{(4)}-y_i^{(4)}| \le 5h^2||x^{(6)}||/4$.

We can deal with (4.2) in a similar fashion to prove that

$$\max_i |z_i^{(2)} - y_i^{(2)}| \leq 3h^4 ||x^{(6)}||/4.$$

It remains to insert these inequalities in the following inequality which is not difficult to prove,

$$||z(t)-y(t)|| \leq h^2 \max_i |z_i^{(2)} - y_i^{(2)}|/8 + 5h^4 \max_i |z_i^{(4)} - y_i^{(4)}|/384$$
$$+ 61h^6 ||x^{(6)}||/4608,$$

to give the result. (Note that we have simplified a multiplier from 71/576 to 1/8.)

We see that if we set $d_0 = d_1 = \ldots = d_N = 0$ we obtain the following result.

<u>Corollary</u>

If $x_0^{(3)} = x_0^{(4)} = x_N^{(3)} = x_N^{(4)} = 0$ and y is the natural quintic spline which agrees with x at the knots then

$$||x-y|| \leq h^6 ||x^{(6)}||/8.$$

<u>Sard's best quadrature of third order.</u>

This formula arises in the same way as the second order formula except that in this case the remainder is required to vanish for all quadratic polynomials. The Peano kernel ℓ of the quadrature formula is given by

$$\ell(t) = \int_0^1 w(s) K_2^+(s-t)ds - \sum_{i=0}^N H_i K_2^+(t_i - t).$$

The proof of the next result is similar to that of theorem 2.2 corollary 2 and so will not be given.

Lemma 4.3

If $m \in C^3[0,1]$ and is such that

$$m^{(3)}(t) = \int_0^1 w(s)K_2^+(s-t)ds + \sum_{i=0}^{N} d_i K_2^+(t_i-t)$$

where $\int_0^1 w(s)s^r ds + \sum_{i=0}^{N} d_i t_i^{\,r} = 0, \quad r = 0,1,2$

then

$$\int_0^1 [\ell(t)]^2 dt = \int_0^1 w(t)[m(t)-y(t)]dt$$

where y is any natural quintic spline such that $m(t_i)=y(t_i)$, $i = 0,1,\ldots,N$. (The function $m-y$ is the Rodrigue function.)

Theorem 4.4

$$\int_0^1 [\ell(t)]^2 dt < h^6 ||w|| \int_0^1 |w(t)|dt/8.$$

Proof

From lemma 4.3 we have, with the same notation,

$$\int_0^1 [\ell(t)]^2 dt = \int_0^1 w(t)[m(t)-y(t)]dt,$$

and so

$$\int_0^1 [\ell(t)]^2 dt \leq ||m-y|| \int_0^1 |w(t)|dt.$$

It remains to note that $m^{(3)}(t_0) = m^{(4)}(t_0) = m^{(3)}(t_N) = m^{(4)}(t_N) = 0$ and use theorem 4.2 with x replaced by

$$-\int_0^1 w(s)K_5^+(s-t)ds.$$

Corollary

If R is the remainder in Sard's best quadrature formula of order three with quadrature points given by $t_i=ih=i/N$, $i=0,1,\ldots,N$,

then

$$|R(x)| \leqslant ||w||.h^3\sqrt{\int_0^1 [x^{(3)}(t)]^2dt}/(2\sqrt{2}).$$

Proof

Since $R(x) = \int_0^1 \ell(t)x^{(3)}(t)dt$ we have by Schwartz's inequality

$$|R(x)|^2 \leqslant \int_0^1 [\ell(t)]^2dt \int_0^1 [x^{(3)}(t)]^2dt$$
$$\leqslant ||w||^2 \, h^6 \int_0^1 [x^{(3)}(t)]^2dt/8$$

which gives the required result.

The next result is also easily proved.

Theorem 4.5

If $x \in C^6[0,1]$, $x_0^{(3)} = x_0^{(4)} = x_N^{(3)} = x_N^{(4)} = 0$

then

$$|R(x)| \leqslant h^6||x^{(6)}|| \int_0^1 |w(t)|dt/8.$$

We conclude this chapter with the proof of a theorem which shows that in general the addition of an extra point in a best quadrature formula gives rise to a better quadrature formula in the sense that the Peano kernel is reduced. In the theorem we assume that $t_0=0$, $t_N=1$ without any further restrictions on the knot spacing.

Theorem 4.6

If

$$\int_0^1 w(t)x(t)dt = \sum_{i=0}^{N} H_i x(t_i) + \int_0^1 x^{(3)}(t)\ell_1(t)dt, \quad N \geqslant 2$$

and

$$\int_0^1 w(t)x(t)dt = \sum_{i=0}^{N} J_i x(t_i)+Jx(t')+\int_0^1 x^{(3)}(t)\ell_2(t)dt$$

are two best quadrature formulae of order three then in general

$$\int_0^1 [\ell_2(t)]^2 dt < \int_0^1 [\ell_1(t)]^2 dt.$$

Proof

We shall show that if $t_0,t_1,\ldots,t_N$ are given then we can choose t' so that the weight J associated with t' is non zero. Suppose otherwise, then for any t' in $0\leq t\leq 1$, where $t'\neq t_i$, $i = 0,1,\ldots,N$ we would have $J=0$. Consequently the two quadrature formulae would be identical and moreover would integrate any natural quintic spline with the knots $t_0,t_1,\ldots,t_N,t'$ exactly.

Let $y(t) = -\sum_{j=0}^{2} c_j K_5^+(t_j-t)+K_5^+(t'-t)$

where $$\sum_{j=0}^{2} c_j\, t_j^{\,r} = t'^{\,r}, \quad r = 0,1,2. \tag{4.5}$$

Then y is a natural quintic spline with the knots $t_0,t_1,\ldots,t_N,t'$, and so for this spline

$$\int_0^1 w(t)y(t)dt = \sum_{i=0}^{N} H_i y(t_i).$$

Now the weights $H_0,H_1,\ldots,H_N$ are independent of t'. Moreover this last equation can be rearranged as

$$\int_0^1 w(t)K_5^+(t'-t)dt = \sum_{i=0}^{N} H_i y(t_i)+\sum_{j=0}^{2} c_j\int_0^1 w(t)K_5^+(t_j-t)dt. \tag{4.6}$$

We notice that (4.5) can be solved for c_0, c_1, c_2 and it is clear that each will be quadratic in t'. It follows that the right hand side of (4.6) is a polynomial of degree five in t'. However the left hand side can never be such a polynomial, consequently the relation (4.6) cannot hold for a continuum of t'. (However there may be values of t' for which (4.6) can hold, but they will not form an interval.)

Thus we have shown that we can choose t' so that $J \neq 0$. The proof of the theorem is completed as follows.

$$\int_0^1 [\ell_1(t)]^2 dt = \int_0^1 [\ell_2(t)]^2 dt + \int_0^1 [\ell_1(t) - \ell_2(t)]^2 dt + 2\int_0^1 \ell_2(t)[\ell_2(t) - \ell_1(t)]dt.$$

Now

$$\ell_2(t) - \ell_1(t) = \sum_{i=0}^{N} (H_i - J_i) K_2^+ (t_i - t) - JK_2^+ (t' - t),$$

which is easily seen to be the third derivative of a natural quintic spline with the knots $t_0, t_1, \ldots, t_N, t'$. However such a spline is integrated exactly by the quadrature formula with Peano kernel ℓ_2 and so the remainder vanishes for it. Hence

$$\int_0^1 [\ell_1(t)]^2 dt = \int_0^1 [\ell_2(t)]^2 dt + \int_0^1 [\ell_1(t) - \ell_2(t)]^2 dt.$$

It follows that

$$\int_0^1 [\ell_2(t)]^2 dt \leq \int_0^1 [\ell_1(t)]^2 dt$$

with equality if and only if $\ell_1 = \ell_2$, that is to say if and only if $J = 0$.

It remains only to remark that since there exist best quadrature with $J \neq 0$ it follows that for these quadrature formulae

$$\int_0^1 [\ell_2(t)]^2 dt < \int_0^1 [\ell_1(t)]^2 dt.$$

In the proof of theorem 4.6 the restriction $N \geq 2$ was necessary for the construction of the spline. However with the same method of proof we can establish the following theorem.

Theorem 4.7

Let $\displaystyle\int_0^1 w(t)\ x(t)dt = \sum_{i=0}^{1} H_i x(t_i) + \int_0^1 x^{(3)}(t)\ell_1(t)dt,$

and let

$$\int_0^1 w(t)x(t)dt = \sum_{i=0}^{1} J_i x(t_i) + J'x(t') + J''x(t'') + \int_0^1 x^{(3)}(t)\ell_2(t)dt$$

be a best quadrature formula of order three. Then in general

$$\int_0^1 [\ell_2(t)]^2 dt < \int_0^1 [\ell_1(t)]^2 dt.$$

Proof

Clearly, as in the proof of theorem 4.6, we have only to show that we can choose t' and t'' so that $|J'|+|J''| \neq 0$. To do this we take a fixed point t' in $[0,1]$ such that $t' \neq t_0, t_1$. (This point t' takes the place of t_2 in the proof of theorem 4.6 when $N=2$). Then we can show exactly as before that the hypothesis that $J'' = 0$ for all choices of t'' in an interval would lead to a contradiction. Consequently there will exist quadrature formulae for which $J'' \neq 0$. The result follows as before.

CHAPTER 5

Natural Quintic Hermitian Splines

We examine the type of natural spline which is associated with the optimum quadrature formula to be discussed in the next chapter. This spline is uniquely determined when it and its first derivative are given at the knots; by analogy with Hermite's polynomial interpolation formula ~~they~~ it ~~are~~ is called a Hermitian spline~~s~~.

An analysis of the convergence of other types of Hermitian spline have been given, see for example Hall [4], however for later purposes we need to discuss the convergence of the natural interpolating quintic Hermitian spline. For brevity we shall call it an H-spline, and denote it by y. The chapter closes with an examination of the qualitative properties of the two fundamental H-splines.

Existence and construction

We assume that $y_i - x_i = y_i^{(1)} - x_i^{(1)} = 0$, $i=0,1,\ldots,N$ and for simplicity we shall write

$$y_i^{(2)} = a_i, \; y_i^{(3)} = b_i, \quad i=0,1,\ldots,N.$$

Since the spline is natural we have immediately that $b_o = b_N = 0$.

Lemma 5.1

If y is an H-spline such that $y_i = x_i$, $y_i^{(1)} = x_i^{(1)}$, $i=0,1,\ldots,N$ then

$$3a_o - a_1 = \frac{20}{h_o^2}(x_1 - x_o) - \frac{4}{h_o}(3x_o^{(1)} + 2x_1^{(1)})$$

$$-\alpha_j a_{j-1} + 3a_j - (1-\alpha_j)a_{j+1} = 20\,\frac{h_{j-1}h_j}{h_{j-1}+h_j}\left[\frac{x_{j+1}-x_j}{h_j^3} - \frac{x_j - x_{j-1}}{h_{j-1}^3}\right]$$

$$+ 8\,\frac{h_{j-1}\,h_j}{h_{j-1}+h_j}\left[\frac{x_{j-1}^{(1)}}{h_{j-1}^2} - \frac{x_{j+1}^{(1)}}{h_j^2}\right] + 12\left(\frac{1}{h_{j-1}} - \frac{1}{h_j}\right)x_j^{(1)},\ j=1,2,\ldots,N-1, \quad (5.1)$$

$$-a_{N-1} + 3a_N = \frac{-20}{h_{N-1}^2}(x_N - x_{N-1}) + \frac{4}{h_{N-1}}(2x_{N-1}^{(1)} + 3x_N^{(1)}).$$

$$-(1-\alpha_j)b_{j-1} + 3b_j - \alpha_j b_{j+1} = \frac{12}{h_{j-1}+h_j}\left\{10\left[\frac{x_{j+1}-x_j}{h_j^2} + \frac{x_j - x_{j-1}}{h_{j-1}^2}\right] - \right.$$

$$\left. - \left[\frac{3}{h_j}x_{j+1}^{(1)} + 7\left(\frac{1}{h_j} + \frac{1}{h_{j-1}}\right)x_j^{(1)} + \frac{3}{h_{j-1}}x_{j-1}^{(1)}\right]\right\} \quad (5.2)$$

$$j=1,2,\ldots,N-1$$

where $b_o = b_N = 0$, and $\alpha_j = h_j/(h_{j-1}+h_j)$.

Proof

These relations can be proved by the method of undetermined coefficients. Alternatively we have for any $z \in C^6[t_j, t_{j+1}]$,

$$z(t) = A(s)z_i + A(1-s)z_{i+1} + B(s)h_i z_i^{(1)} - B(1-s)h_i z_{i+1}^{(1)} + C(s)h_i^3 z_i^{(3)} - C(1-s)h_i^3 z_{i+1}^{(3)}$$

$$+ \frac{h_i^6}{720}s^2(s-1)^2(s^2-s-\tfrac{1}{2})z^{(6)}(t'),\ t_i < t' < t_{i+1}, \quad (5.3)$$

where $s = (t-t_i)/h_i$ and

$$A(s) = \frac{1}{2}(1-s)^2(2+4s+s^2-2s^3),$$

$$B(s) = \frac{1}{4}(1-s)^2 s(4+s-2s^2),$$

$$C(s) = \frac{1}{48}(1-s)^2 s^2(2s-3).$$

If $z = y$, an H-spline, then the remainder term vanishes. This gives therefore a representation of y in $[t_i, t_{i+1}]$ in terms of $y_i, y_{i+1}, y_1^{(1)}, y_{i+1}^{(1)}, y_i^{(3)}, y_{i+1}^{(3)}$. It remains only to impose the conditions of the continuity of $y^{(2)}$ at the knots to give the relations of (5.2). Those of (5.1) can be proved in a similar fashion, for this the representation of y in terms of $y_i, y_{i+1}, y_i^{(1)}$, $y_{i+1}^{(1)}$, a_i, a_{i+1} is required.

The existence and uniqueness of the interpolating spline can be deduced either from (5.1) or (5.2) since the matrix of each set of equations is strictly diagonally dominant. We note that the uniform norm of the inverse of the matrix of either of the equations is bounded above by 1/2.

Convergence

We shall outline the proof of a theorem on the convergence of the interpolating H-spline. This is similar in principle to that of the corresponding result for natural cubic spline interpolation, the algebra is however more complicated and so will be omitted.

First we prove a result which will be found useful in the next chapter.

Lemma 5.2

$$|a_o - x_o^{(2)}|,\ |a_N - x_N^{(2)}| < \tfrac{1}{3}\max(h_o|x_o^{(3)}|, h_{N-1}|x_N^{(3)}|) + h^4||x^{(6)}||/720$$

where $h = \max_i h_i$.

Proof

We sketch this proof.

In the equations (5.1) we write $e_i = a_i - x_i^{(2)}$, $i=0,1,\ldots,N$ and rearrange them to produce a set of N+1 algebraic equations in $e_o, e_1, \ldots, e_N$. The right hand sides of these equations are functions of x alone and with the aid of Peano's method can be written as

$$\frac{1}{3} h_o x_o^{(3)} + \frac{1}{360} h_o^4 x^{(6)}, \quad \frac{h_{j-1}h_j}{360} \frac{(h_{j-1}^3 + h_j^3)}{h_{j-1}+h_j} x^{(6)}, \quad j=1,2,\ldots,N-1,$$

$$-\frac{1}{3} h_{N-1} x_N^{(3)} + \frac{1}{360} h_{N-1}^4 x^{(6)},$$

where we have omitted the arguments of $x^{(6)}$.

Since the norm of the inverse of the matrix of the equations in $e_o, e_1, \ldots, e_N$ is bounded above by 1/2 we obtain the result stated.

<u>Lemma</u> 5.3

If $x \in C^6[t_o, t_N]$ and y is the H-spline such that

$$y_i - x_i = y_i^{(1)} - x_i^{(1)} = 0, \quad i=0,1,\ldots,N$$

then, for $1 \leq i \leq N-1$,

$$|b_i - x_i^{(3)}| < \frac{9}{8} [3^{-i}|x_o^{(3)}| + 3^{i-N}|x_N^{(3)}|] + \frac{h^3}{60} ||x^{(6)}||,$$

where $b_i = y_i^{(3)}$.

<u>Proof</u>

Let $f_i = b_i - x_i^{(3)}$, $i=1,2,\ldots,N-1$. Then (5.2) can be written

$$3f_1 - \alpha_1 f_2 = r_1 - 3x_1^{(3)} + \alpha_1 x_2^{(3)}$$

$$-(1-\alpha_j)f_{j-1} + 3f_j - \alpha_j f_{j+1} = r_j + (1-\alpha_j)x_{j-1}^{(3)} - 3x_j^{(3)} + \alpha_j x_{j+1}^{(3)}, \quad j=2,3,\ldots,N-2,$$

$$-(1-\alpha_{N-1})f_{N-2} + 3f_{N-1} = r_{N-1} + (1-\alpha_{N-1})x_{N-2}^{(3)} - 3x_{N-1}^{(3)}$$

where r_j denotes the right hand side of the j-th equation in (5.2).

Peano's method now leads to the result that

$$
\begin{aligned}
3f_1 - \alpha_1 f_2 &= -(1-\alpha_1)x_o^{(3)} + s_1 \\
-(1-\alpha_j)f_{j-1} + 3f_j - \alpha_j f_{j+1} &= s_j, \quad j=2,3,\ldots,N-2, \\
-(1-\alpha_{N-1})f_{N-2} + 3f_{N-1} &= -\alpha_{N-1}x_N^{(3)} + s_{N-1}
\end{aligned}
\tag{5.4}
$$

where

$$
s_j = \frac{1}{30}\,\frac{[h_j^4 x^{(6)}(t_j'') - h_{j-1}^4 x^{(6)}(t_j')]}{h_j + h_{j-1}}, \quad j=1,2,\ldots,N-1 \tag{5.5}
$$

and $t_{j-1} \leq t_j' \leq t_j \leq t_{j+1}'' \leq t_{j+1}$.

Denote the elements in the inverse of the matrix of the equations (5.4) by C, then, since $||C^{-1}|| \leq \frac{1}{2}$ it follows that

$$
|f_i| \leq |c_{i1}(1-\alpha_1)x_o^{(3)}| + |c_{iN-1}\,\alpha_{N-1}\,x_N^{(3)}| + \frac{1}{60}h^3||x^{(6)}||, \tag{5.6}
$$

$$
i=1,2,\ldots,N-1.
$$

Finally, from (1.4), $|c_{i1}| < 9.3^{-i}/8$, $|c_{iN-1}| < 9.3^{i-N}/8$, which, when used in (5.6) will give the result.

Corollary 1

$$
\max|b_i - x_i^{(3)}| < \tfrac{1}{2}[\,|x_o^{(3)}| + |x_N^{(3)}|\,] + \frac{h^3}{60}||x^{(6)}||, \quad i=1,2,\ldots,N-1.
$$

Corollary 2

For sufficiently large N there exist knots t_p, t_q, where $t_o \leq t_p < t_q \leq t_N$ such that

$$
|b_i - x_i^{(3)}| = O(h^3) \quad p \leq i \leq q.
$$

Moreover $t_p - t_o$, $t_N - t_q$ are $O(h \log h)$ as $h \to 0$, where $\max_i h_i = h$ (We assume t_o, t_N fixed.)

Proof

This is similar to the proof of the corresponding result for natural cubic splines, [9].

Theorem 5.4

If $x \in C^6[t_o, t_N]$, and y is the H-spline such that $y_i - x_i = y_i^{(1)} - x_i^{(1)} = 0$, $i=0,1,\ldots,N$, then for $t_o \leqslant t \leqslant t_N$

$$\max|x(t)-y(t)| \leqslant \frac{h^3}{192} \max(|x_o^{(3)}|, |x_N^{(3)}|) + \frac{7h^6||x^{(6)}||}{46080}$$

where $h = \max_i h_i$.

Proof

In (5.3) set $z = y - x$, then for $t_i \leqslant t \leqslant t_{i+1}$

$$y(t)-x(t) = C(s)(b_i - x_i^{(3)}) - C(1-s)(b_{i+1} - x_{i+1}^{(3)})$$
$$- \frac{h_i^6}{720} s^2(1-s)^2(s^2-s-\tfrac{1}{2})x^{(6)}(t'), \quad t_i < t' < t_{i+1},$$

where $s = (t-t_i)/h_i$ and $C(s) = (1-s)^2 s^2 (2s-3)/48$.

It is easy to show that for $t_i \leqslant t \leqslant t_{i+1}$ the following inequality is true,

$$|y(t)-x(t)| \leqslant \frac{h_i^3}{192} \max(|b_i - x_i^{(3)}|, |b_{i+1} - x_{i+1}^{(3)}|) + h_i^6||x^{(6)}||/15360.$$

Corollary 1 of the previous theorem shows that in $t_o \leqslant t \leqslant t_N$

$$|y(t)-x(t)| \leq \frac{h^3}{192}\max(|x_o^{(3)}|, |x_N^{(3)}|, \tfrac{1}{2}(|x_o^{(3)}|+|x_N^{(3)}|) + \frac{h^3}{60}||x^{(6)}||) + \frac{h^6||x^{(6)}||}{15360}.$$

The result now follows after some simplification.

Note

Corollary 2 of theorem 5.3 can be used to provide a proof of the existence of knots t_p and t_q such that $\max|x(t)-y(t)| = O(h^6)$ for $t_p \leq t \leq t_q$, where t_p-t_o, t_N-t_q are $O(h \log h)$ as $h \to 0$, c.f. [9].

The following theorem is easily proved, in it we assume t_o and t_N fixed.

Theorem 5.5

Given $\varepsilon > 0$, for any $x \in C[t_o, t_N]$ there exists for sufficiently small h a natural quintic Hermitian spline s such that $||x-s|| < \varepsilon$.

Oscillation Properties of the Fundamental H splines

We end this chapter with a discussion of the qualitative properties of the two basic types of H-splines, namely the fundamental H-splines. The results are of interest in themselves and will be invaluable when we discuss the qualitative properties of optimum quadrature.

Definition

(a) P_i is a fundamental H-spline of the first kind if

$$P_i(t_j)=\delta_{ij},\ P_i^{(1)}(t_j)=0, \quad j=0,1,\ldots,N.$$

(b) Q_i is a fundamental H-spline of the second kind if

$$Q_i(t_j)=0,\ Q_i^{(1)}(t_j)=\delta_{ij}, \quad j=0,1,\ldots,N.$$

An immediate consequence of these definitions is that any H-spline can be written as

$$\sum_{i=0}^{N}(y_iP_i + y_i^{(1)}Q_i).$$

Lemma 5.6

$$|P_i^{(2)}(t_0)| + |P_i^{(2)}(t_N)| \neq 0, \quad i=0,1,\ldots,N.$$

Proof

The result is obvious for N=1, since otherwise P_o would be a quintic in $t_o \leq t \leq t_1$ such that $P_o(t_o) = 1$, $P_o^{(r)}(t_o)=0$, $r=1,2,3$, $P_o^{(r)}(t_1)=0$, $r=0,1,2,3$ which is clearly impossible.

We shall now show that for $N \geq 2$ the hypothesis that $P_i^{(2)}(t_0) = P_i^{(2)}(t_N) = 0$ leads to a contradiction. The cases i=0 or N will be considered first, and to be definite we take $i=0$. It follows that $P_o^{(r)}(t_N) = 0$, $r=0,1,2,3$ and $P_o^{(r)}(t_{N-1}) = 0$, $r=0,1$. Consequently P_o is identically zero in $[t_{N-1}, t_N]$ -since it is a polynomial of degree at most five there. This argument can be repeated for the interval $[t_{N-2}, t_{N-1}]$ if $N \neq 2$ etc. until the conclusion is reached that P_o is identically zero for $t \geq t_1$. Hence P_o must satisfy the following $P_o(t_o) = 1$, $P_o^{(r)}(t_o) = 0$, $r=1,2,3$, $P_o^{(r)}(t_1) = 0$, $r=0,1,2,3$. Since these cannot be satisfied by any quintic polynomial the hypothesis that $P_o^{(2)}(t_o) = P_o^{(2)}(t_N)=0$ is false.

We now consider a general value of i for $1 \leq i \leq N-1$, with the hypothesis that $P_i^{(2)}(t_o) = P_i^{(2)}(t_N) = 0$ and we will show that P_i must vanish identically outside the interval $[t_{i-1}, t_{i+1}]$.

If $i \geq 2$ the previous reasoning shows that P_i vanishes identically in $[t_o, t_1]$, and, if $i \geq 3$ it will also be identically zero in $[t_1, t_2]$. We repeat this argument in each interval for which i is not a right hand end point to prove that P_i vanishes for $t \leq t_{i-1}$. In a similar fashion we argue from the knot t_N to deduce that P_i vanishes in any interval so long as t_i is not the left hand end point. Consequently P_i is identically zero except in $[t_{i-1}, t_{i+1}]$. Thus it satisfies

$$P_i^{(r)}(t_{i-1}) = P_i^{(r)}(t_{i+1}) = 0, \quad r=0,1,2,3, \; P_i(t_i)=1, P_i^{(1)}(t_i)=0. \qquad (5.9)$$

We shall show that these cannot be satisfied by an H-spline. For since $P_i^{(1)}$ vanishes at t_{i-1}, t_i, t_{i+1} it follows from Rolle's theorem that $P_i^{(2)}$ vanishes at least twice in (t_{i-1}, t_{i+1}), consequently $P_i^{(2)}$ vanishes at least four times in $[t_{i-1}, t_{i+1}]$.

Another use of Rolle's theorem and of the conditions that $P_i^{(3)}(t_{i-1}) = P_i^{(3)}(t_{i+1}) = 0$ leads to the conclusion that $P_i^{(3)}$ vanishes at least five times in $[t_{i-1}, t_{i+1}]$. Hence since $P_i^{(3)}$ is at most quadratic we see that $P_i^{(3)}$ is identically zero in one of the intervals $[t_{i-1}, t_i]$ or $[t_i, t_{i+1}]$. Let it be the former, then $P_i^{(r)}(t_{i-1}) = 0$ for $r=0,1,2$ implies that $P_i(t) = 0$ for $t_{i-1} \leq t \leq t_i$. However this contradicts $P_i(t_i) = 1$, and so no H-spline satisfies (5.9).

Thus we have shown that the hypothesis that $|P_i^{(2)}(t_o)|+|P_i^{(2)}(t_N)|=0$ leads to contradictions, and so it is false.

Corollary

$$P_o^{(2)}(t_o) \neq 0,\ P_o^{(2)}(t_N) \neq 0,\ P_N^{(2)}(t_o) \neq 0,\ P_N^{(2)}(t_N) \neq 0.$$

Proof

Assume that $P_o^{(2)}(t_o) = 0$, we shall obtain a contradiction. Now P_o vanishes at least N times in $(t_o, t_N]$, consequently $P_o^{(1)}$ vanishes at least 2N times in $[t_o, t_N]$. Hence $P_o^{(2)}$ vanishes not less than 2N times in $[t_o, t_N]$, and so $P_o^{(3)}$ vanishes at least 2N+1 times in $[t_o, t_N]$. However this means that P_o is quadratic in one of the intervals. Clearly this cannot be $[t_o, t_1]$ because the conditions $P_o(t_o) = 1$, $P_o^{(1)}(t_o) = P_o^{(2)}(t_o) = P_o^{(1)}(t_N) = 0$ cannot be satisfied by a quadratic.

Moreover if the interval in which $P_o^{(3)}$ vanishes is $[t_k, t_{k+1}]$, $k \geq 1$, then this would imply that P_o would be identically zero in that interval. In this case P_o would be identically zero for $t \geq t_k$ which leads to the conclusion that $P_o^{(2)}(t_N) = 0$. This contradicts lemma 4.1, and so $P_o^{(2)}(t_o) \neq 0$.

The proofs of the other results follow in a similar fashion.

Theorem 5.7

If $P_i^{(2)}(t_N) = 0$ then

(i) $P_i(t) = 0$ for $t \geq t_{i+1}$,

(ii) $P_i(t) \geq 0$, $-\infty < t < \infty$,

(iii) $P_i(t) = 0$ only for $t = t_o, t_1, \ldots, t_{i-1}, t_{i+1}$ in $-\infty < t \leq t_{i+1}$.

There is a similar set of conclusions when $P_i^{(2)}(t_o) = 0$.

Proof

From the corollary to lemma 4.1 if $P_i^{(2)}(t_N) = 0$ then $i \leq N-1$.

(i) This follows as in the proof of lemma 5.6. We note that $i \geq 1$ and $P_i^{(2)}(t_o) \neq 0$, and further that P_i is an H-spline with knots $t_o, t_1, \ldots, t_{i+1}$ and such that $P_i^{(r)}(t_{i+1}) = 0$, $r=0,1,2,3$.

(ii) We shall show first that the only zeros of P_i in $[t_o, t_{i+1}]$ are at the knots $t_o, t_1, \ldots, t_{i-1}, t_{i+1}$. Assume the contrary, then there is another point in (t_o, t_{i+1}) at which P_i vanishes. Consequently P_i has at least $i+2$ zeros in $[t_o, t_{i+1}]$, and so, by Rolle's theorem, $P_i^{(1)}$ must have at least $i+1$ zeros in (t_o, t_{i+1}). One of these zeros will be at the knot t_i, however the remaining ones lie between the knots. Hence $P_i^{(1)}$ vanishes at least $2i+2$ times in $[t_o, t_{i+1}]$. Since $P_i^{(2)}(t_{i+1}) = 0$ we can deduce with the aid of Rolle's theorem that $P_i^{(2)}$ vanishes at least $2i+2$ times in (t_o, t_{i+1}). Another application of Rolle's theorem, together with the use of the conditions $P_i^{(3)}(t_o) = P_i^{(3)}(t_N) = 0$ leads to the result that $P_i^{(3)}$ vanishes at least 2i+3 times in $[t_o, t_{i+1}]$. However $P_i^{(3)}$ is quadratic in each of the i+1 intervals and so must vanish identically in at least one of them. This interval cannot be either $[t_{i-1}, t_i]$ or $[t_i, t_{i+1}]$ since we must have $P_i(t_i) = 1$. Hence P_i vanishes identically in the same interval as $P_i^{(3)}$. We can now continue the property of vanishing in an interval down to the end interval $[t_o, t_1]$. However this would mean that $P_i^{(2)}(t_o) = 0$, which contradicts lemma 4.1. It follows that the only zeros of P_i in $[t_o, t_{i+1}]$ are at $t_o, t_1, \ldots, t_{i-1}, t_{i+1}$.

We next show that $P_i^{(2)}$ does not vanish at $t_o, t_1, \ldots, t_{i-1}$. Again we assume the contrary, that is, $P_i^{(2)}$ vanishes at a knot which is not t_i nor t_{i+1}. Rolle's theorem can now be used to show that since $P_i^{(1)}$ vanishes at least 2i+1 times in $[t_o, t_{i+1}]$ then $P_i^{(2)}$ will vanish at least 2i times in (t_o, t_{i+1}). This will be at points all of which, except perhaps one, are not knots. Consequently from this assumption it would follow that $P_i^{(2)}$ would vanish at least 2i+2 times in $[t_o, t_{i+1}]$. But then we would have that $P_i^{(3)}$ vanished at least 2i+3 times in $[t_o, t_{i+1}]$. The rest of the proof follows on the same line as before.

Thus P_i cannot change sign at the knots, and since $P_i(t_i) = 1$ we conclude that $P_i(t) \geq 0$, $-\infty < t \leq t_{i+1}$ if $i \geq 1$.

It remains to consider the case $i = 0$, that is we have to prove that if $P_o(t_o) = 1$, $P_o^{(1)}(t_o) = 0$, $P_o(t_1) = P_o^{(1)}(t_1) = P_o^{(2)}(t_2) = P_o^{(3)}(t_3) = 0$ then $P(t) > 0$ for $t \leq t_o$. However it is easy to show that $P_o(t) = \frac{(t_1-t)^4}{h_o^4}[1 - 4\frac{(t-t_o)}{h_o}]$ for $t_o \leq t \leq t_1$.

Hence $P_o^{(2)}(t_o) = 140/h_o^2$, and so $P_o(t) = 1 + 70(t-t_o)^2/h_o^2$, $t \leq t_o$, which is positive.

Theorem 5.8

If $P_i^{(2)}(t_o) \neq 0 \neq P_i^{(2)}(t_N)$ then P_i has at most one zero which is not at a knot, and P_i will change sign only at this zero.

Proof

We note first that P_i cannot vanish identically in any interval in $[t_o, t_N]$. This is obviously true for the intervals $[t_{i-1}, t_i]$ and $[t_i, t_{i+1}]$. And if P_i were to vanish identically in any other interval then we would continue the identically zero conditions down to one of the end intervals which would contradict the hypothesis of the theorem.

Let us now assume that P_i vanishes at two distinct points in (t_o, t_N) neither of which are knots. Then, with the aid of Rolle's theorem, we can deduce that $P_i^{(3)}$ must be identically zero in an interval. If this interval does not have t_i as an end point this would imply that P_i would vanish identically in this interval. This however is not possible by the above. Consequently if P_i vanishes at two distinct non knot points then $P_i^{(3)}$ must vanish in one or both of the intervals which have t_i as an end point. Let the interval be $[t_i, t_{i+1}]$. Then since

$$P_i(t_{i+1}) = P_i^{(1)}(t_{i+1}) = P_i^{(3)}(t_{i+1}) = 0$$

it follows from the uniqueness of the interpolatory natural quintic spline that P_i vanishes identically for $t_{i+1} \leq t \leq t_N$. But since

$P_i^{(2)}$ is continuous this implies that $P_i^{(2)}(t_{i+1}) = 0$. Consequently P_i is a quadratic in $[t_i, t_{i+1}]$ and such that $P_i^{(r)}(t_{i+1}) = 0$, $r=0,1,2$. But this implies that $P_i(t_i) = 0$, which is a contradiction.

In a similar fashion if $P_i^{(3)}$ vanishes identically in $[t_{i-1}, t_i]$ we reach the same contradiction. Therefore we conclude that P_i can vanish at no more than one point in (t_o, t_N) in addition to the knots at which it is constrained to vanish.

If $1 \leqslant i \leqslant N-1$ then P_i cannot vanish outside $[t_o, t_N]$ since, for $t \leqslant t_o$ P_i is proportional to $(t-t_o)^2$. When $i=0, P_o$ may vanish outside $[t_o, t_N]$, however it cannot vanish in $[t_o, t_N]$ except at the knots $t_1, t_2, \ldots, t_N$. This follows from the observation that in the count of the zeros of $P_o^{(1)}$ we do not need to allow for the possibility that the knot t_o is counted twice. Since $P_o(t_o) = 1$, $P_o^{(1)}(t_o) = 0$ it follows that $P_o(t) = 1 + (t-t_o)^2 P_o^{(2)}(t_o)/2$ for $t \leqslant t_o$. Clearly this can have at most one zero in $(-\infty, t_o)$. Similarly P_N has at most one zero which is not a knot and this zero lies above t_N.

We complete the proof of the theorem by showing that P_i does not change sign at the knots at which it vanishes. For this to be true we have to prove that $P_i^{(2)}$ does not vanish at any knot except perhaps at t_i. The cases $i=0, N$ need to be treated separately from the general case. Let $i=0$, then P_o has N zeros. Consequently $P_o^{(1)}$ has at least $N-1$ zeros in (t_o, t_N), and these are at points which are not knots. Therefore $P_o^{(1)}$ has at least 2N zeros in $[t_o, t_N]$. It follows that $P_o^{(2)}$ has at least $2N-1$ zeros which lie between the knots, and so has at least 2N zeros in $(t_o, t_N]$ if $P_o^{(2)}$ vanishes at a knot. We now deduce that $P_o^{(3)}$ has at least $2N+1$

zeros in $[t_o, t_N]$. But, as before, this implies that P_o is quadratic in at least one interval. It is easily shown that this will contradict the hypothesis that $P_o(t_o) = 1$. The proof for $i=N$ is similar.

When $1 \leq i \leq N-1$ we begin the count of zeros with the fact that P_i vanishes $N+1$ times in $[t_o, t_N]$. The rest of the proof by contradiction follows familiar lines.

Since P_i can change sign only where it vanishes and since $P_i^{(2)}$ is non zero at each of the knots where it vanishes we conclude that it can change sign only at the zero which is not at a knot. Indeed it is easily proved by the same method that if P_i has a zero at a point which is not a knot then it will certainly change sign there. This concludes the proof.

The theorem shows that P_i has the same qualitative behaviour as the first fundamental polynomial in the usual Hermite's interpolation formula. We next investigate Q_i and we shall see that its behaviour is similar to the second fundamental polynomial in Hermite's interpolation formula.

<u>Lemma</u> 5.9

$Q_i^{(3)}$ cannot be identically zero in any interval in $[t_o, t_N]$.

<u>Proof</u>

The case $N=1$ will be taken first. Now $Q_o^{(3)}(t) = 0$, $t_o \leq t \leq t_1$, would imply that Q_i is at most quadratic in this interval. But this, taken together with $Q_o(t_o) = Q_o(t_1) = Q_o^{(1)}(t_1) = 0$, would give the result that $Q_o(t) = 0$, $t_o \leq t \leq t_1$. Hence since

$Q_o^{(1)}(t_o) = 1$ it follows that $Q_o^{(3)}$ cannot vanish identically in $[t_o, t_1]$. A similar argument holds for Q_1.

When $N \geq 2$ we must distinguish between $i=0,N$ and $1 \leq i \leq N-1$. First let $1 \leq i \leq N-1$, and assume that $Q_i^{(3)}(t) = 0$ for $t_k \leq t \leq t_{k+1}$. We note immediately that $i \neq k,k+1$; if $i=k$ for example then Q_i would be quadratic in $[t_i, t_{i+1}]$ and $Q_i(t_i) = Q_i(t_{i+1}) = Q_i^{(1)}(t_{i+1}) = 0$ which would imply that $Q_i(t) = 0$, $t_i \leq t \leq t_{i+1}$, a contradiction. In addition we also notice that since $Q_i^{(3)}(t_k) = Q_i^{(3)}(t_{k+1}) = 0$ the spline Q_i can be regarded as the union of two splines, one with knots $t_o, t_1, \ldots, t_k$, the other with knots $t_{k+1}, \ldots, t_N$. One of these splines is identically zero (since a spline is uniquely determined by its interpolation conditions). Let this be the spline with the knots $t_{k+1}, t_{k+2}, \ldots, t_N$, that is, $i < k$. Now if $i \neq k-1$ we see that $Q_i(t) = 0$ for $t_{k-1} \leq t \leq t_k$. This follows from the requirements that Q_i should be a polynomial of degree at most five in $[t_{k-1}, t_k]$ together with $Q_i(t_{k-1}) = Q_i^{(1)}(t_{k-1}) = 0$, $Q_i^{(r)}(t_k) = 0$, $r=0,1,2,3$. Clearly this argument can be repeated until we reach the conclusion that $Q_i(t) = 0$ for $t \geq t_{i+1}$.

Now Q_i vanishes at $i+2$ knots, $t_o, t_1, \ldots, t_{i+1}$, and so by Rolle's theorem $Q_i^{(1)}$ vanishes at least $i+1$ times between knots. Hence $Q_i^{(1)}$ vanishes at least $2i+1$ times in $[t_o, t_{i+1}]$. Hence, again by Rolle's theorem, $Q_i^{(2)}$ vanishes at not less than $2i+2$ points in $[t_o, t_{i+1}]$ (we have used $Q_i^{(2)}(t_{i+1}) = 0$). Finally, Rolle's theorem together with the constraints $Q_i^{(3)}(t_o) = Q_i^{(3)}(t_{i+1}) = 0$ gives the result that $Q_i^{(3)}$ has not less than $2i+3$ zeros in $[t_o, t_{i+1}]$. However $Q_i^{(3)}$ is quadratic in each of the $i+1$ intervals and so can vanish at no more than $2i+2$ points unless it vanishes identically in one of the intervals. Consequently $Q_i^{(3)} = 0$ in at least one of the intervals in $[t_o, t_{i+1}]$.

Clearly we can now use the previous argument to conclude that Q_i vanishes identically except for $t_{i-1} \leq t \leq t_{i+1}$. However we have the following conditions to be satisfied,

$$Q_i^{(r)}(t_{i-1}) = Q_i^{(r)}(t_{i+1}) = 0, \ r=0,1,2,3, \ Q_i(t_i)=0, \ Q_i^{(1)}(t_i)=1.$$

Rolle's theorem can be used again to show that these conditions cannot hold simultaneously. Indeed Q_i vanishes at $t = t_{i-1}, t_i, t_{i+1}$ implies that $Q_i^{(1)}$ vanishes at least four times in $[t_{i-1}, t_{i+1}]$. Consequently $Q_i^{(2)}$ vanishes at least five times in the same range and $Q_i^{(3)}$ vanishes at least six times also in this range. We conclude that Q_i is quadratic in one of the intervals $[t_{i-1}, t_i]$ or $[t_i, t_{i+1}]$. Clearly this precludes $Q_i^{(1)}(t_i) = 1$. Hence the hypotehsis that $Q_i^{(3)}$ vanishes identically in an interval leads to a contradiction.

It remains to prove the result for $i=0,N$. However it is easily seen that Q_o, for example, would be identically zero except in $[t_o, t_1]$. This case is completed by noticing that this situation is identical with the one treated in the first paragraph,(when $N = 1$).

<u>Theorem</u> 5.10

(a) $Q_i^{(2)}(t_j) \neq 0, \ j \neq i, \ 0 \leq i,j \leq N,$

(b) $Q_o^{(2)}(t_o) \neq 0 \neq Q_N^{(2)}(t_N),$

(c) $(t-t_i)Q_i(t) \geq 0, \ -\infty < t < \infty.$

<u>Proof</u>

We shall show first that Q_i vanishes in $[t_o, t_N]$ only at the knots. Suppose the contrary, that is, Q_i vanishes in (t_o, t_N) at some point which is not a knot. Then this would mean that Q_i would vanish at least N+2 times in $[t_o, t_N]$. Three applications of Rolle's theorem will lead to the conclusion that $Q_i^{(3)}$ would have to vanish identically in an interval. This is not possible by lemma 5.9, and so the supposition is false.

We show next that $Q_i^{(2)}(t_j) \neq 0$, $j \neq i$. Two applications of Rolle's theorem will show that $Q_i^{(2)}$ vanishes at least 2N times in $[t_o, t_N]$ at points all of which except one (which may be $t=t_i$) lie between knots. Suppose now $Q_i^{(2)}(t_j) = 0$ for some knot t_j, where $i \neq j$. Then $Q_i^{(2)}$ would vanish at not less than 2N+1 points in $[t_o, t_N]$. Rolle's theorem and the conditions $Q_i^{(3)}(t_o) = Q_i^{(3)}(t_N)=0$ lead to the conclusion that in this case we would have $Q_i^{(3)} = 0$ in an interval. This contradicts lemma 6.1 and so the hypothesis that $Q_i^{(2)}(t_j) = 0$ for some $t_j \neq t_i$ is false.

If $i = 0, N$ we can obtain the stronger results stated in (b). The proof follows the same path as in the previous paragraph, however the reservation that the 2N points at which $Q_o^{(2)}$ can be shown to vanish might include the knot t_o need no longer be made. Consequently the hypothesis that $Q_o^{(2)}(t_o) = 0$ will lead to a contradiction.

If i satisfies $1 \leq i \leq N-1$ the proof of (c) is now straightforward, for Q_i vanishes only at the knots $t_o, t_1, \ldots, t_N$, and so, since $Q_i^{(2)}$ does not vanish at the knots $t_o, t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_N$, it follows that Q_i changes sign in $[t_o, t_N]$ only when $t = t_i$. When $t \leq t_o$ we have $Q_i(t) = (t-t_o)^2 Q_i^{(2)}(t_o)/2$, and so Q_i does not vanish for $t < t_o$. Similarly for $t \geq t_N$.

We shall complete the proof by showing that Q_o does not vanish for $t < t_o$; since $Q_o(t) = (t-t_o) + (t-t_o)^2 Q_o^{(2)}(t_o)/2$

this can be accomplished by proving that $Q_o^{(2)}(t_o) < 0$. Suppose the contrary is true, that is that $Q_o^{(2)}(t_o) > 0$ (we have seen that $Q_o^{(2)}(t_o) \neq 0$). We use Rolle's theorem to show that this would imply that there would be a point t', where $t_o < t' < t_1$, at which $Q_o^{(1)}(t') = 0$. Consequently $\int_{t_o}^{t'} Q_o^{(2)}(t)dt + 1 = 0$, and since $Q_o^{(2)}(t_o) > 0$ it follows that $Q_o^{(2)}$ would vanish at some point in (t_o, t'). We can use the existence of this zero of $Q_o^{(2)}$ to lead, again by Rolle's theorem, to the contradictory conclusion that $Q_o^{(3)}$ would vanish in some interval. Hence $Q_o^{(2)}(t_o) < 0$ and Q_o cannot vanish in $(-\infty, t_o)$.

This completes the proof of the theorem.

Theorem 5.11

For each i, $0 \leqslant i \leqslant N$, there exists a constant α_i such that

$$P_i(t) + \alpha_i Q_i(t) \geqslant 0, \quad -\infty < t < \infty.$$

Proof

Let $R_i = P_i + \alpha_i Q_i$, we choose α_i so that $R_i^{(2)}(t)$ vanishes at one of the end knots. If $P_i^{(2)}$ vanishes at one of the end knots then we take $\alpha_i = 0$, and theorem ~~6.2~~ 5.7 gives the required result.

In the general case when $P_i^{(2)}(t_o) \neq 0 \neq P_i^{(2)}(t_N)$ choose α_i so that $R_i^{(2)}(t_N) = 0$.* The proof that $R_i(t) \geqslant 0$ for $-\infty < t < \infty$ is accomplished in two steps. First it is clear that R_i is identically zero for $t \geqslant t_{i+1}$ (if i = N then this is not relevant), the hypothesis that R_i vanishes in $[t_o, t_{i+1}]$ at points which are not knots is easily shown to be false by the use of Rolle's theorem.

* This is possible since $Q_i^{(2)}(t_o) \neq 0 \neq Q_i^{(2)}(t_N)$, theorem 5.10.

Rolle's theorem can be used again to prove that $R_i^{(2)}$ does not vanish at any of the knots $t_o, t_1, \ldots, t_{i-1}$. The proof is completed by showing, by calculation, that R_i cannot vanish for $t < t_o$. Then, since $R_i(t_i) = 1$, it will follow that $R_i(t) \geq 0$ in $(-\infty, \infty)$.

CHAPTER 6

Third Order Optimum Quadrature

We turn now to the investigation of optimum quadrature of order three. The quadrature formula with remainder will be written

$$\int_0^1 w(t)x(t)dt = \sum_{i=o}^{N} H_i x(t_i) + \int_a^b k(t)x^{(3)}dt$$

where $a = \min(0, t_o)$, $b = \max(1, t_N)$ and

$$k(t) = \int_0^1 w(s)K_2^+ (s-t)ds - \sum_{i=o}^{N} H_i K_2^+(t_i-t), \quad a \leqslant t \leqslant b.$$

It will be recalled that we wish to find quadrature points $t_o, t_1, \ldots, t_N$ and weights $H_o, H_1, \ldots, H_N$ so as to minimize

$$\int_a^b [k(t)]^2 dt$$

subject to $\displaystyle\int_0^1 w(t)t^r dt = \sum_{i=o}^{N} H_i t_i^r$, $r=0,1,2$.

We shall say that an optimum formula with N+1 points exists if there is a set of <u>distinct</u> quadrature points $t_o, t_1, \ldots, t_N$ and <u>nonzero</u> quadrature weights $H_o, H_1, \ldots, H_N$ which minimize $\int_a^b [k(t)]^2 dt$. It has been shown by Karlin [6] that an optimum quadrature formula exists for $N \geqslant 2$ when w=1 (see also Schoenberg [15]). However it is not clear from his proof (which is sketched) that it can be adapted for the case when w is not constant. We present here a proof in this case.

<u>Theorem</u> 6.1

If, for $N \geqslant 1$, an optimum formula with N+1 points exists and if $w(t) \geq 0$, $0 \leq t \leq 1$, then

(i) $0 < t_i < 1$, $i=0,1,\ldots,N$,

(ii) $H_i > 0$, $i=0,1,\ldots,N$.

Proof

If P_i and Q_i, $i=0,1,\ldots,N$ are the fundamental H-splines and if y is any H-spline then

$$y = \sum_{i=0}^{N} y_i P_i + y_i^{(1)} Q_i. \tag{6.1}$$

(i) Since the optimum formula integrates exactly any H-spline with the knots $t_0,t_1,\ldots,t_N$ it follows that $\int_0^1 w(t)Q_i(t)dt = 0$, $i=0,1,\ldots,N$. Consequently Q_i must change sign in $(0,1)$ since w is of constant sign there. But, from theorem 5.10(c), Q_i changes sign only at the knot t_i. Hence t_i must satisfy $0 < t_i < 1$.

(ii) The quadrature weights $H_0,H_1,\ldots,H_N$ are given by

$$H_i = \int_0^1 w(t)P_i(t)dt = \int_0^1 w(t)[P_i(t)+\alpha_i Q_i(t)]dt, \quad i=0,1,\ldots,N.$$

From theorem 5.11 we can choose α_i so that $P_i+\alpha_i Q_i \geq 0$ in $(-\infty,\infty)$. Hence, as $w(t) \geq 0$, the result follows.

Corollary

If an optimum formula exists with N+1 points then it minimizes $\int_0^1 [k(t)]^2 dt$, (i.e. $a=0,b=1$), subject to the constraints.

Lemma 6.2

If, for $N \geq 2$, an optimum formula with N+1 points exists then an optimum formula with N+2 points will exist.

Proof

From the previous corollary it is clear that we need consider only quadrature points in $(0,1)$.

Let k be the Peano kernel of the N+1 point formula, if there is more than one such formula we choose the one which gives the smallest value of $\int_0^1 [k(t)]^2 dt$. Let k_1 be the Peano kernel of any N+2 point quadrature formula (for the same integral) which is exact for quadratics and whose quadrature points lie in $[0,1]$.

Suppose that an N+2 point optimum formula does not exist. Then the problem of minimizing $\int_0^1 [k_1(t)]^2 dt$ subject to the constraints would not have a solution for which there would be N+2 distinct quadrature points. However by supposition an N+1 point formula exists and so the N+2 point minimization problem would have the N+1 point formula as its solution. Hence, for all choices of N+2 quadrature points we would have $\int_0^1 [k_1(t)]^2 dt \geqslant \int_0^1 [k(t)]^2 dt$. But we have seen, theorem 4.6, that there exist N+2 point formulae such that $\int_0^1 [k_1(t)]^2 dt < \int_0^1 [k(t)]^2 dt$. It follows that an N+2 point formula will exist. The restriction $N \geqslant 2$ is necessary in this proof; however we have the following result when N=1.

Lemma 6.3

If an optimum formula with two points exists then there will exist an optimum formula with four points.

Proof

The proof of this uses the same arguments as were used in the proof of lemma 6.2 except that we appeal to theorem 4.7 instead of theorem 4.6.

Theorem 6.4

If $w(t) \geq 0$, $0 \leq t \leq 1$, then optimum quadrature formulae exist with 2, 4, 5, 6, ... points.

Proof

We show first that a two point formula exists, that is to say, the quadrature points are distinct and the weights are non zero. The quadrature formula is exact for quadratics and so

$$\int_0^1 w(t)t^r dt = H_o t_o^r + H_1 t_1^r, \quad r=0,1,2.$$

Clearly neither weight is zero, otherwise these equations would not be satisfied. Moreover the elimination of H_o and H_1 gives $\int_0^1 w(t)(t-t_o)(t-t_1)dt = 0$, and so if $w \geq 0$ the quadrature points cannot coalesce. Hence a two point formula exists.

If a three point formula exists then we can use lemma 6.2 to prove the existence of a four point formula etc. On the other hand if a three point formula does not exist we can now use lemma 6.3 to prove the existence of a four point formula. Hence optimum formulae with 4, 5, ... points exist. This concludes the proof.

It is unfortunate that we are unable to prove the existence of a three point formula for the optimum quadrature of $\int_0^1 w(t)x(t)dt$. Fortunately Karlin's result will fill this gap when $w=1$, moreover a calculation in this case shows that the optimum points are given by $\frac{1}{2}, \frac{1}{2}(1 \pm t)$ where t is that zero of $2t^3 - 9t^2 + 15t - 7$ which lies in $[0,1]$. In fact $2t = 3 + [(17^{\frac{1}{2}}-4)^{\frac{1}{3}} - (17^{\frac{1}{2}}+4)^{\frac{1}{3}}]$, and the quadrature points are

0.128 186, 0.5, 0.871 814 and the quadrature weights $H_0 = H_2 = 0.301\ 396$, $H_1 = 0.397\ 208$.

Convergence of optimum quadrature

If k is the Peano kernel for the optimum formula with N+1 points it follows that for any other quadrature formula with N+1 points in [0,1] we must have

$$\int_0^1 [k(t)]^2 dt < \int_0^1 [\ell(t)]^2 dt$$

where ℓ is the Peano kernel for the other quadrature formula. We can use theorem 4.4to give the following result immediately. (This theorem was suggested by S. Michaelson.)

Theorem 6.5

If $H = 1/N$ and if k is the Peano kernel for the optimum formula then

$$\int_0^1 [k(t)]^2 dt < H^6 ||w||. \int_0^1 |w(t)| dt/8.$$

The order of the convergence of the optimum formula will clearly depend on the distribution of quadrature points, in particular the size of t_0 and of $1-t_N$ will be crucial. The next lemma goes some way to providing this information but it is clearly unsatisfactory.

Lemma 6.6

If $t_0, t_1, \ldots, t_N$ are the quadrature points in the N+1 point optimum formula with $w(t) = 1$, $0 \leq t \leq 1$ then

$$t_o,\ (1-t_N) < 1.75H^{6/7},\ t_{i+1}-t_i < 4.25H^{6/7},\ i=0,1,\ldots,N-1$$

where $H = 1/N$.

Proof

The Peano kernel k of the quadrature formula is given by

$$k(t) = \frac{(1-t)^3}{6} - \sum_{i=o}^{N} H_i K_2^+(t_i-t),\quad 0 \leq t \leq 1.$$

It follows that $k(t) = \frac{(1-t)^3}{6}$ for $t_N \leq t \leq 1$.

However

$$\int_{t_N}^{1} [k(t)]^2 dt < \int_0^1 [k(t)]^2 dt < H^6/8$$

from theorem 6.3. A simple calculation leads to the result that $(1-t_N)^7 < 63H^6/2$ which gives the required upper bound. It is easily seen that $k(t) = \frac{-t^3}{6}$ for $0 \leq t \leq t_o$, and the required inequality will follow from an argument similar to the previous one.

In order to find an upper bound for $\left(t_{i+1} - t_i\right)$ set

$$x(t) = (t_{i+1}-t)^{3+\alpha}(t-t_i)^{3+\alpha}\ \text{for}\ t_i \leq t \leq t_{i+1},\ \alpha > 0$$

and zero otherwise. Then $x \in C^3[0,1]$ and the quadrature formula gives

$$\int_{t_i}^{t_{i+1}} x(t)dt = \int_{t_i}^{t_{i+1}} k(t)x^{(3)}(t)dt.$$

Hence

$$\left| \int_{t_i}^{t_{i+1}} x(t)dt \right|^2 \leq \int_{t_i}^{t_{i+1}} [x^{(3)}(t)]^2 dt \int_0^1 [k(t)]^2 dt.$$

A simple calculation leads to the result stated.

Theorem 6.7

If $x \in C^6[0,1]$ and if R denotes the remainder in the optimum quadrature formula with N+1 points and with unit weight function then

$$|R(x)| = O(H^{24/7}),$$

where $H = 1/N$.

Proof

Since R vanishes for any H-spline with the knots $t_0, t_1, \ldots, t_N$ it follows that

$$R(x) = R(x-y) = \int_0^1 k(t)[x^{(3)}(t)-y^{(3)}(t)]dt$$

where y is the H-spline which agrees with x at the knots. Schwartz's inequality gives

$$|R(x)|^2 \leq \int_0^1 [k(t)]^2 dt \int_0^1 [x^{(3)}(t)-y^{(3)}(t)]^2 dt. \qquad (6.2)$$

But $y^{(3)}(t) = 0$ for $t \leq t_0$, $t \geq t_N$, and so with the result of theorem 6.5 we deduce that

$$|R(x)|^2 \leq \frac{H^6}{8}\left\{\int_0^{t_0} [x^{(3)}(t)]^2 dt + \int_{t_0}^{t_N} [x^{(3)}(t)-y^{(3)}(t)]^2 dt \right.$$

$$\left. + \int_{t_N}^1 [x^{(3)}(t)]^2 dt \right\} \qquad (6.3)$$

where $H = 1/N$.

Three integrations by parts leads to the result that

$$\int_{t_0}^{t_N} [x^{(3)}(t)-y^{(3)}(t)]^2 dt = x_N^{(3)}[x_N^{(2)}-y_N^{(2)}]-x_o^{(3)}[x_o^{(2)}-y_o^{(2)}]$$

$$- \int_{t_0}^{t_N} x^{(6)}(t)[x(t)-y(t)]dt$$

$$<|x_N^{(3)}|.|x_N^{(2)}-y_N^{(2)}|+|x_o^{(3)}|.|x_o^{(2)}-y_o^{(2)}|+||x^{(6)}||. \max_{t_o\leq t\leq t_N} |x(t)-y(t)|. \quad (6.4)$$

From lemma 5.2

$$|x_o^{(2)}-y_o^{(2)}|,|x_N^{(2)}-y_N^{(2)}| < \tfrac{1}{3} \max\ (h_o|x_o^{(3)}|,h_{N-1}|x_N^{(3)}|)+ \frac{h^4}{720} ||x^{(6)}||$$

and from theorem 5.3

$$\max_{t_o\leq t\leq t_N} |x(t)-y(t)| < \frac{h^3}{192} [\max(|x_o^{(3)}|,|x_N^{(3)}|)+ \frac{7h^3}{46080} ||x^{(6)}||].$$

The use of these in (6.4) leads to

$$\int_{t_0}^{t_N} [x^{(3)}(t)-y^{(3)}(t)]^2 dt < K_1h+K_2h^3+K_3h^4+K_4h^6,$$

where $K_1,\ldots,K_4$ are constants. Consequently, with the aid of lemma 6.6 we obtain the result.

In the case of a general weight function when we do not have any information about the spacing of the quadrature points we have the next result.

<u>Theorem</u> 6.8

If $x \in C^6[0,1]$, and if R denotes the remainder in the optimum quadrature formula with N+1 points then, if the weight function is continuous, $w > 0$,

$$|R(x)| = O(H^3)$$

where $H = 1/N$.

Proof

This is the same as that of the previous theorem except that in this case we can suppose

$$\int_0^1 [x^{(3)}(t)-y^{(3)}(t)]^2 dt$$

is merely bounded. The estimate for $\int_0^1 [k(t)]^2 dt$ is taken from theorem 6.5.

Appendix

Two Cubic Splines and Related Best Quadrature Formulae

We shall consider finally two types of cubic spline which have more favourable convergence properties than the natural cubic spline when they are used to interpolate $x \in C^4[0,1]$.

Definition

A <u>cubic spline</u> with the knots $t_0, t_1, \ldots, t_N$ is in $C^2[0,1]$ and is a polynomial of degree at most three in $[t_j, t_{j+1}]$, $j = 0,1,\ldots,N-1$, $(t_0=0,\ t_N=1)$.

Denote a cubic spline by y and let $y_j = x_j$, $j = 0,1,\ldots,N$, where $x \in C^4[0,1]$. Then if $y_j^{(1)} = \lambda_j$, $j = 0,1,\ldots,N$ we can show from Hermite's interpolation formula that with the uniform norm on $[0,1]$,

$$||x-y|| \leq \frac{1}{4}h \max_j |x_j^{(1)} - \lambda_j| + \frac{1}{384} h^4 ||x^{(4)}||, \text{ where } 0 \leq j \leq N. \tag{A.1}$$

It is clear from this that in general the best order of approximation of the spline y to x which can be expected is $O(h^4)$ and that the precise order depends on $\max_j |x_j^{(1)} - \lambda_j|$.

For the moment let y be a natural cubic spline, and rewrite the equations of (2.4) as

$$\begin{aligned}
2(\lambda_0 - x_0^{(1)}) + (\lambda_1 - x_1^{(1)}) &= 3[t_0,t_1]x - 2x_0^{(1)} - x_1^{(1)} \\
\alpha_j(\lambda_{j-1} - x_{j-1}^{(1)}) + 2(\lambda_j - x_j^{(1)}) + (1-\alpha_j)(\lambda_{j+1} - x_{j+1}^{(1)}) &= 3\alpha_j[t_{j-1},t_j]x + 3(1-\alpha_j)[t_j,t_{j+1}]x \\
&\quad -\alpha_j x_{j-1}^{(1)} - 2x_j^{(1)} - (1-\alpha_j)x_{j+1}^{(1)}, \\
&\qquad j = 1,2,\ldots,N-1, \\
(\lambda_{N-1} - x_{N-1}^{(1)}) + 2(\lambda_N - x_N^{(1)}) &= 3[t_{N-1},t_N]x - x_{N-1}^{(1)} - 2x_N^{(1)},
\end{aligned} \tag{A.2}$$

where $\alpha_j = h_j/(h_{j-1}+h_j)$.

The right hand sides of the first and last of these equations are easily shown to be $O(h)$ (the precise values are not important here). However, with the use of Peano's method, we obtain for the remaining equations,

$$\alpha_j(\lambda_j-x_{j-1}^{(1)})+2(\lambda_j-x_j^{(1)})+(1-\alpha_j)(\lambda_{j+1}-x_{j+1}^{(1)})=\frac{1}{24}\frac{h_{j-1}h_j}{(h_{j-1}+h_j)}[h_{j-1}^2x^{(4)}-h_j^2x^{(4)}] \quad (A.3)$$

$$= O(h^3).$$

A rough estimation of $\max_j|\lambda_j-x_j^{(1)}|$ leads to the conclusion that it is $O(h)$, which, when inserted in (A.1) gives the convergence for the natural cubic spline as $O(h^2)$. Clearly this is due solely to the effect of the $O(h)$ of the first and last equations. The aim is to produce equations to replace these which will give better convergence for the spline.

<u>Third order cubic spline</u> $N \geq 2$

Let $\lambda_0, \lambda_1, y_0, y_1$ be connected by a relation of the form

$$\lambda_0+a\lambda_1 = by_0+cy_1,$$

then, with $y_0 = x_0, y_1 = x_1$, we have

$$(\lambda_0-x_0^{(1)})+a(\lambda_1-x_1^{(1)}) = bx_0+cx_1-x_0^{(1)}-ax_1^{(1)}.$$

Choose now a,b,c so that the order of the right hand side is as high as possible. Some simple analysis shows that if $a=1$, $b=-c=-2/h_0$ then we have

$$(\lambda_0-x_0^{(1)})+(\lambda_1-x_1^{(1)}) = -\frac{1}{6}h_0^2\,x^{(3)}(t'). \quad (A.4a)$$

A similar argument for $\lambda_{N-1},\lambda_N,y_{N-1},y_N$ leads to

$$(\lambda_{N-1}-x_{N-1}^{(1)}) + (\lambda_N-x_N^{(1)}) = -\frac{1}{6}h_{N-1}^2\, x^{(3)}(t''). \qquad (A.4b)$$

Consequently we are lead to a cubic spline for which

$$\lambda_0+\lambda_1 = 2[t_0,t_1]y$$

$$\alpha_j\lambda_{j-1}+2\lambda_j+(1-\alpha_j)\lambda_{j+1}=3\alpha_j[t_{j-1},t_j]y+3(1-\alpha_j)[t_j,t_{j+1}]y,\ j=1,2,\ldots,N-1, \qquad (A.5)$$

$$\lambda_{N-1}+\lambda_N = 2[t_{N-1},t_N]y.$$

The qualitative meaning of the first equation in (A.5) becomes clear after an examination of Hermite's interpolation formula with the interpolation points t_0,t_1. For it is easily seen that it is simply the condition that y should be quadratic in $[t_0,t_1]$. Similarly the last equation implies that y is quadratic in $[t_{N-1},t_N]$. This leads to the following definition.

Definition

A third order cubic spline with the knots $t_0,t_1,\ldots,t_N$ $N\geq 2$ is in $C^2[0,1]$ and such that

(a) it is a polynomial of degree at most three in $[t_j,t_{j+1}]$, $j = 0,1,\ldots,N-1$,

(b) $y^{(3)}(t_0) = y^{(3)}(t_N) = 0$.

Theorem A.1

A third order cubic spline y is determined uniquely by its values at the knots for $N\geq 2$.

Proof

The uniqueness of the spline depends on the solvability of the equations of (A.5). For $N\geq 2$ we can eliminate λ_0,λ_N from these equations to give

$$(2-\alpha_1)\lambda_1+(1-\alpha_1)\lambda_2=\alpha_1[t_0,t_1]y+3(1-\alpha_1)[t_1,t_2]y$$

$$\alpha_j\lambda_{j-1}+2\lambda_j+(1-\alpha_j)\lambda_{j+1}=3\alpha_j[t_{j-1},t_j]y+3(1-\alpha_j)[t_j,t_{j+1}]y,\quad j=2,3,\ldots,N-2, \qquad (A.6)$$

$$\alpha_{N-1}\lambda_{N-2}+(1+\alpha_{N-1})\lambda_{N-1}\quad =3\alpha_{N-1}[t_{N-2},t_{N-1}]y+(1-\alpha_{N-1})[t_{N-1},t_N]y.$$

Since $0<\alpha_j<1$ these equations have a strictly diagonally dominant matrix and so there is a unique solution $\lambda_1,\lambda_2,\ldots,\lambda_{N-1}$. The values of λ_0,λ_N can then be found from (A.5). Hence the spline is determined uniquely by its knot values.

Theorem A.2

If y is a third order cubic spline such that $y_i=x_i$, $i = 0,1,\ldots,N$ where $x \in C^4[0,1]$ then $||y-x|| = O(h^3)$.

Proof

Eliminate $\lambda_0-x_0^{(1)}$ and $\lambda_N-x_N^{(1)}$ from (A.3), (A.4a), (A.4b). This will give a set of linear algebraic equations to be solved for $\lambda_j-x_j^{(1)}$, $j=1,1,N-1$ which has the same matrix as that of (A.6). Clearly the right hand sides of these equations are at least $O(h^2)$, and as the uniform norm of the matrix is bounded by unity it follows that $|\lambda_j-x_j^{(1)}|=O(h^2)$ for $j = 1,2,\ldots,N-1$.

From (A.4a) we see that $|\lambda_0 - x_0^{(1)}| = O(h^2)$, similarly $|\lambda_N - x_N^{(1)}| = O(h^2)$. Consequently $|\lambda_j - x_j^{(1)}| = O(h^2)$, $j = 0,1,\ldots,N$.

It remains to insert these in (A.1) to give the result. The next two theorems will be stated only since the proofs follow familiar lines.

Theorem A.3

The unique $z \in C^2[0,1]$ such that

(a) $z_j = x_j$, $j = 1,2,\ldots,N-1$, $N \geq 2$

(b) $z_0^{(1)} = x_0^{(1)}$, $z_N^{(1)} = x_N^{(1)}$

(c) $\int_0^1 [z^{(2)}(t)]^2 dt$ is a minimum

is a third order cubic spline with the knots $t_0, t_1, \ldots, t_N$.

Theorem A.4

If

$$\int_0^1 w(t)x(t)dt = J_0 x_0^{(1)} + \sum_{i=1}^{N-1} H_i x_i + J_N x_N^{(1)} + \int_0^1 k(t)x^{(2)}(t)dt$$

is a best quadrature formula with remainder - in the sense that $J_0, H_1, \ldots, H_{N-1}, J_N$ are chosen so that $\int_0^1 [k(t)]^2\,dt$ is a minimum - then the remainder vanishes for any third order cubic spline with the knots $t_0, t_1, \ldots, t_N$, $N \geq 2$.

Uniform cubic spline $N \geq 3$.

Let $\lambda_0, \lambda_1, y_0, y_1, y_2$ be connected by a relation of the form

$$a\lambda_0 + \lambda_1 = b[t_0, t_1]y + c[t_1, t_2]y,$$

then if $y_j = x_j$, $j = 0,1,2$, we can rewrite this as

$$a(\lambda_0 - x_0^{(1)}) + (\lambda_1 - x_1^{(1)}) = b[t_0,t_1]x + c[t_1,t_2]x - a x_0^{(1)} - x_1^{(1)}.$$

a, b, c are given values so that the right hand side is $O(h^3)$. This leads to

$$\alpha_1(\lambda_0 - x_0^{(1)}) + (\lambda_1 - x_1^{(1)}) = (3\alpha_1 - \alpha_1^2)[t_0,t_1]x + (1-\alpha_1)^2[t_1,t_2]x - \alpha_1 x_0^{(1)} - x_1^{(1)} \qquad (A.7a)$$
$$= \frac{1}{24} h_0^2 h_1 \; x^{(4)}.$$

Similarly the appropriate equation at the other end is

$$(\lambda_{N-1} - x_{N-1}^{(1)}) + (1-\alpha_{N-1})(\lambda_N - x_N^{(1)}) = (2-\alpha_{N-1}-\alpha_{N-1}^2)[t_{N-1},t_N]x + \alpha_{N-1}^2[t_{N-2},t_{N-1}]x$$
$$- x_{N-1}^{(1)} - (1-\alpha_{N-1})x_N^{(1)} \qquad (A.7b)$$
$$= -\frac{1}{24} h_{N-1}^2 \; h_{N-2} \; x^{(4)}.$$

Consequently we arrive at a spline for which the following relations hold,

$$\alpha_1\lambda_0 + \lambda_1 = \alpha_1(3-\alpha_1)[t_0,t_1]y + (1-\alpha_1)^2[t_1,t_2]y,$$
$$\alpha_j\lambda_{j-1} + 2\lambda_1 + (1-\alpha_j)\lambda_{j+1} = 3\alpha_j[t_{j-1},t_j]y + 3(1-\alpha_j)[t_j,t_{j+1}]y, \quad j=1(1)N-1, \qquad (A.8)$$
$$\lambda_{N-1} + (1-\alpha_{N-1})\lambda_N = (1-\alpha_{N-1})(2+\alpha_{N-1})[t_{N-1},t_N] + \alpha_{N-1}^2[t_{N-2},t_{N-1}]y.$$

It can be shown without much difficulty that the first two of these equations imply that this cubic spline is such that $y^{(3)}$ is continuous at $t=t_1$. Similarly the last two equations imply that $y^{(3)}$ is continuous at $t=t_{N-1}$. In other words the cubic spline is a single cubic in $[t_0,t_2]$ and a single cubic in $[t_{N-2},t_N]$.

Definition

A uniform cubic spline with the knots $t_0, t_1, \ldots, t_N$, $N \geq 3$ is in $C^2[0,1]$ and such that

(a) it is a polynomial of degree at most three in $[t_j, t_{j+1}]$, $j = 0,1,\ldots,N-1$

(b) $y^{(3)}$ is continuous at t_1, t_{N-1}.

Theorem A.5

A uniform cubic spline is determined uniquely by its values at the knots for $N \geq 3$.

Proof

The uniqueness depends on the solvability of the equations (A.8). For $N \geq 3$ eliminate λ_0, λ_N from these equations. This leads to

$$\lambda_1 + (1-\alpha_1)\lambda_2 = \alpha_1^2 [t_0, t_1]y + (1-\alpha_1)(2+\alpha_1)[t_1, t_2]y$$

$$\alpha_j \lambda_{j-1} + 2\lambda_j + (1-\alpha_j)\lambda_{j+1} = 3\alpha_j [t_{j-1}, t_j]y + 3(1-\alpha_j)[t_j, t_{j+1}]y \quad j = 2,3,\ldots,N-2,$$

$$\alpha_{N-1}\lambda_{N-2} + \lambda_{N-1} = \alpha_{N-1}(3-\alpha_{N-1})[t_{N-2}, t_{N-1}]y + (1-\alpha_{N-1})^2 [t_{N-1}, t_N]y.$$

Since $0 < \alpha_1, \alpha_{N-1} < 1$ these equations have a strictly diagonally dominant matrix and so $\lambda_1, \lambda_2, \ldots, \lambda_{N-1}$ can be found if $y_0, y_1, \ldots, y_N$ are known. λ_0, λ_N can be calculated from (A.8). Hence the spline is uniquely determined.

Theorem A.6

If y is a uniform cubic spline such that $y_i = x_i$, $i = 0,1,\ldots,N$

where $x \in C^4[0,1]$ then, for $N \geq 4$,

$$|x(t)-y(t)| = O(h^4) + \frac{h_0^2}{h_1} O(h^3), \quad t_0 \leq t \leq t_1$$

$$|x(t)-y(t)| = O(h^4), \qquad t_1 \leq t \leq t_{N-1}$$

$$|x(t)-y(t)| = O(h^4) + \frac{h_{N-1}^2}{h_{N-2}} O(h^3), t_{N-1} \leq t \leq t_N.$$

Proof

Let $e_i = \lambda_i - x_i^{(1)}$, $i = 0,1,N$, and eliminate e_0, e_1, e_{N-1}, e_N from equations (A.3) and (A.7a), (A.7b). This will lead to the following equations

$$[2-\alpha_2(1-\alpha_1)]e_2 + (1-\alpha_2)e_3 = O(h^3)$$

$$\alpha_j e_{j-1} + 2e_j + (1-\alpha_j)e_{j+1} = O(h^3), \qquad j = 2,3,\ldots,N-2$$

$$\alpha_{N-2}e_{N-3} + [2-\alpha_{N-1}(1-\alpha_{N-2})]e_{N-2} = O(h^3).$$

The matrix of these equations has the strict diagonal dominance property and the uniform norm of its inverse is bounded by unity. Hence we have

$$\max|e_j| = O(h^3), \quad j = 2,3,\ldots,N-2.$$

Further, since $e_1 + (1-\alpha_1)e_2 = O(h^3)$, it follows that $|e_1| = O(h^3)$, with a similar result for $|e_{N-1}|$. However the bound for $|e_0|$ is to be found from (A.7a). That is

$$\alpha_1 |e_0| \leq |e_1| + \frac{1}{24} h_0^2 h_1 M_4,$$

consequently

$$|e_0| = \frac{1}{\alpha_1} O(h^3) = \left(1 + \frac{h_0}{h_1}\right) O(h^3).$$

Similarly $|e_N| = \left(1 + \frac{h_{N-1}}{h_{N-2}}\right) O(h^3)$.

Now, from (A.7a), in $[t_0,t_1]$ we have

$$|x(t)-y(t)| \leq \frac{1}{4} h_0 \max(|e_0|,|e_1|) + O(h^4)$$

$$= O(h^4) + \left(1 + \frac{h_0^2}{h_1}\right) O(h^4).$$

There is a similar result for the interval $[t_{N-1},t_N]$. The bounds in the remaining intervals depend on $|e_1|,\ldots,|e_{N-1}|$ which are simply $O(h^3)$. The result follows easily.

The uniform cubic spline arises also as the solution of an interpolation problem different from the one which introduced it, moreover it is connected with a type of best quadrature formula. The next two theorems will be stated only.

Theorem A.7

The unique $z \in C^2[0,1]$ such that

(a) $z_j=x_j$, $j = 0,1,\ldots,N$, $j \neq 1,N-1$,

(b) $z_0^{(1)}=x_0^{(1)}$, $z_N^{(1)}=x_N^{(1)}$,

(c) $\int_0^1 [z^{(2)}(t)]^2 dt$ is a minimum,

is a uniform cubic spline with the knots $t_0,t_1,\ldots,t_N$.

Theorem A.8

If

$$\int_0^1 m(t)x(t)dt = J_0 x_0^{(1)} + \sum_{\substack{i=0 \\ i\neq 1,N-1}}^{N} H_i x_i + J_N x_N^{(1)} + \int_0^1 k(t)x^{(2)}(t)dt$$

is a best quadrature formula with remainder (c.f. theorem A.8), then the remainder vanishes when x is any uniform cubic spline with the knots $t_0, t_1, \ldots, t_N$.

Uniform cubic spline quadrature

Since the uniform cubic spline can be used for interpolation it will create a quadrature formula with the knots as quadrature points. When the knots are equi-spaced the quadrature weights can be calculated explicitly. If there are four knots the formula will be the three eighths rule, and with five knots it will be once repeated Simpson's rule. The general results are contained in the next theorem. The proof of this is very similar to that of theorem 3.8 and so only a sketch will be given.

Theorem A.9

If

$$\int_0^1 y(t)dt = h \sum_{j=0}^{N} H_j y(jh), \quad h = 1/N, \ N \geq 4,$$

is true for any uniform cubic spline with the knots $0, h, 2h, \ldots, Nh$ then the quadrature weights are as follows.

$$N=5, \quad H_0 = H_5 = \frac{41}{120}, \ H_1 = H_4 = \frac{152}{120}, \ H_2 = H_3 = \frac{107}{120}.$$

$$N \geq 6, \quad H_0 = H_N = \frac{1}{2} + \frac{1}{72}[2+3T_{N-3}-2T_{N-1}]/U_{N-3}$$

$$H_1 = H_{N-1} = 1 + \frac{1}{18}[T_{N-1}-3T_{N-3}-4]/U_{N-3}$$

$$H_2 = H_{N-2} = 1 + \frac{1}{72}[T_{N-1}+72U_{N-4}+74]/U_{N-3}$$

$$H_i = 1 + [U_{N-i-2}+U_{i-2}]/U_{N-3}, \quad i = 3,4,\ldots,N-3,$$

where the argument of the Chebyshev polynomials is -2.

Proof

From (3.10) and (A.8), with $\alpha_j = \frac{1}{2}$, $j = 0,1,N-1$, we see that it is necessary to calculate $\lambda_N - \lambda_0$ from the equations

$$h[\tfrac{1}{2}\lambda_0+\lambda_1] = (-5y_0+4y_1+y_2)/4$$

$$h[\lambda_{j-1}+4\lambda_j+\lambda_{j+1}] = 3(y_{j+1}-y_{j-1}), \qquad j = 1,2,\ldots,N-1,$$

$$h[\lambda_{N-1}+\tfrac{1}{2}\lambda_N] = (5y_N-4y_{N-1}-y_{N-2})/4.$$

The inverse of this set of equations can be calculated from (1.5) and with its help it will be found that

$$3hU_{N-3}(\lambda_N-\lambda_0)=T_{-1}\left\{(T_{N-1}-T_{-1})\frac{(-5y_0+4y_1+y_2)}{4}+3\sum_{j=1}^{N-1}[T_{N-j-1}-T_{j-1}](y_{j+1}-y_{j-1}) + (T_{-1}-T_{N-1})\frac{(5y_N-4y_{N-1}-y_{N-2})}{4}\right\}.$$

When this is inserted in (3.10) the result will be found after some rearrangement and simplification.

The proof of the final theorem follows easily from theorem A.6

Theorem A.10

If $x \in C^4[0,1]$, and $H_0,H_1,\ldots,H_N$ are chosen so that

$$\int_0^1 m(t)y(t)dt = \sum_{j=0}^{N} H_j y_j$$

is exact for any uniform cubic spline with the knots $t_0,t_1,\ldots,t_N$, $N\geq 3$, then

$$\left|\int_0^1 m(t)x(t)dt - \sum_{j=0}^{N} H_j x_j\right| = O(h^4) + \left(\frac{h_0^3}{h_1} + \frac{h_{N-1}^3}{h_{N-2}}\right) O(h^3).$$

References

[1] Ahlberg, J. H., Nilson, E. N. and Walsh, J. L. The Theory of Splines and Their Applications. Academic Press (1967).

[2] Cheney, E. W. and Schurer, F. A Note on the Operators Arising in Spline Approximation. J. Approx. Theory 1 (1968) pp. 94-102.

[3] Greville, T. N. E. Introduction to Spline Functions. In 'Theory and Applications of Spline Functions'. Edited by T. N. E. Greville. Academic Press (1969).

[4] Hall, C. A. On Error Bounds for Spline Interpolation. J. Approx. Theory 1 (1968) pp. 209-218.

[5] Handscomb, D. C. Optimal Approximation of Linear Functionals. In 'Methods of Numerical Approximation'. Edited by D. C. Handscomb. Pergamon Press (1966).

[6] Karlin, S. The Fundamental Theorem of Algebra for Monosplines Satisfying Certain Boundary Conditions and Applications to Optimal Quadrature Formulas. In 'Approximations with Special Emphasis on Spline Functions'. Edited by I. J. Schoenberg. Academic Press (1969).

[7] Kershaw, D. The Explicit Inverses of Two Commonly Occurring Matrices. Math. Comp. 23 (1969) pp. 189-191. (Sewn in)

[8] Kershaw, D. Inequalities on the Elements of the Inverse of a Certain Tridiagonal Matrix. Math. Comp. 24 (1970) pp. 155-158. (Sewn in)

[9] Kershaw, D. A Note on the Convergence of Interpolatory Cubic Splines. SIAM J. Numerical Analysis 8 (1971) pp. 67-74. (Sewn in)

[10] Meir, A. and Sharma, A. On Uniform Approximation by Cubic Splines. J. Approx. Theory 2 (1969) pp. 270-274.

[11] Nord, S. Approximation properties of the spline fit. BIT 7 (1967) pp. 132-144.

[12] Sard, A. Best approximate integration formulas; best approximation formulas. Amer. J. Math. 71 (1949) pp. 80-91.

[13] Schoenberg, I. J. Contributions to the problem of approximation of equidistant data by analytic functions. Qu. Appl. Math. 4 (1946), pp. 45-99, 112-141.

[14] Schoenberg, I. J. Spline interpolation and best quadrature formulae. Bull. Amer. Math. Soc. 70 (1964), pp. 143-148.

[15] Schoenberg, I. J. Monosplines and Quadrature Formulae. In 'Theory and Applications of Spline Functions'. Edited by T. N. E. Grenville. Academic Press (1969).

The Orders of Approximation of the First Derivative of Cubic Splines at the Knots

By D. Kershaw

Reprinted from
MATHEMATICS OF COMPUTATION
January, 1972, Vol. 26, No. 117

MATHEMATICS OF COMPUTATION, VOLUME 26, NUMBER 117, JANUARY, 1972

The Orders of Approximation of the First Derivative of Cubic Splines at the Knots

By D. Kershaw

Abstract. The order of approximation of the first derivative of four types of interpolating cubic splines are found. The splines are defined by a variety of endpoint conditions and include the natural cubic spline and the periodic cubic spline. It is found that for two types there is an increase in the order of approximation when equal intervals are used, and that for a special distribution of knots the same order can be realized for the natural spline.

1. Introduction. The cubic spline is now a well established tool for smooth interpolation in a table of a function defined at a discrete set of points. A useful account of the basic properties of this spline and an algorithm for constructing it can be found in [1], and an analysis of the convergence of the spline to the function it interpolates is given in [4].

The present paper is devoted to an investigation of the problem of finding how well the first derivative, taken at the knots, of the spline approximates the first derivative of the interpolated function there. It was shown in [4] that there is $O(h^3)$ approximation uniformly over the range of the knots, as the maximum interval tends to zero, but as it is often the case that the derivative is taken at the knots, it is felt that the results may be of some value.

2. Notation. The set of real numbers, $t_0, t_1, \cdots, t_N$, will be called *knots* and will satisfy

$$-\infty < t_0 < t_1 < \cdots < t_{N-1} < t_N < \infty, \qquad N \geqq 2.$$

The *interval* $t_i \leqq t \leqq t_{i+1}$ will have *length* $h_i = t_{i+1} - t_i$, $i = 0(1)N - 1$, and the *maximum interval length* will be h, that is,

$$h = \max_{0 \leqq i \leqq N-1} h_i.$$

y will denote a *cubic spline* with the above knots. As stated in Section 1, more than one kind of spline will be considered but they will have the common property that each is a member of $C^2(-\infty, \infty)$ and that in each interval they are polynomials of degree at most three.

x will be a member of $C^5[t_0, t_N]$ and will be the function with which the spline agrees at the knots. For brevity, define

$$x_i^{(r)} = \left(\frac{d}{dt}\right)^r x(t), \quad \text{for } t = t_i,\ i = 0(1)N,\ r = 0(1)5.$$

Received September 22, 1970.
AMS 1970 *subject classifications*. Primary 65D25; Secondary 41A15.
Key words and phrases. Interpolation, cubic splines.

Then

$$y_i = x_i, \qquad i = 0(1)N.$$

The norms which will be used are the uniform norms for functions, vectors and matrices, namely,

$$\|x\| = \max_{t_0 \le t \le t_N} |x(t)|, \qquad \|\mathbf{x}\| = \max_i |x_i|, \qquad \|A\| = \max_i \sum_j |a_{ij}|.$$

The domain of the suffixes in the vector and matrix norms will be clear from the context.

It is convenient to define here

$$M_4 = \|x^{(4)}\|, \qquad M_5 = \|x^{(5)}\|.$$

The first and last columns of the $(n + 1) \times (n + 1)$ unit matrix will be written respectively as $\mathbf{e}_0$, $\mathbf{e}_n$; the jth element of the vector $\mathbf{x}$ will be denoted by $[\mathbf{x}]_j$.

3. The Cubic Splines. Four types of cubic splines will be described in this section. Cubic splines are usually characterized by the value of their second derivative at each of the knots (see for example [1]), but for the purpose of this note, an alternative method will be used.

Let

$$\lambda_i = y_i^{(1)}, \qquad i = 0(1)N,$$

then, if $y(t)$ takes the same value as $x(t)$ at each of the knots, it follows from Hermite's two point interpolation formula that, for $t_i \leqq t \leqq t_{i+1}$,

$$\begin{aligned} y(t) = {} & \left[3\left(\frac{t_{i+1} - t}{h_i}\right)^2 - 2\left(\frac{t_{i+1} - t}{h_i}\right)^3\right]x_i + \left[3\left(\frac{t - t_i}{h_i}\right)^2 - 2\left(\frac{t - t_i}{h_i}\right)^3\right]x_{i+1} \\ & + h_i\left[\left(\frac{t_{i+1} - t}{h_i}\right)^2 - \left(\frac{t_{i+1} - t}{h_i}\right)^3\right]\lambda_i - h_i\left[\left(\frac{t - t_i}{h_i}\right)^2 - \left(\frac{t - t_i}{h_i}\right)^3\right]\lambda_{i+1}, \end{aligned} \tag{1}$$

$$i = 0(1)N - 1.$$

A simple calculation shows that

$$\begin{aligned} h_i^2 y_i^{(2)} &= 6(x_{i+1} - x_i) - h_i(4\lambda_i + 2\lambda_{i+1}), \\ h_i^2 y_{i+1}^{(2)} &= -6(x_{i+1} - x_i) + h_i(2\lambda_i + 4\lambda_{i+1}). \end{aligned} \tag{2}$$

Now, as $y \in C^2(-\infty, \infty)$, the two expressions for $y_i^{(2)}$ from the equations which arise from the intervals (t_{i-1}, t_i), (t_i, t_{i+1}) must be equal. The identification gives the equations:

$$\frac{\lambda_{i-1} + 2\lambda_i}{h_{i-1}} + \frac{2\lambda_i + \lambda_{i+1}}{h_i} = 3\left[\frac{x_{i+1} - x_i}{h_i^2} + \frac{x_i - x_{i-1}}{h_{i-1}^2}\right], \qquad i = 1(1)N - 1. \tag{3}$$

It is convenient to define

$$\alpha_i = h_{i-1}/(h_{i-1} + h_i),$$

then the equations become

$$(1-\alpha_i)\lambda_{i-1} + 2\lambda_i + \alpha_i\lambda_{i+1} = 3\left[\alpha_i \frac{(x_{i+1}-x_i)}{h_i} + (1-\alpha_i)\frac{(x_i - x_{i-1})}{h_{i-1}}\right],$$

$$i = 1(1)N-1,$$

which can be written as

$$(1-\alpha_i)(\lambda_{i-1} - x_{i-1}^{(1)}) + 2(\lambda_i - x_i^{(1)}) + \alpha_i(\lambda_{i+1} - x_{i+1}^{(1)})$$
$$= -(1-\alpha_i)x_{i-1}^{(1)} - 2x_i^{(1)} - \alpha_i x_{i+1}^{(1)} + 3\left[\alpha_i \frac{(x_{i+1}-x_i)}{h_i} + (1-\alpha_i)\frac{(x_i - x_{i-1})}{h_{i-1}}\right],$$

$$i = 1(1)N-1.$$

Finally, the use of Peano's method for finding remainders gives the result that

$$(1-\alpha_i)(\lambda_{i-1} - x_{i-1}^{(1)}) + 2(\lambda_i - x_i^{(1)}) + \alpha_i(\lambda_{i+1} - x_{i+1}^{(1)})$$

$$= \frac{1}{24} h_{i-1}h_i(h_{i-1} - h_i)x_i^{(4)} - \frac{1}{60} h_{i-1}h_i(h_{i-1}^2 + h_i^2 - h_{i-1}h_i)x^{(5)}(\tau_i), \tag{4}$$

where $t_{i-1} \leqq \tau_i \leqq t_{i+1}$, $i = 1(1)N-1$.

The sets of Eqs. (3), (4) are satisfied by $\lambda_0, \lambda_1, \cdots, \lambda_N$ for each of the splines to be considered. Clearly, two further relations are needed in order that a unique interpolating spline may be found. The equations (3) are the useful ones for the actual calculation of the splines and, for completeness, the two relations to be adjoined to (3) will be given for the different types of splines to be described. For this note, however, (4) are the useful equations and these relations will have to be written in a form similar to (4).

(A) *Natural Cubic Spline*. The relations which help to define this spline are [1]

$$y_0^{(2)} = y_N^{(2)} = 0,$$

whence, from (2), the equations additional to (3) are

$$\text{(5a)}\qquad \begin{aligned} 2\lambda_0 + \lambda_1 &= \frac{3}{h_0}(x_1 - x_0), \\ \lambda_{N-1} + 2\lambda_N &= \frac{3}{h_{N-1}}(x_N - x_{N-1}). \end{aligned}$$

With the aid of Peano's method these can be written

$$\text{(5b)}\qquad \begin{aligned} 2(\lambda_0 - x_0^{(1)}) + (\lambda_1 - x_1^{(1)}) &= \tfrac{1}{2}h_0 x_0^{(2)} - \frac{1}{24} h_0^3 x_0^{(4)} - \frac{1}{60} h_0^4 x^{(5)}(\tau_0), \\ & \qquad t_0 \leqq \tau_0 \leqq t_1, \quad \text{and} \\ (\lambda_{N-1} - x_{N-1}^{(1)}) + 2(\lambda_N - x_N^{(1)}) &= -\tfrac{1}{2}h_{N-1}x_N^{(2)} + \frac{1}{24} h_{N-1}^3 x_N^{(4)} + \frac{1}{60} h_{N-1}^4 x^{(5)}(\tau_N), \\ & \qquad t_{N-1} \leqq \tau_N \leqq t_N. \end{aligned}$$

These equations together with (4) are, in matrix form,

$$
\text{(5c)}\qquad \begin{bmatrix} 2 & 1 & 0 & \cdots & 0 \\ 1-\alpha_1 & 2 & \alpha_1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 2 \end{bmatrix} (\boldsymbol{\lambda} - \mathbf{x}^{(1)})
$$

$$
= \frac{1}{2} h_0 \left[x_0^{(2)} - \frac{1}{12} h_0^2 x_0^{(4)} \right] \mathbf{e}_0 - \frac{1}{2} h_{N-1} \left[x_N^{(2)} - \frac{1}{12} h_{N-1}^2 x_N^{(4)} \right] \mathbf{e}_N + \mathbf{x}^{(4)} + \mathbf{x}^{(5)}
$$

where

$$
\boldsymbol{\lambda} - \mathbf{x}^{(1)} = [\lambda_0 - x_0^{(1)} \; \lambda_1 - x_1^{(1)} \; \cdots \; \lambda_N - x_N^{(1)}]^T,
$$

$$
\text{(5d)}\qquad \mathbf{x}^{(4)} = \frac{1}{24} [0 \; h_0 h_1 (h_0 - h_1) x_1^{(4)} \; \cdots \; h_{N-2} h_{N-1} (h_{N-2} - h_{N-1}) x_{N-1}^{(4)} \; 0]^T,
$$

$$
\text{(5e)}\qquad \mathbf{x}^{(5)} = -\frac{1}{60} [h_0^4 x^{(5)}(\tau_0) \; h_0 h_1 (h_0^2 + h_1^2 - h_0 h_1) x^{(5)}(\tau_1) \; \cdots \; -h_{N-1}^4 x^{(5)}(\tau_N)]^T.
$$

(B) *Cubic Spline D*1. Here, $y_0^{(1)}$ and $y_N^{(1)}$ are fitted exactly, and so

$$
\text{(6a)}\qquad \lambda_0 = x_0^{(1)}, \qquad \lambda_N = x_N^{(1)}
$$

are the equations to be put with (3) for the calculation of this spline. Further, (4) can now be written as

$$
\text{(6b)}\qquad \begin{bmatrix} 2 & \alpha_1 & 0 & \cdots & 0 & 0 \\ 1-\alpha_2 & 2 & \alpha_2 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1-\alpha_{N-1} & 2 \end{bmatrix} (\boldsymbol{\lambda} - \mathbf{x}^{(1)}) = \mathbf{x}^{(4)} + \mathbf{x}^{(5)},
$$

where

$$
\boldsymbol{\lambda} - \mathbf{x}^{(1)} = [\lambda_1 - x_1^{(1)} \; \cdots \; \lambda_{N-1} - x_{N-1}^{(1)}]^T,
$$

$$
\text{(6c)}\qquad \mathbf{x}^{(4)} = \frac{1}{24} [h_0 h_1 (h_0 - h_1) x_1^{(4)} \; \cdots \; h_{N-2} h_{-N-1} (h_{N-2} - h_{N-1}) x_{N-1}^{(4)}]^T,
$$

$$
\text{(6d)}\qquad \mathbf{x}^{(5)} = -\frac{1}{60} [h_0 h_1 (h_0^2 + h_1^2 - h_0 h_1) x^{(5)}(\tau_1)
$$

$$
\cdots h_{N-2} h_{N-1} (h_{N-2}^2 + h_{N-1}^2 - h_{N-2} h_{N-1}) x^{(5)}(\tau_{N-1})]^T.
$$

(C) *Cubic Spline D*2. If

$$
y_0^{(2)} = x_0^{(2)}, \qquad y_N^{(2)} = x_N^{(2)},
$$

then, from (2), the equations additional to (3) are

$$
\text{(7)}\qquad \begin{aligned} 2\lambda_0 + \lambda_1 &= \frac{3}{h_0} (x_1 - x_0) - h_0 x_0^{(2)}, \\ \lambda_{N-1} + 2\lambda_N &= \frac{3}{h_{N-1}} (x_N - x_{N-1}) + h_{N-1} x_N^{(2)}. \end{aligned}
$$

Peano's theorem gives the results

$$2(\lambda_0 - x_0^{(1)}) + (\lambda_1 - x_1^{(1)}) = -\frac{1}{24} h_0^3 x_0^{(4)} - \frac{1}{60} h_0^4 x^{(5)}(\tau_0),$$

$$(\lambda_{N-1} - x_{N-1}^{(1)}) + 2(\lambda_N - x_N^{(1)}) = \frac{1}{24} h_{N-1}^3 x_N^{(4)} - \frac{1}{60} h_{N-1}^4 x^{(5)}(\tau_N).$$

On comparison with the corresponding ones for the cubic spline, namely (5b), it is seen that the matrix equation for this spline is identical with (5c) except that the terms $x_0^{(2)}$ and $x_N^{(2)}$ are replaced by zero.

(D) *Periodic Cubic Spline.* When x has period $t_N - t_0$ and $x_0^{(r)} = x_N^{(r)}, r = 0, 1, \cdots,$ then the spline can be taken to be periodic in the sense that

(8a) $$y_0^{(r)} = y_N^{(r)}, \qquad r = 0, 1, 2.$$

The Eqs. (3) remain valid but in the first λ_0, x_0 can be replaced by λ_N, x_N, respectively. An additional equation arises from the observation that $y_0^{(2)} = y_N^{(2)}$ and is, after simplification,

(8b) $$\beta\lambda_1 + (1 - \beta)\lambda_{N-1} + 2\lambda_N = 3\left[\beta\left(\frac{x_1 - x_0}{h_0}\right) + (1 - \beta)\left(\frac{x_N - x_{N-1}}{h_{N-1}}\right)\right].$$

In the required form, this is

(8c) $$\begin{aligned}&\beta(\lambda_1 - x_1^{(1)}) + (1 - \beta)(\lambda_{N-1} - x_{N-1}^{(1)}) + 2(\lambda_N - x_N^{(1)})\\ &\quad = \frac{1}{24} h_{N-1}h_0(h_{N-1} - h_0)x_0^{(4)} - \frac{1}{60} h_{N-1}h_0(h_{N-1}^2 + h_0^2 - h_{N-1}h_0)x^{(5)}(\pi)\end{aligned}$$

where $\beta = h_{N-1}/(h_0 + h_{N-1})$, and $t_0 - h_{N-1} \leqq \pi \leqq t_1$.

Thus, the matrix equation is

(8d) $$\begin{bmatrix} 2 & \alpha_1 & 0 & \cdots & 0 & 1 - \alpha_1 \\ 1 - \alpha_2 & 2 & \alpha_2 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \beta & 0 & 0 & \cdots & 1 - \beta & 2 \end{bmatrix} (\boldsymbol{\lambda} - \mathbf{x}^{(1)}) = \mathbf{x}^{(4)} + \mathbf{x}^{(5)},$$

where

$$\boldsymbol{\lambda} - \mathbf{x}^{(1)} = [\lambda_1 - x_1^{(1)} \cdots \lambda_N - x_N^{(1)}]^T,$$

(8e) $$\mathbf{x}^{(4)} = \frac{1}{24} [h_0h_1(h_0 - h_1)x_1^{(4)} \cdots h_{N-1}h_0(h_{N-1} - h_0)x_N^{(4)}]^T,$$

(8f) $$\mathbf{x}^{(5)} = -\frac{1}{60} [h_0h_1(h_0^2 + h_1^2 - h_0h_1)x^{(5)}(\tau_1) \cdots h_{N-1}h_0(h_{N-1}^2 + h_0^2 - h_{N-1}h_0)x^{(5)}(\pi)]^T.$$

4. Error in the First Derivatives of the Splines at the Knots. It will be noticed that the matrices which occurred in Section 3 for each of the splines are strictly diagonally dominant, and so the equations can be solved. Further, if A represents any of them then, with the uniform norm $||A^{-1}|| \leqq 1$. This follows from the observation that if $||A\mathbf{x}|| \geqq 1$ for $||\mathbf{x}|| = 1$, then $||A^{-1}|| \leqq 1$. Now, $A = 2I + B$, where $||B|| \leqq 1$ and so $||A\mathbf{x}|| \geqq 2||\mathbf{x}|| - ||B\mathbf{x}||$ and, as $||B\mathbf{x}|| \leqq ||B|| \leqq 1$, the result is proved.

THEOREM 1. *If y is either a cubic spline $D1$ or a periodic cubic spline, then*

$$\|\lambda - \mathbf{x}^{(1)}\| \leqq \frac{1}{24} h^2 \max_i |h_{i-1} - h_i| \cdot M_4 + \frac{1}{60} h^4 M_5. \tag{9}$$

Proof. In (6b) and (8d), multiply by the inverse of the respective matrices, and take the uniform norm of each side. Then,

$$\|\lambda - \mathbf{x}^{(1)}\| \leqq \|\mathbf{x}^{(4)}\| + \|\mathbf{x}^{(5)}\|,$$

where $\mathbf{x}^{(4)}$, $\mathbf{x}^{(5)}$ are defined by (6c), (6d) for the $D1$ spline and by (8e), (8f) for the periodic spline.

The results now follow on taking the uniform norms of $\mathbf{x}^{(4)}$, $\mathbf{x}^{(5)}$.

COROLLARY. *If $h_i = h$, $i = 0(1)N - 1$, then, if y is either a cubic spline $D1$ or a periodic cubic spline, then*

$$\|\lambda - \mathbf{x}^{(1)}\| \leqq \frac{1}{60} h^4 M_5. \tag{10}$$

The remaining types of splines will be taken together as the analysis is common to them both. The equations for the natural cubic spline are given by (5c). Denote by A the matrix. Then, after multiplying (5c) by A^{-1} it will easily be seen that

$$\begin{aligned} |\lambda_j - x_j^{(1)}| \leqq {} & h_0[C_1 + h_0^2 D_1]\, |[A^{-1}\mathbf{e}_0]_j| + h_{N-1}[C_2 + h_{N-1}^2 D_2]\, |[A^{-1}\mathbf{e}_N]_j| \\ & + \frac{1}{24} h^2 \max |h_{i-1} - h_i| \cdot M_4 + \frac{1}{60} h^4 M_5, \qquad j = 0(1)N, \end{aligned} \tag{11}$$

where

$$C_1 = \frac{1}{2} |x_0^{(2)}|, \qquad C_2 = \frac{1}{2} |x_N^{(2)}|, \qquad D_1 = \frac{1}{24} |x_0^{(4)}|, \qquad D_2 = \frac{1}{24} |x_N^{(4)}|.$$

The corresponding inequalities for the cubic spline $D2$ are found by putting $C_1 = C_2 = 0$ in (11) and are

$$\begin{aligned} |\lambda_j - x_j^{(1)}| \leqq {} & h_0^3 D_1\, |[A^{-1}\mathbf{e}_0]_j| \\ & + h_{N-1}^3 D_2\, |[A^{-1}\mathbf{e}_N]_j| + \frac{1}{24} h^2 \max_i |h_{i-1} - h_i|\, M_4 + \frac{1}{60} h^4 M_5. \end{aligned} \tag{12}$$

Clearly, the nonvanishing of the multipliers of $[A^{-1}\mathbf{e}_0]_j$, $[A^{-1}\mathbf{e}_N]_j$ have an adverse effect on the approximations in (12) when the intervals are equal, and for the natural spline this is apparently disastrous, even when the intervals are equal. But, on examination, it is seen that to increase the order of approximation in both cases it is necessary only to make the first and last intervals small enough. The situations can be saved a little in the general case of unequal intervals as shown in the following theorems.

THEOREM 2. *If y is a natural cubic spline, $h < 1$ and if $N \geqq 2 - 2r \log h/\log \alpha$, there exist integers p, q, $0 \leqq p < q \leqq N$, such that, for $p \leqq j \leqq q$,*

$$\begin{aligned} |\lambda_j - x_j^{(1)}| \leqq {} & \tfrac{2}{3} h_0 h^r [C_1 + h_0^2 D_1] + \tfrac{2}{3} h_{N-1} h^r [C_2 + h_{N-1}^2 D_2] \\ & + \frac{1}{24} h^2 \max_i |h_{i-1} - h_i| \cdot M_4 + \frac{1}{60} h^4 M_5, \end{aligned}$$

where the real number α *is*

(i) $2 + \sqrt{3}$ *if* $h_i = h$,

(ii) 2 *when the intervals are unequal.*

Also

$$t_p - t_0 < h[1 - r \log h/\log \alpha], \qquad t_N - t_q < h[1 - r \log h/\log \alpha].$$

Proof. This depends on results from [2], where it is shown that for equal intervals

$$|[A^{-1}\mathbf{e}_0]_j| = U_{N-j}(2)/U_{N+1}(2), \qquad |[A^{-1}\mathbf{e}_N]_j| = U_j(2)/U_{N+1}(2), \qquad j = 0(1)N,$$

and from [3], where it is shown that when the intervals are not equal

$$|[A^{-1}\mathbf{e}_0]_j| \leqq \tfrac{2}{3}\cdot 2^{j-N}, \qquad |[A^{-1}\mathbf{e}_N]_j| \leqq \tfrac{2}{3}\cdot 2^{-j}.$$

Now,

$$\frac{U_j(2)}{U_{N+1}(2)} = \frac{(2+\sqrt{3})^{j+1} - (2-\sqrt{3})^{j+1}}{(2+\sqrt{3})^{N+2} - (2-\sqrt{3})^{N+2}} < (2+\sqrt{3})^{j-N-1}, \qquad j = 0(1)N,$$

and similarly,

$$U_{N-j}(2)/U_{N+1}(2) < (2+\sqrt{3})^{-j-1}, \qquad j = 0(1)N.$$

Hence, (11) can be replaced by

$$|\lambda_j - x_j^{(1)}| \leqq h_0[C_1 + h_0^2 D_1]\alpha^{-j} + h_{N-1}[C_2 + h_{N-1}^2 D_2]\alpha^{j-N-1}$$

$$+ \frac{1}{24} h^2 \max_i |h_{i-1} - h_i| \cdot M_4 + \frac{1}{60} h^5 M_5, \qquad j = 0(1)N, \tag{13}$$

where $\alpha = 2 + \sqrt{3}$ if $h_i = h$ and $\alpha = 2$ otherwise.

(For simplicity of presentation, the factor $\frac{2}{3}$ which should occur in these inequalities when $\alpha = 2$ and the factor $2 - \sqrt{3}$ when $\alpha = 2 + \sqrt{3}$ have been replaced by unity.)

As $\alpha > 1$, it follows that α^{-j} decreases with increasing j, and so $\alpha^{-j} \leqq h^r$ for all $j \geqq p$ where the integer p satisfies $\alpha^{-p} \leqq h^r < \alpha^{-p+1}$, that is

$$-r \log h/\log \alpha \leqq p < 1 - r \log h/\log \alpha.$$

Similarly, $\alpha^{j-N} \leqq h^r$ for all $j \leqq q$ where the integer q satisfies

$$N - 1 + r \log h/\log \alpha < q \leqq N + r \log h/\log \alpha.$$

In order that $p < q$, it is sufficient that

$$N - 1 - r \log h/\log \alpha - 1 - r \log h/\log \alpha \geqq 0$$

which is equivalent to

$$N \geqq 2 - 2r \log h/\log \alpha.$$

It remains to note that

$$t_p - t_0 \leqq ph < h[1 - r \log h/\log \alpha],$$

$$t_N - t_q \leqq (N - q)h < h[1 - r \log h/\log \alpha].$$

(The inequality $N \geqq 2 - 2r \log h/\log \alpha$ will be satisfied for sufficiently large N as $Nh \geqq t_N - t_0$.)

COROLLARY. *If y is a cubic spline (D2), $h < 1$ and if $N \geqq 2 - 2r \log h/\log \alpha$ then there exist integers p, q, $0 \leqq p < q \leqq N$ such that, for $p \leqq j \leqq q$,*

$$|\lambda_i - x_i^{(1)}| \leqq \frac{2}{3} h_0^3 h^r D_1 + \frac{2}{3} h_{N-1}^3 h^r D_2 + \frac{1}{24} h^2 \max_i |h_{i-1} - h_2| \cdot M_4 + \frac{1}{60} h^4 M_5,$$

where α is

(i) $2 + \sqrt{3}$ *if $h_i = h$,*

(ii) 2 *when the intervals are unequal.*

Also

$$t_p - t_0 < h[1 - r \log h/\log \alpha], \qquad t_N - t_q < h[1 - r \log h/\log \alpha].$$

Proof. This follows from Theorem 2 on setting $C_1 = C_2 = 0$.

Conclusions. The approximation of the first derivative at the knots is best when equal intervals are used both for the cubic spline $D1$ and the periodic cubic spline. In each case, the approximation is $O(h^4)$. When unequal intervals are used, it drops to $O(h^3)$. For the cubic spline $D2$, the order is generally $O(h^3)$ whether the intervals are equal or not, but with equal intervals and for a large enough number of points, the order is $O(h^4)$ at a number of internal knots.

The first derivative of the natural cubic spline is only an $O(h)$ approximation to the first derivative of the interpolated function at the knots, although for a sufficiently large number of knots the order can be made $O(h^3)$ or $O(h^4)$ at a range of internal points if the intervals are respectively unequal or equal.

Similar theorems can be proved for other types of cubic splines with mixed end conditions. It is worth remarking that if one end only is 'natural', for example $y_N^{(2)} = 0$, then the effect of this on the approximation will decrease rapidly as this point is left (by a factor of $2 - \sqrt{3}$ for equal intervals and 0.5 for unequal intervals).

Department of Computer Science
University of Edinburgh
Edinburgh, Scotland

1. T. N. E. GREVILLE, "Spline functions, interpolation, and numerical quadrature," in *Mathematical Methods for Digital Computers*. Vol. 2, A. Ralston and H. S. Wilf (Editors), Wiley, New York, 1967. MR 35 #2516.

2. D. KERSHAW, "The explicit inverses of two commonly occurring matrices," *Math. Comp.*, v. 23, 1969, pp. 189–191. MR 38 #6754.

3. D. KERSHAW, "Inequalities on the elements of the inverse of a certain tridiagonal matrix," *Math. Comp.*, v. 24, 1970, 155–158. MR 41 #2907.

4. D. KERSHAW, "A note on the convergence of interpolatory cubic splines," *SIAM J. Numer. Anal.*, v. 8, 1971, pp. 67–74.

A reprint from:

Mathematics of Computation

Volume 23, Number 105

January, 1969

Published by the American Mathematical Society

PROVIDENCE, RHODE ISLAND

The Explicit Inverses of Two Commonly Occurring Matrices

By D. Kershaw

Abstract. Explicit formulae are given for the inverses of certain tridiagonal scalar and block matrices.

During an investigation into the convergence properties of natural splines it was found convenient to have the explicit forms for the inverses of certain tridiagonal matrices. Special forms of these matrices arise in other branches of numerical analysis and so it may be useful to record them.

The derivation is elementary but complicated and will only be indicated. ■

Notes.

(i) The matrices will be of order $n \times n$.

(ii) T_r, U_r will denote the Chebyshev polynomials of the first and second kinds respectively, both with argument λ.

(iii) The elements of A, A^{-1} will be denoted by a_{rs}, a_{rs}^{-1} respectively. Similarly for B, B^{-1}.

(iv) $U_{-1} = 0$, $U_{-2} = -U_0$.

Matrix $A, n \geqq 2$.

$$\begin{aligned} a_{rs} &= -a \quad && r = s = 1, n\,, \\ &= -2\lambda && r = s = 2(1)n-1\,, \\ &= 1 && |r - s| = 1\,, \\ &= 0 && \text{otherwise}\,, \end{aligned}$$

$$a_{rs}^{-1} = \frac{-1}{a^2U_{n-2} - 2aU_{n-3} + U_{n-4}} [aU_{r-2} - U_{r-3}][aU_{n-s-1} - U_{n-s-2}]\,, \qquad 1 \leqq r \leqq s \leqq n\,,$$

$$a_{sr}^{-1} = a_{rs}^{-1}\,.$$

Matrix $B, n \geqq 3$.

$$b_{rs} = a_{rs}\,, \qquad \text{except that } b_{1n} = b_{n1} = 1\,,$$

$$b_{rs}^{-1} = \frac{-1}{a^2U_{n-2} - 2aU_{n-3} - 2(1 + T_{n-2})} \left\{ \begin{matrix} [aU_{r-2} - U_{r-3}][aU_{n-s-1} - U_{n-s-2}] \\ + U_{s-r-1} - U_{r-2}U_{n-s-1} \end{matrix} \right\}, \qquad 1 \leqq r \leqq s \leqq n,$$

$$b_{sr}^{-1} = b_{rs}^{-1}\,.$$

Special Forms of A.

Received May 17, 1968.

$$a = 2\lambda\,,$$

$$A = \begin{bmatrix} -2\lambda & 1 & 0 \cdots 0 & 0 \\ 1 & -2\lambda & 1 \cdots 0 & 0 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ 0 & 0 & 0 \cdots 1 & -2\lambda \end{bmatrix},$$

$$a_{rs}^{-1} = -\frac{1}{U_n}\cdot U_{r-1}U_{n-s}\,, \quad 1 \leqq r \leqq s \leqq n,\, a_{sr}^{-1} = a_{rs}^{-1}\,,$$

$$a = \lambda\,,$$

$$A = \begin{bmatrix} -\lambda & 1 & 0 \cdots & 0 & 0 \\ 1 & -2\lambda & 1 \cdots & 0 & 0 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot \\ 0 & 0 & 0 \cdots & -2\lambda & 1 \\ 0 & 0 & 0 \cdots & 1 & -\lambda \end{bmatrix},$$

$$a_{rs}^{-1} = \frac{1}{(1-\lambda^2)U_{n-2}}\cdot T_{r-1}T_{n-s}\,, \quad 1 \leqq r \leqq s \leqq n,\, a_{sr}^{-1} = a_{rs}^{-1}\,.$$

Special Forms of B.

$$a = 2\lambda\,,$$

$$B = \begin{bmatrix} -2\lambda & 1 & 0 \cdots & 0 & 1 \\ 1 & -2\lambda & 1 \cdots & 0 & 0 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot \\ 0 & 0 & 0 \cdots & -2\lambda & 1 \\ 1 & 0 & 0 \cdots & 1 & -2\lambda \end{bmatrix},$$

$$b_{rs}^{-1} = \frac{1}{2[1-T_n]}\,[U_{n-s+r-1} + U_{s-r-1}]\,, \quad 1 \leqq r \leqq s \leqq n\,,\, b_{sr}^{-1} = b_{rs}^{-1}\,,$$

$$a = \lambda\,,$$

$$B = \begin{bmatrix} -\lambda & 1 & 0 \cdots & 0 & 1 \\ 1 & -2\lambda & 1 \cdots & 0 & 0 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot \\ 0 & 0 & 0 \cdots & -2\lambda & 1 \\ 1 & 0 & 0 \cdots & 1 & -\lambda \end{bmatrix},$$

$$b_{rs}^{-1} = \frac{1}{2 + (2-\lambda^2)U_{n-2}}\,[T_{r-1}T_{n-s} - U_{r-2}U_{n-s-1} + U_{s-r-1}]\,,$$
$$1 \leqq r \leqq s \leqq n,\, b_{sr}^{-1} = b_{rs}^{-1}\,.$$

Outlines of Proof for A. The columns of A^{-1} are the solutions of

$$A\mathbf{x} = \mathbf{e}_s\,, \quad s = 1, 2, \cdots, n\,,$$

where $\mathbf{e}_s$ is the sth unit vector. In recurrence form this matrix equation becomes

$$\begin{gathered} -ax_1 + x_2 = 0\,, \\ x_{r-1} - 2\lambda x_r + x_{r+1} = \delta_{rs}\,, \quad r = 1, 2, \cdots, n-1\,, \\ x_{n-1} - ax_n = 0\,. \end{gathered} \tag{1}$$

Note now that all except the first, last and sth equations are satisfied by either Chebyshev polynomial with argument λ. If one assumes, for example, that

$$
\begin{aligned}
x_r &= AT_r + BU_{r-1}\,, \quad r = 1, 2, \cdots, s-1\,, \\
x_r &= CT_r + DU_{r-1}\,, \quad r = s+1, \cdots, n\,,
\end{aligned} \tag{2}
$$

then each equation in (1) will be identically satisfied except for

$$
\begin{aligned}
-ax_1 + x_2 &= 0\,, \\
x_{s-2} - 2\lambda x_{s-1} + x_s &= 0\,, \\
x_{s-1} - 2\lambda x_s + x_{s+1} &= 1\,, \\
x_s - 2\lambda x_{s+1} + x_{s+2} &= 0\,, \\
x_{n-1} - ax_n &= 0\,.
\end{aligned}
$$

These give sufficient equations to solve for the unknowns A, B, C, D, x_s. A similar technique holds for B.

Generalization to Partitioned Matrices. If, in the matrices A and B, the scalars a, λ, 1 are replaced by the $m \times m$ matrices Γ, Λ, I, (I being the unit matrix), respectively, then the results given above will still be valid if $\Gamma\Lambda = \Lambda\Gamma$, and the reciprocals which occur are replaced by the inverses of the corresponding matrices.

For example, the inverse of A, in block form, will be

$$
-[\Gamma^2 U_{n-2} - 2\Gamma U_{n-3} + U_{n-4}]^{-1}[\Gamma U_{r-2} - U_{r-3}][\Gamma U_{n-s-1} - U_{n-s-2}]
$$

where the argument of the Chebyshev polynomials is now Λ.

Department of Computer Science
University of Edinburgh
Edinburgh 8, Scotland

A reprint from:

Mathematics of Computation

Volume 24, Number 109

January, 1970

Published by the American Mathematical Society

PROVIDENCE, RHODE ISLAND

Inequalities on the Elements of the Inverse of a Certain Tridiagonal Matrix

By D. Kershaw

Abstract. Inequalities are obtained for the elements in the inverse of a tridiagonal matrix with positive off-diagonal elements.

During an investigation into the convergence properties of natural splines it was found useful to have bounds on the inverse of a tridiagonal matrix with positive off-diagonal elements. Matrices of this type arise in other branches of numerical analysis, in particular in the discrete analogue of certain second-order differential operators, and so it may be useful to record these results. The matrix is

$$A = \begin{bmatrix} \lambda_1 & 1-\alpha_1 & 0 & \cdots & 0 & 0 \\ \alpha_2 & \lambda_2 & 1-\alpha_2 & \cdots & 0 & 0 \\ 0 & \alpha_3 & \lambda_3 & \cdots & 0 & 0 \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot & \cdot\cdot \\ 0 & 0 & 0 & \cdots & \lambda_{n-1} & 1-\alpha_{n-1} \\ 0 & 0 & 0 & \cdots & \alpha_n & \lambda_n \end{bmatrix}$$

where $0 < \alpha_r < 1$, $r = 1(1)_n$ and $\lambda_r\lambda_{r+1} > 1$, $r = 1(1)n - 1$.

If the elements of A^{-1} are denoted by

$$a_{rs}^{-1}, \qquad r, s = 1(1)n$$

then the following inequalities hold:

$$1 < a_{ss}^{-1}\lambda_s < \mu_s/(\mu_s - 1), \qquad s = 1(1)n$$

$$0 < (-1)^{r-s}a_{rs}^{-1} \prod_{t=t_1}^{t_2} \lambda_t < \frac{\mu_s}{\mu_s - 1}, \qquad r, s = 1(1)n, \quad r \neq s,$$

where $t_1 = \min(r, s)$, $t_2 = \max(r, s)$, and

$$\mu_s = \min(\lambda_{s-1}\lambda_s, \lambda_s\lambda_{s+1}), \qquad s = 2(1)n - 1,$$

with $\mu_1 = \lambda_1\lambda_2$, $\mu_n = \lambda_{n-1}\lambda_n$.

The proof is elementary and will be indicated only. The last column of A^{-1} is given by the solution of the equations:

$$\lambda_1 x_1 + (1 - \alpha_1)x_2 = 0,$$

$$\alpha_r x_{r-1} + \lambda_r x_r + (1 - \alpha_r)x_{r+1} = 0, \qquad r = 2(1)n - 1, \tag{1}$$

$$\alpha_n x_{n-1} + \lambda_n x_n = 1,$$

Received January 20, 1969, revised March 9, 1969.
AMS Subject Classifications. Primary 1515, 1558; Secondary 6550, 6562, 6580.
Key Words and Phrases. Tridiagonal matrices.

(where $\{a_{rn}^{-1}\}$ has been replaced by $\{x_r\}$ for simplicity). Now x_2 cannot vanish, otherwise recursively from the first $n - 1$ equations it would follow that

$$x_r = 0\,, \qquad r = 1(1)n\,,$$

contradicting the final equation. Hence the first equation can be written

$$-\lambda_1 \frac{x_1}{x_2} = 1 - \alpha_1$$

giving

$$0 < -\lambda_1 \frac{x_1}{x_2} < 1\,.$$

It will now be shown by induction that

$$(2) \qquad 0 < -\lambda_r \frac{x_r}{x_{r+1}} < 1\,, \qquad r = 2(1)n - 1\,.$$

Assume that these inequalities hold for $r = 2(1)p - 1$, so that in particular

$$0 < -\lambda_{p-1} x_{p-1}/x_p < 1\,.$$

Now

$$\alpha_p x_{p-1} + \lambda_p x_p + (1 - \alpha_p) x_{p+1} = 0$$

and, as $x_p \neq 0$, this can be written, after multiplication by λ_{p-1},

$$-\lambda_{p-1}\lambda_p = \alpha_p \lambda_{p-1} x_{p-1}/x_p + (1 - \alpha_p)\lambda_{p-1} x_{p+1}/x_p\,,$$

from which it follows that

$$(3) \quad \min\,(\lambda_{p-1}x_{p-1}/x_p,\ \lambda_{p-1}x_{p+1}/x_p) < -\lambda_{p-1}\lambda_p < \max\,(\lambda_{p-1}x_{p-1}/x_p,\ \lambda_{p-1}x_{p+1}/x_p)\,.$$

Consideration of the inequalities

$$-\lambda_{p-1}\lambda_p < -1\,, \qquad \lambda_{p-1}x_{p-1}/x_p > -1$$

shows that (3) can be more precisely written as

$$\lambda_{p-1}x_{p+1}/x_p < -\lambda_{p-1}\lambda_p < \lambda_{p-1}x_{p-1}/x_p\,.$$

The lower inequality is easily seen to be equivalent to

$$0 < -\lambda_p x_p/x_{p+1} < 1\,,$$

thus completing the proof of (2).

Next consider the last equation of (1) which can be written

$$\alpha_n \lambda_{n-1} x_{n-1}/x_n = -\lambda_{n-1}\lambda_n + \lambda_{n-1}/x_n\,,$$

but as

$$0 < -\alpha_n \lambda_{n-1} x_{n-1}/x_n < \alpha_n < 1$$

it follows that

$$0 < \lambda_{n-1}\lambda_n - \lambda_{n-1}/x_n < 1$$

which can be rewritten, replacing $\lambda_{n-1}\lambda_n$ by μ_n, as

$$1 < \lambda_n x_n < \mu_n/(\mu_n - 1)\,. \tag{4}$$

It is now a simple matter to prove by induction using (2) and (4) that

$$0 < (-1)^{n-r} x_r \lambda_r \lambda_{r+1} \cdots \lambda_n < \mu_n/(\mu_n - 1)\,, \qquad r = n - 1(-1)1\,.$$

For, if this is true when $r = p$, then

$$0 < (-1)^{n-p} x_p \lambda_p \cdots \lambda_n < \mu_n/(\mu_n - 1)\,,$$

but from (2)

$$0 < (-)^{n-p-1} \frac{\lambda_{p-1}\lambda_p \cdots \lambda_n x_{p-1}}{(-1)^{n-p}\lambda_p \cdots \lambda_n x_p} < 1\,,$$

and so

$$0 < (-1)^{n-p-1}\lambda_{p-1}\lambda_p \cdots \lambda_n x_{p-1} < (-1)^{n-p}\lambda_p \cdots \lambda_n x_p < \mu_n/(\mu_n - 1)\,,$$

completing the induction. In an identical fashion it can be shown that the elements in the first column of A^{-1} satisfy

$$1 < a_{11}^{-1}\lambda_1 < \mu_1/(\mu_1 - 1)\,, \quad \text{where } \mu_1 = \lambda_1\lambda_2\,,$$

and

$$0 < (-1)^{r-1} a_{r1}^{-1}\lambda_1\lambda_2 \cdots \lambda_r < \mu_1/(\mu_1 - 1)\,, \qquad r = 2(1)n\,.$$

To prove the inequalities in the general case the following equations for the elements of the sth column of A^{-1} must be considered:

$$\begin{aligned} \lambda_1 x_1 + (1 - \alpha_1) x_2 &= 0\,, \\ \alpha_r x_{r-1} + \lambda_r x_r + (1 - \alpha_r) x_{r+1} &= \delta_{rs}\,, \qquad r = 2(1)n - 1\,, \\ \alpha_n x_{n-1} + \lambda_n x_n &= 0\,. \end{aligned} \tag{5}$$

In order to use the previous line of argument it must be shown that neither x_2 nor X_{n-1} vanish. Now if $X_{n-1} = 0$, then using the last $n - s + 1$ equations of (5) it would follow that

$$x_n = x_{n-1} = \cdots = x_{s+1} = x_s = 0\,, \qquad \alpha_s x_{s-1} = 1\,. \tag{6}$$

If $x_2 = 0$ the first $s + 1$ equations would give the contradictory conclusion

$$x_1 = x_2 = \cdots = x_{s-1} = x_s = 0\,, \qquad (1 - \alpha_s) x_{s+1} = 1\,.$$

Alternatively, if $x_2 \neq 0$, then the argument used to derive (2) could be used again to prove that

$$0 < -\lambda_r x_r / x_{r+1} < 1\,, \qquad r = 1(1)s - 1\,, \tag{7}$$

and the last of these inequalities contradicts (6).

Similarly, the assumption that $x_2 = 0$ will lead to contradictions, and so $x_2 \cdot x_{n-1} \neq 0$. It follows that (7) holds, and also, coming back from the nth equation of (5),

$$0 < -\lambda_{r+1} x_{r+1} / x_r < 1\,, \qquad r = s(1)n\,. \tag{8}$$

In particular, from (7) and (8)

$$0 < -\lambda_{s-1}x_{s-1}/x_s < 1\,, \qquad 0 < -\lambda_{s+1}x_{s+1}/x_s < 1\,,$$

which, as $\lambda_{s-1}\lambda_s > 1$, $\lambda_s\lambda_{s+1} > 1$, are equivalent to

$$(9) \qquad 0 < -x_{s-1}/\lambda_s x_s < 1/\lambda_{s-1}\lambda_s\,, \qquad 0 < -x_{s+1}/\lambda_s x_s < 1/\lambda_s\lambda_{s+1}\,.$$

Now the sth equation of (5) can be rewritten as

$$1 - \frac{1}{\lambda_s x_s} = -\alpha_s \frac{x_{s-1}}{\lambda_s x_s} - (1 - \alpha_s)\frac{x_{s+1}}{\lambda_s x_s}$$

and so

$$\min\,(-x_{s-1}/\lambda_s x_s,\ -x_{s+1}/\lambda_s x_s) < 1 - 1/\lambda_s x_s < \max\,(-x_{s-1}/\lambda_s x_s,\ x_{s+1}/\lambda_s x_s)\,,$$

and, using (9), this implies that

$$(10) \qquad 0 < 1 - 1/\lambda_s x_s < \max\,(1/\lambda_{s-1}\lambda_s,\ 1/\lambda_s\lambda_{s+1})\,.$$

If now

$$\mu_s = \min\,(\lambda_{s-1}\lambda_s,\ \lambda_s\lambda_{s+1})$$

then (10) becomes

$$0 < 1 - 1/\lambda_s x_s < 1/\mu_s\,,$$

from which it follows that

$$(11) \qquad 1 < \lambda_s x_s < \mu_s/(\mu_s - 1)\,.$$

It remains to use (11) to translate the inequalities (7), (8) into inequalities on the elements themselves. This can be done by induction as was indicated in the case when the last column of A^{-1} was considered and need not be described.

(**Note Added in Proof.** The conditions on α_1, α_n can be relaxed to $0 \leqq \alpha_1 < 1$, $0 < \alpha_n \leqq 1$, in which case $1 < a_{ss}^{-1}\lambda_s \leq \mu_s/(\mu_s - 1)$ for $s = 1, n$.)

University of Edinburgh
Edinburgh
Scotland

SIAM J. NUMER. ANAL.
Vol. 8, No. 1, March 1971

A NOTE ON THE CONVERGENCE OF INTERPOLATORY CUBIC SPLINES*

D. KERSHAW†

Abstract. It is shown that if $x \in C^4[a, b]$ is approximated by a natural cubic spline, then the error is $O(h^4)$ in a closed interval which is asymptotic to $[a, b]$ as h, the maximum interval length, decreases to zero. A by-product of the technique used is that if exact end conditions are imposed, then the error is $O(h^4)$ in $[a, b]$.

1. Introduction. The genesis and characterization of cubic splines are described in detail in [1]. For the purposes of this note it will be sufficient to state that a cubic spline with *knots*

$$t_0, t_1, \cdots, t_N,$$

where $a = t_0 < t_1 < \cdots < t_N = b$, $N \geqq 1$, is a member of $C^2(-\infty, \infty)$ and is a cubic polynomial in each interval (t_i, t_{i+1}), $i = 0, 1, \cdots, N-1$. If the cubic spline is to be uniquely determined by its values at the knots, and so be useful for interpolatory purposes, then two further conditions need to be imposed. The conditions which give rise to the *natural cubic spline* are that it should be linear in $(-\infty, a)$, (b, ∞). These are not the only conditions which will make the spline unique, and in this note three other types will be considered. These are the ones which arise when the spline is constrained to have its first or second derivatives at the endpoints $t = a, b$, the same as the respective ones of the function to be interpolated. In addition, the periodic cubic spline will be defined.

The purpose of the note is to find the orders of approximation of these different types of cubic splines to the functions with which they agree at the knots, each when $\max (t_{i+1} - t_i)$ tends to zero. This has been done for the natural spline by Atkinson [2] when the knot spacing is fixed. Although the results for the other types of splines are probably known, they will be dealt with here for completeness.

In the following, y will denote a cubic spline with the knots defined above, and x will be a function with which y agrees at these knots. It will be assumed that $x \in C^4[a, b]$ and that

$$\|x^{(4)}\| \triangleq \max_{a \leqq t \leqq b} |x^{(4)}(t)| = M.$$

In addition,

$$h_i \triangleq t_{i+1} - t_i, \quad i = 0, 1, \cdots, N-1; \qquad h \triangleq \max_{0 \leqq i \leqq N-1} h_i.$$

The main result of this note can be summarized as the following theorem.

THEOREM 1. *If y is a natural cubic spline, $x \in C^4[a, b]$, and*

$$y(t_i) = x(t_i), \qquad i = 0, 1, \cdots, N,$$

* Received by the editors June 3, 1969, and in revised form April 15, 1970.

† Department of Computer Science, University of Edinburgh, Edinburgh 8, Scotland.

then there exist knots t_p, t_q for sufficiently large N, where

$$a < t_p < t_q < b,$$

and a constant K such that for $t_p \leqq t \leqq t_q$,

$$\max |x(t) - y(t)| \leqq Kh^4,$$
$$\max |x^{(1)}(t) - y^{(1)}(t)| \leqq 4Kh^3,$$
$$\max |x^{(2)}(t) - y^{(2)}(t)| \leqq 8Kh^2.$$

Further,

$$t_p - a = O(h \log h), \qquad b - t_q = O(h \log h) \quad \textit{as } h \to 0.$$

For the other types of splines it will be shown that the same orders of approximation hold in the full range $a \leqq t \leqq b$. These results are contained in Theorem 2 below.

2. Interpolating cubic splines. A derivation of the defining equations for the cubic spline can be found in [3], but for completeness and to set the terminology it will be outlined here.

As the spline y is cubic in (t_i, t_{i+1}), it follows that for $t_i \leqq t \leqq t_{i+1}$,

$$(1) \qquad h_i y^{(2)}(t) = (t_{i+1} - t)y^{(2)}(t_i) + (t - t_i)y^{(2)}(t_{i+1}), \qquad i = 0, 1, \cdots, N-1.$$

Let $y^{(2)}(t_i) = \kappa_i$ in (1); integrate twice and impose the interpolation conditions

$$y(t_i) = x(t_i) \triangleq x_i, \qquad i = 0, 1, \cdots, N,$$

to give

$$h_i y(t) = \tfrac{1}{6}[(t_{i+1} - t)^3\kappa_i + (t - t_i)^3\kappa_{i+1}] + (t_{i+1} - t)[x_i - \tfrac{1}{6}h_i^2\kappa_i]$$
$$+ (t - t_i)[x_{i+1} - \tfrac{1}{6}h_i^2\kappa_{i+1}]$$

for $t_i \leqq t \leqq t_{i+1}, i = 0, 1, \cdots, N-1$. Now use the condition that $y^{(1)}(t)$ is continuous at the knots $t_1, t_2, \cdots, t_{N-1}$ to give the well-known result that

$$\tfrac{1}{6}[h_{i-1}\kappa_{i-1} + 2(h_{i-1} + h_i)\kappa_i + h_i\kappa_{i+1}] = \frac{x_{i+1} - x_i}{h_i} - \frac{x_i - x_{i-1}}{h_{i-1}},$$

$i = 1, 2, \cdots, N-1$. Now if

$$\alpha_i \triangleq h_{i-1}/(h_{i-1} + h_i),$$

then these can be written in the form

$$(2) \qquad \alpha_i\kappa_{i-1} + 2\kappa_i + (1 - \alpha_i)\kappa_{i+1} = 6[t_{i-1}, t_i, t_{i+1}]x, \qquad i = 1, 2, \cdots, N-1.$$

(Here $[t_{i-1}, t_i, t_{i+1}]x$ denotes, in Ostrowski's notation, the second divided difference of x at the points t_{i-1}, t_i, t_{i+1}.) Clearly, as there are two more unknowns than equations in (2), restrictions must be imposed to determine a unique cubic spline. Only the following will be considered here.

(a) *Natural cubic spline.* If y is to be linear in $(-\infty, a], [b, \infty)$, then

$$y^{(2)}(t_0) = y^{(2)}(t_N) = 0,$$

whence, in (2),

$$\kappa_0 = \kappa_N = 0.$$

The equations for this in matrix form are then

$$\begin{bmatrix} 2 & 1-\alpha_1 & 0 & \cdots & 0 \\ \alpha_2 & 2 & 1-\alpha_2 & \cdots & 0 \\ 0 & \alpha_3 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 \end{bmatrix} \begin{bmatrix} \kappa_1 \\ \kappa_2 \\ \kappa_3 \\ \vdots \\ \kappa_{N-1} \end{bmatrix} = 6 \begin{bmatrix} [t_0, t_1, t_2]x \\ [t_1, t_2, t_3]x \\ [t_2, t_3, t_4]x \\ \vdots \\ [t_{N-2}, t_{N-1}, t_N]x \end{bmatrix}. \tag{3}$$

(b) D1 *spline.* In this case the first derivatives are fitted at t_0, t_N; that is,

$$y_0^{(1)} = x_0^{(1)}, \qquad y_N^{(1)} = x_N^{(1)}.$$

It is easily shown that these give rise to two equations in addition to (2), namely,

$$\begin{aligned} 2\kappa_0 + \kappa_1 &= 6[x_1 - x_0 - h_0 x_0^{(1)}]/h_0^2, \\ \kappa_{N-1} + 2\kappa_N &= -6[x_N - x_{N-1} - h_{N-1} x_N^{(1)}]/h_{N-1}^2. \end{aligned} \tag{4}$$

The resulting equation can be written

$$\begin{bmatrix} 2 & 1 & 0 & \cdots & 0 & 0 \\ \alpha_1 & 2 & 1-\alpha_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 2 \end{bmatrix} \begin{bmatrix} \kappa_0 \\ \kappa_1 \\ \vdots \\ \kappa_N \end{bmatrix} = 6 \begin{bmatrix} [x_1 - x_0 - h_0 x_0^{(1)}]/h_0^2 \\ [t_0, t_1, t_2]x_1 \\ \vdots \\ -[x_N - x_{N-1} - h_{N-1} x_N^{(1)}]/h_{N-1}^2 \end{bmatrix}. \tag{5}$$

(c) D2 *spline.* If the second derivatives are fitted at $t = t_0, t_N$ then in (2),

$$\kappa_0 = x_0^{(2)}, \qquad \kappa_N = x_N^{(2)},$$

whence for this spline,

$$\begin{aligned} &\begin{bmatrix} 2 & 1-\alpha_1 & 0 & \cdots & 0 \\ \alpha_2 & 2 & 1-\alpha_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 \end{bmatrix} \begin{bmatrix} \kappa_1 \\ \kappa_2 \\ \vdots \\ \kappa_{N-1} \end{bmatrix} \\ &\qquad = 6 \begin{bmatrix} [t_0, t_1, t_2]x - \alpha_1 x_0^{(2)} \\ [t_1, t_2, t_3]x \\ \vdots \\ [t_{N-2}, t_{N-1}, t_N]x - (1 - \alpha_{N-1}) x_N^{(2)} \end{bmatrix}. \end{aligned} \tag{6}$$

(d) *Periodic spline.* The function x is now assumed to have period $b - a$, that is,

$$x(t + b - a) = x(t).$$

Thus the periodic cubic spline will be made to satisfy

$$y^{(r)}(a) = y^{(r)}(b), \qquad r = 0, 1, 2.$$

It is clear that the equations (2) still hold for $i = 1, 2, \cdots, N - 1$, but in the first equation κ_0 can be replaced by κ_N. In addition, the condition that $y^{(1)}$ is continuous at $t = t_N$ has to be imposed. This gives, after some simplification,

$$\beta\kappa_1 + (1 - \beta)\kappa_{N-1} + 2\kappa_N = 6[t_{N-1}, t_N, t_N + h_0]x,$$

where

$$\beta \triangleq h_0/(h_0 + h_{N-1}).$$

The equations now have the form

$$\begin{bmatrix} 2 & 1-\alpha_1 & 0 & 0 & \alpha_1 \\ \alpha_2 & 2 & 1-\alpha_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta & 0 & 0 & 1-\beta & 2 \end{bmatrix} \begin{bmatrix} \kappa_1 \\ \kappa_2 \\ \vdots \\ \kappa_N \end{bmatrix} = 6 \begin{bmatrix} [t_0, t_1, t_2]x \\ [t_1, t_2, t_3]x \\ \vdots \\ [t_{N-1}, t_N, t_N + h_0]x \end{bmatrix}. \tag{7}$$

It will be noticed that each of these square matrices is strictly diagonally dominant, and so each type of spline is determined uniquely by its knot values. Further, they share the property that if A denotes any one of them, then with the uniform matrix norm,

$$\|A^{-1}\| \leqq 1.$$

This is easily proved from the observation:

$$\textit{if } \|A\mathbf{x}\| \geqq m, \quad m > 0, \quad \|\mathbf{x}\| = 1, \quad \textit{then } \|A^{-1}\| \leqq 1/m.$$

Now each matrix can be written as $2I + B$, where $\|B\| \leqq 1$; and so as $\|B\mathbf{x}\| \leqq \|B\| \cdot \|\mathbf{x}\| \leqq 1$, it follows that

$$\|A\mathbf{x}\| = \|2\mathbf{x} + B\mathbf{x}\| \geqq 2\|\mathbf{x}\| - \|B\mathbf{x}\| \geqq 1,$$

which gives the result.

3. Errors in interpolation.

LEMMA 1. *If* $z \in C^2[t', t'']$, $z(t') = z(t'') = 0$, *then, for* $t' \leqq t \leqq t''$,

(i) $$\max|z(t)| \leqq \tfrac{1}{8}(t'' - t')^2 \max|z^{(2)}(t)|,$$

(ii) $$\max|z^{(1)}(t)| \leqq \tfrac{1}{2}(t'' - t') \max|z^{(2)}(t)|.$$

Proof. (i) This follows immediately from Lagrange's interpolation formula.

(ii) It is easily verified that

$$(t'' - t')z^{(1)}(t) = \int_{t'}^{t} (s - t')z^{(2)}(s)\,ds - \int_{t}^{t''} (t'' - s)z^{(2)}(s)\,ds,$$

and so for $t' \leqq t \leqq t''$,

$$(t'' - t')\max|z^{(1)}(t)| \leqq \max|z^{(2)}(t)| \max\left\{\int_{t'}^{t}(s - t')\,ds + \int_{t}^{t''}(t'' - s)\,ds\right\}$$
$$= \tfrac{1}{2}(t'' - t')^2 \max|z^{(2)}(t)|.$$

COROLLARY. *If y is any of the cubic splines, then, for $t_i \leqq t \leqq t_{i+1}$, $i = 0, 1, \cdots, N-1$,*

$$\max|x(t) - y(t)| \leqq \tfrac{1}{8}h_i^2 L_i,$$
$$\max|x^{(1)}(t) - y^{(1)}(t)| \leqq \tfrac{1}{2}h_i L_i,$$
$$\max|x^{(2)}(t) - y^{(2)}(t)| \leqq L_i,$$

where $L_i = \frac{1}{8}h_i^2 M + \max\{|x_i^{(2)} - \kappa_i|, |x_{i+1}^{(2)} - \kappa_{i+1}|\}$.

Proof. From the lemma it is sufficient to prove only the last inequality. As $y^{(2)}$ is linear in $t_i \leqq t \leqq t_{i+1}$, then from Lagrange's linear interpolation formula,

$$h_i[y^{(2)}(t) - x^{(2)}(t)] = (t_{i+1} - t)(\kappa_i - x_i^{(2)}) + (t - t_i)(\kappa_{i+1} - x_{i+1}^{(2)})$$
$$+ \tfrac{1}{2}h_i(t - t_i)(t - t_{i+1})x^{(4)}(\tau),$$

where $t_i < \tau < t_{i+1}$. Hence

$$h_i \max|y^{(2)}(t) - x^{(2)}(t)| \leqq h_i \max\{|x_i^{(2)} - \kappa_i|, |x_{i+1}^{(2)} - \kappa_{i+1}|\}$$
$$+ \tfrac{1}{2}h_i \cdot M \cdot \max|(t - t_i)(t - t_{i+1})|$$

for $t_i \leqq t \leqq t_{i+1}$, from which the result follows.

The inequalities in the corollary are valid for each type of spline, and so it remains only to estimate L_i for each type. To do this (2) will be rewritten as

$$\alpha_i[x_{i-1}^{(2)} - \kappa_{i-1}] + 2[x_i^{(2)} - \kappa_i] + (1 - \alpha_i)[x_{i+1}^{(2)} - \kappa_{i+1}]$$
$$= \alpha_i x_{i-1}^{(2)} + 2x_i^{(2)} + (1 - \alpha_i)x_{i+1}^{(2)} - 6[t_{i-1}, t_i, t_{i+1}]x, \qquad i = 1, 2, \cdots, N-1.$$

A routine use of Peano's method for finding remainders shows that

$$\alpha_i[x_{i-1}^{(2)} - \kappa_{i-1}] + 2[x_i^{(2)} - \kappa_i] + (1 - \alpha_i)[x_{i+1}^{(2)} - \kappa_{i+1}]$$
$$= \tfrac{1}{4}[\alpha_i h_{i-1}^2 + (1 - \alpha_i)h_i^2]x^{(4)}(\sigma_i) \triangleq r_i, \tag{8}$$

where $t_{i-1} \leqq \sigma_i \leqq t_{i+1}$, $i = 1, 2, \cdots, N-1$. For the D1 spline a similar rewriting and finding of remainders is required for (4). The results are easily shown to be

$$2[x_0^{(2)} - \kappa_0] + [x_1^{(2)} - \kappa_1] = 2x_0^{(2)} + x_1^{(2)} - 6[x_1 - x_0 - h_0 x_0^{(1)}]/h_0^2$$
$$= \tfrac{1}{4}h_0^2 x^{(4)}(\sigma_0) = r_0, \qquad t_0 \leqq \sigma_0 \leqq t_1, \tag{9}$$
$$[x_{N-1}^{(2)} - \kappa_{N-1}] + 2[x_N^{(2)} - \kappa_N] = x_{N-1}^{(2)} + 2x_N^{(2)} + 6[x_N - x_{N-1} - h_{N-1}x_N^{(1)}]/h_{N-1}^2$$
$$= \tfrac{1}{4}h_{N-1}^2 x^{(4)}(\sigma_N) = r_N, \qquad t_{N-1} \leqq \sigma_N \leqq t_N.$$

65

4. Orders of convergence. In order to use these last results, the matrix equations (3), (5), (6), (7) will be rewritten so that the vector of unknowns is formed from $x_i^{(2)} - \kappa_i$. That is, if A is a typical matrix, then $A\mathbf{x}^{(2)}$ is subtracted from each side, where $\mathbf{x}^{(2)}$ is a vector of appropriate second derivatives of x at the knots. From the results (8), (9) it is clear that for (5), (6), (7) the modulus of each element in the resulting right-hand side vectors will be bounded by $\frac{1}{4}h^2M$.

It follows, after multiplying by the respective inverses and taking the uniform norm, that

$$\max|x_i^{(2)} - \kappa_i| \leqq \tfrac{1}{4}h^2M$$

for the respective ranges of the suffix. This gives, with the corollary, Theorem 2.

THEOREM 2. *If y is a* D1, D2 *or a periodic spline which agrees with $x \in C^4[a, b]$ at the knots, then for $a \leqq t \leqq b$,*

$$\max|x(t) - y(t)| \leqq \tfrac{3}{64} Mh^4,$$

$$\max|x^{(1)}(t) - y^{(1)}(t)| \leqq \tfrac{3}{16} Mh^3,$$

$$\max|x^{(2)}(t) - y^{(2)}(t)| \leqq \tfrac{3}{8} Mh^2.$$

The result stated in the introduction will now be proved.

Proof of Theorem 1. In the procedure of rewriting (3) so that the vector of the unknowns is formed from $x_i^{(2)} - \kappa_i, i = 1, 2, \cdots, N - 1$, the resulting matrix equation has the form

$$A(\mathbf{x}^{(2)} - \boldsymbol{\kappa}) = \mathbf{r} - \alpha_1 x_0^{(2)}\mathbf{e}_1 - (1 - \alpha_{N-1})x_N^{(2)}\mathbf{e}_{N-1},$$

where

$$\mathbf{x}^{(2)} - \boldsymbol{\kappa} = [x_1^{(2)} - \kappa_1 \quad x_2^{(2)} - \kappa_2 \cdots x_{N-1}^{(2)} - \kappa_{N-1}]^T, \qquad \mathbf{r} = [r_1, r_2, \cdots, r_{N-1}]^T,$$

and $\mathbf{e}_1, \mathbf{e}_{N-1}$ are the first and last columns of the $N - 1 \times N - 1$ unit matrix. It follows that

$$\mathbf{x}^{(2)} - \boldsymbol{\kappa} = A^{-1}\mathbf{r} - \alpha_1 x_0^{(2)}[A^{-1}\mathbf{e}_1] - (1 - \alpha_{N-1})x_N^{(2)}[A^{-1}\mathbf{e}_{N-1}],$$

and so, for $i = 1, 2, \cdots, N - 1$, with the uniform matrix norm,

$$(10) \quad |x_i^{(2)} - \kappa_i| \leqq \|A^{-1}\mathbf{r}\| + \alpha_1|x_0^{(2)}|\cdot|[A^{-1}\mathbf{e}_1]_i| + (1 - \alpha_{N-1})|x_N^{(2)}|\cdot|[A^{-1}\mathbf{e}_{N-1}]_i|,$$

where $[\cdot]_i$ denotes the ith element of the vector inside the brackets. If $x_0^{(2)}, x_N^{(2)}$ do not vanish (otherwise the natural spline would also be a D2 spline), then estimates of $|[A^{-1}\mathbf{e}_1]_i|, |[A^{-1}\mathbf{e}_{N-1}]_i|$ are required. These can be found from [5], whence it is easily deduced that

$$(11) \qquad |[A^{-1}\mathbf{e}_1]_i| \leqq \tfrac{4}{3}\cdot 2^{-i}, \quad |[A^{-1}\mathbf{e}_{N-1}]_i| \leqq \tfrac{4}{3}\cdot 2^{i-N}, \qquad i = 1, 2, \cdots, N - 1.$$

Thus (10) becomes, after using the result that $\|A^{-1}\mathbf{r}\| \leqq \|\mathbf{r}\| \leqq \frac{1}{4}Mh^2$,

$$(12) \qquad |x_i^{(2)} - \kappa_i| \leqq \tfrac{1}{4}Mh^2 + \tfrac{4}{3}[\alpha_1|x_0^{(2)}|\cdot 2^{-i} + (1 - \alpha_{N-1})|x_N^{(2)}|\cdot 2^{N-i}],$$

$i = 1, 2, \cdots, N - 1$. For convenience,

$$C \triangleq \tfrac{4}{3}\alpha_1|x_0^{(2)}|, \qquad D \triangleq \tfrac{4}{3}(1 - \alpha_{N-1})|x_N^{(2)}|;$$

then, from the corollary to the lemma, for $t_i \leqq t \leqq t_{i+1}$, $i = 1, 2, \cdots, N-2$,

$$\max|x^{(2)}(t) - y^{(2)}(t)| \leqq \tfrac{3}{8}Mh^2 + \max\{[C \cdot 2^{-i} + D \cdot 2^{i-N}],\ [C \cdot 2^{-i-1} + D \cdot 2^{-N+i+1}]\}.$$

Now $C \cdot 2^{-t} + D \cdot 2^{t-N}$ is convex towards the t-axis and so for $t_p \leqq t \leqq t_q$, $1 \leqq p < q \leqq N-1$,

$$\max|x^{(2)}(t) - y^{(2)}(t)| \leqq \tfrac{3}{8}Mh^2 + \max\{[C \cdot 2^{-p} + D \cdot 2^{p-N}], [C \cdot 2^{-q} + D \cdot 2^{q-N}]\},$$

which, as $p < q$, can be replaced by

$$\max|x^{(2)}(t) - y^{(2)}(t)| \leqq \tfrac{3}{8}Mh^2 + C \cdot 2^{-p} + D \cdot 2^{q-N}, \qquad t_p \leqq t \leqq t_q. \tag{13}$$

It remains to be shown that for large enough N, the integers p, q can be chosen so that

$$C \cdot 2^{-p} \leqq h^2, \qquad D \cdot 2^{q-N} \leqq h^2.$$

For these equalities to hold,

$$p \geqq (\log Ch^{-2})/\log 2, \qquad q \leqq N - (\log Dh^{-2})/\log 2,$$

and so p, q can be defined by

$$p = 1 + \max\{0, (\log Ch^{-2})/\log 2\}, \quad q = \min\{N-1, N - (\log Dh^{-2})/\log 2\}, \tag{14}$$

where $[z]$ here denotes the integral part of z.

Finally in order that $p < q$, it is sufficient that

$$N - 2 \geqq [\log(Ch^{-2}) + \log(Dh^{-2})]/\log 2;$$

that is,

$$2^{N-2} \geqq CDh^{-4}, \tag{15}$$

which, as $h \geqq (b-a)/N$, will be satisfied for large enough N. Hence if (15) holds, an interval (t_p, t_q) can be chosen so that in it

$$\max|x^{(2)}(t) - y^{(2)}(t)| \leqq \tfrac{3}{8}Mh^2 + 2h^2 \triangleq 8Kh^2,$$

which, with the corollary to the lemma, gives the inequalities in the statement of Theorem 1.

Further, from (14) (the trivial case where $p = 1, q = N-1$ will not be considered),

$$(\log Ch^{-2})/\log 2 < p \leqq 1 + (\log Ch^{-2})/\log 2,$$

$$N - 1 - (\log Dh^{-2})/\log 2 \leqq q < N - (\log Dh^{-2})/\log 2,$$

and so

$$t_p - a = t_p - t_0 \leqq ph \leqq h[1 + (\log Ch^{-2})/\log 2] = O(h \log h)$$

and

$$b - tq = t_N - t_q \leqq (N - q)h \leqq h[1 + (\log Dh^{-2})/\log 2] = O(h \log h).$$

This completes the proof of Theorem 1.

5. Conclusions. The results show that although the convergence of the interpolating cubic spline is not uniformly $O(h^4)$ as is the case when exact endpoint conditions are fitted, for a sufficiently large number of points it is $O(h^4)$ except in two end intervals, which tend to zero as the maximum interval tends to zero.

A result worth noting comes from (12), where it is clear that if α_0 and $1 - \alpha_{N-1}$ are each $O(h^2)$, then the order of approximation is immediately $O(h^4)$. This implies that only the first and last intervals need to be small compared with the maximum interval length.

It is also worth remarking that if equal intervals were used, then the estimates (11) could be replaced by (see [5])

$$|[A^{-1}\mathbf{e}_1]_i| = 2U_{N-i-1}(2)/U_{N-1}(2) < 2(2 - \sqrt{3})^i,$$
$$|[A^{-1}\mathbf{e}_{N-1}]_i| = 2U_{i-1}(2)/U_{N-1}(2) < 2(2 - \sqrt{3})^{N-i},$$

where U_i is the ith Chebyshev polynomial of the second kind.

Finally, it will be noticed that the above approach can be used to investigate the convergence of splines with a mixture of end conditions. For example, if $y_0^{(2)} = x_0^{(2)}$ and $y_N^{(1)} = x_N^{(1)}$, it is obvious that the approximation would be $O(h^4)$. If one end is natural and the other with a fitted derivative, the situation is easily dealt with by using (11).

REFERENCES

[1] J. H. Ahlberg, E. N. Nilson and J. L. Walsh, *The Theory of Splines and their Applications*, Academic Press, New York, 1967.

[2] K. Atkinson, *On the order of convergence of natural cubic spline interpolation*, this Journal, 5 (1968), pp. 89–101.

[3] T. N. E. Greville, *Spline functions*, Mathematical Methods for Digital Computers, vol. 2, A. Ralston and H. S. Wilf, eds., John Wiley, New York, 1967, Chap. 8.

[4] D. Kershaw, *The explicit inverse of two commonly occurring matrices*, Math. Comp., 23 (1969), pp. 189–191.

[5] ———, *Inequalities on the elements of the inverse of a certain tridiagonal matrix*, Ibid., 24 (1970), pp. 155–158.