SHAPE PRESERVING PIECEWISE RATIONAL INTERPOLATION

by

ROGER DELBOURGO

Thesis submitted for the Degree of Doctor of Philosophy

Brunel University

Department of Mathematics and Statistics,

Brunel University.

July 1984

BEST COPY

AVAILABLE

Variable print quality

CONTENTS

Page

ABSTRACT

This thesis considers rational spline methods for solving shape preserving interpolation problems. Data (x_i, f_i), $i=1,\ldots,n$, is given which satisfies monotonic and/or convex constraints. A function $s \in C^1[x_1, x_n]$ is then sought, piecewise defined on the partition $x_1 < x_2 < \ldots < x_n$, which preserves the shape of the data and is such that $s(x_i)=f_i$, $s^1(x_i)=d_i$, $i=1,\ldots,n$. The derivative parameters must satisfy necessary conditions imposed by the shape constraints.

The thesis begins with a review of existing methods which attempt to solve the shape-preserving interpolation problem. In the remainder of the thesis, new methods are studied based on the use of piecewise defined rational functions. It is shown that a piecewise rational quadratic function solves the monotonic interpolation problem. An accurate error bound can be expected for a good choice of the derivative parameters. This choice is considered in detail and includes a C^2 rational spline approach. Here, second derivative continuity constraints give rise to a system of non-linear equations, the solution and properties of which are studied in detail.

A generalisation of the piecewise rational quadratic form to cubic form is then considered. The cubic, in each interval $[x_i, x_{i+1}]$, depends on a parameter r_i. It degenerates to the rational quadratic for one choice of r_i. It is shown that another choice of r_i gives an accurate solution to the convex interpolation problem. Finally, the C^2 rational cubic spline theory corresponding to this convex case is studied.

ACKNOWLEDGEMENT

I wish to record my gratitude to Dr.J.A.Gregory, for his guidance and constant encouragement during the period which spanned these investigations. Dr. Gregory was first to suggest to me ideas on rational interpolation, and the pages herein contain a development of these ideas.

NOTATION

The following notation will be adopted and adhered to throughout the present work. (See Figure 0.1)

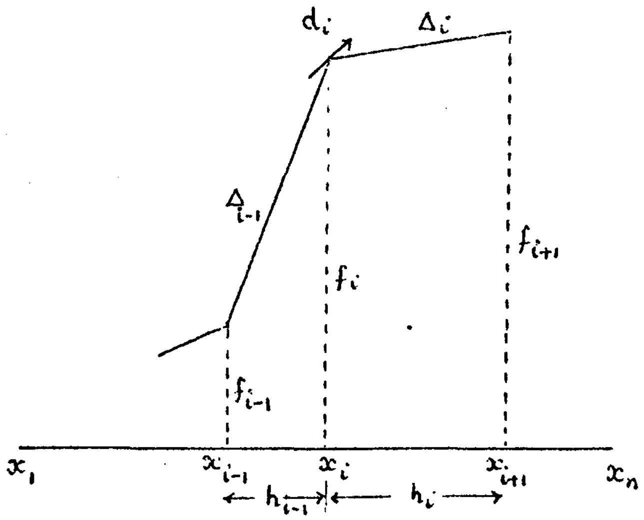


Fig.0.1

$\pi = \{x_1, x_2, \ldots, x_n\}$	: a partition of $I = [x_1, x_n] \subset R$ i.e. $x_1 < x_2 < \ldots < x_n$
$I_i = [x_i, x_{i+1}]$	: i th subinterval of I
f_i	: data value at x_i, (x_i, f_i) is a data point $(i=1,\ldots,n)$
$h_i = x_{i+1} - x_i$	: length of I_i
$h = \max\{h_i\}$	
$\Delta_i = \Delta_{i,i+1} = (f_{i+1} - f_i)/h_i$	: chord slope joining points 'i' and 'i+1'
$\Delta_{i,j} = \Delta_{j,i} = (f_j - f_i)/(x_j - x_i)$	: chord slope joining points 'i' and 'j'
$C^r[x_1, x_n]$	: class of functions with continuous r th derivatives on I.
$s: I \to R$	: interpolant under consideration, with $s(x_i) = f_i$ $(i=1,\ldots,n)$
$d_i = s^{(1)}(x_i)$.	: derivative of s at x_i, i.e. $s'(x_i)$

$\theta = (x-x_i)/h_i$	: for $x \in I_i$; $0 \leq \theta \leq 1$
$f: I \rightarrow R$	: a function of given continuity class
$f_i^{(1)}, f_i^{(2)}$	: first, second derivatives of f at x_i, i.e. $f'(x_i), f''(x_i)$.
$\lambda_i = d_i - f_i^{(1)}$	
$\|F\| = \|F\|_\infty$	: the uniform norm of F, i.e. $\max\{\|F(x)\|: x \in I\}$.

Remark

Attention should be drawn especially to the convention regarding the labelling of data points. The first and last are (x_1, f_1), (x_n, f_n). There are $(n-1)$ subintervals, of lengths $h_1, \ldots, h_{n-1}$, supporting chord slopes $\Delta_1, \ldots, \Delta_{n-1}$. An interpolant to the data has end slopes d_1, d_n.

Chapter 1

INTRODUCTION

1.1 Introduction to Rational Interpolation

The subject of these investigations is the determination of interpolants which preserve the monotonic and/or convex shape of a given set of data. This will be achieved by means of piecewise defined rational functions (that is, ratios of polynomials). This section introduces the reader to that notion.

Data will be supplied in the form of a set of points (x_i, f_i), $i=1,2,\ldots,n$, with $n \geq 3$ and $x_1 < x_2 < \ldots < x_n$. The values f_i are assumed to be not prone to error, even if they are the result of experiment, and the data will be

either (a) monotonic increasing (decreasing)

or (b) convex (concave)

or both. A definition of these terms is given below.

On the closed interval $[x_1, x_n]$, a function s will be defined in the following manner:

(i) In an arbitrary subinterval $[x_i, x_{i+1}]$, $s(x)$ is a rational function of the form specified in (iii) below with

$s(x_i) = f_i$, $s(x_{i+1}) = f_{i+1}$.

(ii) $s^{(1)}$ is continuous, and $s^{(1)}(x_i) = d_i$, $s^{(1)}(x_{i+1}) = d_{i+1}$, where d_i, d_{i+1} are derivative values to be estimated.

Thus, the function $s \in C^1[x_1, x_n]$; it consists of a union of rational polynomial pieces, it interpolates the data and at the points x_i its derivative takes the values d_i. In a given problem, these derivatives d_i can be estimated by a variety of explicit finite difference formulae, but other settings for them may be achieved by implicit methods, which will include iteration.

(iii) In a representative interval $[x_i, x_{i+1}]$, the form of s will be taken as

$$s(x) = \frac{\sum_{k=0}^{N} a_k^{(i)} (x_{i+1}-x)^{N-k} (x-x_i)^k}{\sum_{k=0}^{N} b_k^{(i)} (x_{i+1}-x)^{N-k} (x-x_i)^k} ,$$

where N= 2, N= 3 for rational quadratic, rational cubic interpolation, respectively.

The coefficients $a_k^{(i)}$, $b_k^{(i)}$ are non-zero and depend on the i th interval. The denominator in s(x) is not allowed to assume zero or negative values. For a symmetrical form of s, we shall set $b_o^{(i)} = b_N^{(i)} = 1$. Then the interpolation problem requires that $a_o^{(i)} = f_i$, $a_N^{(i)} = f_{i+1}$. Remaining coefficients are determined from further conditions on s, for instance from the requirements (ii) above.

Consider the case of a rational quadratic interpolant (N=2).

In $[x_i, x_{i+1}]$,

$$s(x) = \frac{f_i (x_{i+1}-x)^2 + a_1^{(i)} (x_{i+1}-x)(x-x_i) + f_{i+1}(x-x_i)^2}{(x_{i+1}-x)^2 + b_1^{(i)} (x_{i+1}-x)(x-x_i) + (x-x_i)^2} .$$

We make the substitutions

$$x = x_i + \theta h_i \quad , \; h_i = x_{i+1} - x_i .$$

Then

$$s(x) = s(x_i + \theta h_i) = S(\theta) \; , \text{ where}$$

$$S(\theta) = \frac{f_i (1-\theta)^2 + a_1^{(i)} \theta(1-\theta) + f_{i+1}\theta^2}{(1-\theta)^2 + b_1^{(i)} \theta(1-\theta) + \theta^2} \; ; \quad 0 \leq \theta \leq 1 .$$

$b_1^{(i)} \geq 0$ is sufficient to ensure that the denominator remains of constant (positive) sign. In calculations, it is simpler to use the representation $S(\theta)$. Thus, since $s^{(1)}$ and $S^{(1)}$ are related by $s^{(1)}(x) = h_i^{-1} S^{(1)}(\theta)$, we have $d_i = h_i^{-1} S^{(1)}(0)$, and $d_{i+1} = h_i^{-1} S^{(1)}(1)$.

In treating rational cubic interpolation ($N=3$) similar considerations will apply.

Parts of this thesis have appeared in the literature.(References [7], [14]) Other parts have been/will be submitted for publication (References [8], [9], [10]). We will show in this thesis that rational quadratic interpolation is appropriate in general for data which is monotonic. We will also show that the rational quadratic is a special case of rational cubic interpolation and this generalisation allows us to deal with convex data.

In Chapter 2, we will review some of the existing methods of interpolation. These do not use rational functions, and most are concerned with the construction of shape-preserving interpolants using quadratic and cubic polynomials.

We end this section by defining the terms 'monotonic' and 'convex':

Definition 1.1.1

The data (x_i, f_i), $i=1,\ldots,n$ is strictly monotonic increasing if and only if

$$f_1 < f_2 < \ldots < f_n$$

(equivalently, $\Delta_1, \Delta_2, \ldots, \Delta_{n-1} > 0$, where $\Delta_i = (f_{i+1} - f_i)/h_i$).

Definition 1.1.2

The data (x_i, f_i), $i=1,\ldots,n$ is strictly convex if and only if

$$\Delta_1 < \Delta_2 < \ldots < \Delta_{n-1} \quad .$$

Definitions of 'strictly monotonic decreasing' and strictly concave' can be stated by reversing the inequality signs. The qualification 'strictly' will not apply in either definition if '<' is replaced by '$\leq$' in the sequence of inequalities.

1.2 Monotonic and Convex DATA sets used

The DATA sets employed in our experiments on rational interpolation are listed below. Most are practical sets considered by a number of authors. They are all either monotonic increasing or convex. In fact, there is no essential (theoretical) difference between increasing and decreasing data, or between convex and concave data.

The sets will be subdivided into three categories; those which are
- (a) monotonic, but not convex: labelled M
- (b) monotonic and convex: labelled MC
- and (c) convex but not monotonic: labelled C.

(M1) [AKIMA,'70] [FRITSCH & CARLSON, '80]

x	0	2	3	5	6	8	9	11	12	14	15
f	10	10	10	10	10	10	10.5	15	50	60	85

(M2) [BURDEN, FAIRES & REYNOLDS,'78]

x (Year)	1920	1930	1940	1950	1960	1970
f (Popul'n of U.S.A. in millions)	105.711	123.203	131.669	150.697	179.323	203.212

(M3) [PRUESS,'79]

x	22	22.5	22.6	22.7	22.8	22.9	23	23.1	23.2	23.3	23.4	23.5	24
f	523	543	550	557	565	575	590	620	860	915	944	958	986

(M4) [FRITSCH & CARLSON,'80]

x	7.99	8.09	8.19	8.7	9.2	10	12	15	20
f	0	.0000276429	.0437498	.169183	.469428	.943740	.998636	.999919	.99999

(M5) The Normal distribution function

$$\phi(x)=\frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{x} e^{-t^2/2}dt$$

x	-4	-3	-2	-1	0	1	2	3	4
$\phi(x)$	.00003	.00135	.0228	.1587	.5	.8413	.9772	.99865	.99997

(MC1) McALLISTER, PASSOW & ROULIER, '77

x	-2	-1	-.3	-.2
$f(x)= 1/x^2$	.25	1	$11.\dot{1}$	25

(MC2) Exponential Test function

$f(x) = \exp x$; domain $[0,1]$.

$d_1 = 1$, $d_n = e$. Equal intervals, h.

(i) h= .2 (n=6) (ii) h= .1 (n=11)

(iii) h= .05 (n=21) (iv) h= .025 (n=41).

(MC3) Quarter Circle

Seven points (x_r, f_r), equally spaced over a quarter circle:

$x_r = \sin\{\frac{\pi}{12}(r-1)\}$, $f_r = 1- \cos\{\frac{\pi}{12}(r-1)\}$; r=1,2,...,7.

(C1) Half Circle

Thirteen points (x_r, f_r) equally spaced over a half circle:

$x_r = -\cos\{\frac{\pi}{12}(r-1)\}$, $f_r = 1 - \sin\{\frac{\pi}{12}(r-1)\}$;r=1,2,...,13.

Chapter 2

A REVIEW OF SOME INTERPOLATION METHODS

Many situations in practice present interpolation problems to be solved. For these, it is desirable to obtain interpolation functions which reflect global properties of the given set of interpolation points. Often, for instance, there are demands that monotonic or/and convex data produce also a monotonic or/and convex interpolant. For most interpolation methods these demands can be fulfilled only in special cases. As an example, when cubic interpolation spline schemes are used, it is well-known that unexpected oscillations may be produced in specific cases. There is thus a special need for shape preserving interpolation methods.

Shape preserving spline interpolation methods have been proposed by several authors recently. See references [13], [15], [17], [19], [20], [21], in particular. The methods were devised to produce interpolatory splines preserving properties such as monotonicity or convexity that are present in the data. Some of the methods solve the problem by adding, where necessary, spline knots between interpolation points in such a way that the number of free parameters are enough to guarantee the existence of the solution. However, the exact positions of the additional knots are unknown a priori.

In this section we review a sèlection of interpolation methods which attempt to preserve the shape of the data. If further details on these are required, they may be found in the original references.

To facilitate comparisons between different methods we must agree on a common notation, and we shall use the one we adopted at the outset.

2.1 Akima's method

This method (see [1]) can be applied to arbitrary data (x_i, f_i). The interpolant is constructed using piecewise defined cubics. At the knots x_i, the derivatives d_i are calculated explicitly, and are set as follows:

At the end points x_1, x_n,

$$d_1 = \Delta_1 + h_1(\Delta_1 - \Delta_2)/(h_1+h_2) \ ,$$
$$d_n = \Delta_{n-1} + h_{n-1}(\Delta_{n-1} - \Delta_{n-2})/(h_{n-1}+h_{n-2}) \ .$$

At the 'next-to-end' points x_2, x_{n-1},

$$d_2 = (h_1\Delta_2 + h_2\Delta_1)/(h_1+h_2) \ ,$$
$$d_{n-1} = (h_{n-1}\Delta_{n-2} + h_{n-2}\Delta_{n-1})/(h_{n-1}+h_{n-2}) \ .$$

At the interior points corresponding to $i=3,\ldots,n-2$,

$$d_i = \begin{cases} \Delta_{i-1} \ , & \text{if } \Delta_{i-1} = \Delta_i \\ \tfrac{1}{2}(\Delta_{i-1} + \Delta_i), & \text{if } \Delta_{i-2} = \Delta_{i-1} \text{ and } \Delta_i = \Delta_{i+1} \\ \dfrac{|\Delta_{i+1}-\Delta_i|\cdot\Delta_{i-1} + |\Delta_{i-2}-\Delta_{i-1}|\cdot\Delta_i}{|\Delta_{i+1}-\Delta_i| + |\Delta_{i-2}-\Delta_{i-1}|} \ , & \text{otherwise} \end{cases}$$

We recognise the formulae for d_1, d_2, d_{n-1}, d_n as ordinary 3-point difference approximations. The formulae for d_i, $i=3,\ldots,n-2$ are motivated by consideration of a series of diagrams which illustrate different possibilities that might arise in the data. Expressions for d_i are weighted averages ensuring that $\min\{\Delta_{i-1},\Delta_i\} \leq d_i \leq \max\{\Delta_{i-1},\Delta_i\}$.

Once the derivatives d_i are set, the Hermite cubic polynomials can be constructed. On an arbitrary interval $[x_i, x_{i+1}]$,$(i=1,\ldots n-1)$ the cubic is

$$s(x) = f_i + d_i(x-x_i) + h_i^{-1}(3\Delta_i - 2d_i - d_{i+1})(x-x_i)^2 + h_i^{-2}(d_i + d_{i+1} - 2\Delta_i)(x-x_i)^3.$$

The resulting interpolant is usually a 'tight' curve and unwanted inflexions are rarely produced. However, we must remark that the method is not specifically designed to reflect properties of monotonicity or convexity present in the data. It is in the nature of the method that the accuracy of the interpolant cannot be stated. Further, second derivative discontinuities may sometimes be large.

2.2 Schweikert's tension spline

This spline (see [20]) overcomes the problem created by the production of unnecessary inflexion points. The spline method, or some of its later variants, can be applied to monotonic data to produce a monotonic interpolant.

The tension spline, $s(x)$, satisfies $s(x_i)=f_i, i=1,\ldots,n$ in addition to the conditions that

$$s^{(2)}(x)-\sigma^2 s(x)$$

is continuous in $[x_1,x_n]$ and linear on each interval $[x_i,x_{i+1}]$, $i=1,\ldots,n-1$. The factor σ is a non-zero tension parameter. The larger σ is taken the 'tighter' will be the fit to the data and as $\sigma\to\infty$ $s(x)$ approaches a piecewise linear fit.

When the relevant differential equations are solved we obtain, for $x\in[x_i,x_{i+1}]$,

$$s(x)=\frac{s^{(2)}(x_i)}{\sigma^2}\cdot\frac{\sinh\{\sigma(x_{i+1}-x)\}}{\sinh \sigma h_i}+\frac{s^{(2)}(x_{i+1})}{\sigma^2}\cdot\frac{\sinh\{\sigma(x-x_i)\}}{\sinh \sigma h_i}$$

$$+\left(f_i-\frac{s^{(2)}(x_i)}{\sigma^2}\right)\frac{x_{i+1}-x}{h_i}+\left(f_{i+1}-\frac{s^{(2)}(x_{i+1})}{\sigma^2}\right)\frac{x-x_i}{h_i}$$

The unknown second derivatives $s^{(2)}(x_i)$ are obtained by differentiating this equation and matching $s^{(1)}(x)$ at the end points of intervals. This leads to a tridiagonal system of linear equations to be solved:

$$e_{i-1}s^{(2)}(x_{i-1})+(c_{i-1}+c_i)s^{(2)}(x_i)+e_is^{(2)}(x_{i+1})=\Delta_i-\Delta_{i-1} \quad ,$$

$$i=2,\ldots,n-1 \quad ,$$

where
$$e_i = \frac{1}{\sigma}\left(\frac{1}{\sigma h_i} - \frac{1}{\sinh \sigma h_i} \right) \quad ,$$

$$c_i = \frac{1}{\sigma}\left(\frac{\cosh \sigma h_i}{\sinh \sigma h_i} - \frac{1}{\sigma h_i} \right) .$$

If the 3-point quadratic approximations for the end derivatives are used (as, for instance, in Akima's method), then as end equations we can use

$$c_1s^{(2)}(x_1) + e_1s^{(2)}(x_2) = h_1(\Delta_2-\Delta_1)/(h_1+h_2) \quad ,$$

$$e_{n-1}s^{(2)}(x_{n-1}) + c_{n-1}s^{(2)}(x_n) = h_{n-1}(\Delta_{n-1}-\Delta_{n-2})/(h_{n-1}+h_{n-2}) .$$

The complete tridiagonal system for the $s^{(2)}(x_i), i=1,\ldots,n$ can then be solved in the standard manner.

It follows that the form of $s(x)$ in any interval is then known and $s(x)$ can be constructed. We note that, because of the involvement of hyperbolic functions, the computational simplicity of cubic polynomials is lost. Also we must note that the results depend on the choice of σ.

2.3 Fritsch, Carlson and Butland methods

Fritsch and Carlson, in their 1980 paper (see [13]), consider monotonic increasing data and examine the conditions under which an interpolant constructed using piecewise defined cubics can also be monotonic increasing. These conditions involve choosing the derivatives d_i so that certain inequalities are satisfied.

The Fritsch and Carlson analysis is made by a study of various cases. The quadratic and linear expressions for $s^{(1)}(x)$ and $s^{(2)}(x)$, respectively, in a representative interval $[x_i,x_{i+1}]$ lead to the necessary and sufficient conditions.

We assume $\Delta_i > 0$ in $[x_i,x_{i+1}]$. Otherwise, if $\Delta_i = 0$, we must set

$d_i = d_{i+1} = 0$ at the ends of the interval.

Necessary conditions for a monotonic interpolant in $[x_i, x_{i+1}]$ are $d_i \geq 0$, $d_{i+1} \geq 0$.

Sufficient conditions are best described using the variables

$$\alpha_i = d_i/\Delta_i \ , \ \beta_i = d_{i+1}/\Delta_i \ .$$

A detailed analysis shows that these are separable into 3 sets of possibilities:

(a) $\alpha_i + \beta_i - 2 \leq 0$

(b) $\alpha_i + \beta_i - 2 > 0$ and $(2\alpha_i + \beta_i - 3 < 0$ or $\alpha_i + 2\beta_i - 3 < 0)$

(c) $\alpha_i + \beta_i - 2 > 0$ and $(2\alpha_i + \beta_i - 3 > 0$ and $\alpha_i + 2\beta_i - 3 > 0$ and $\alpha_i^2 + \alpha_i\beta_i + \beta_i^2 - 6\alpha_i - 6\beta_i + 9 \leq 0)$.

These sets of inequalities have a simple pictorial interpretation. An α-β diagram is drawn as in Figure 2.3.1. We see that the pair (α_i, β_i) must lie within the shaded region M if the interpolant in the interval $[x_i, x_{i+1}]$ is to be monotonic.

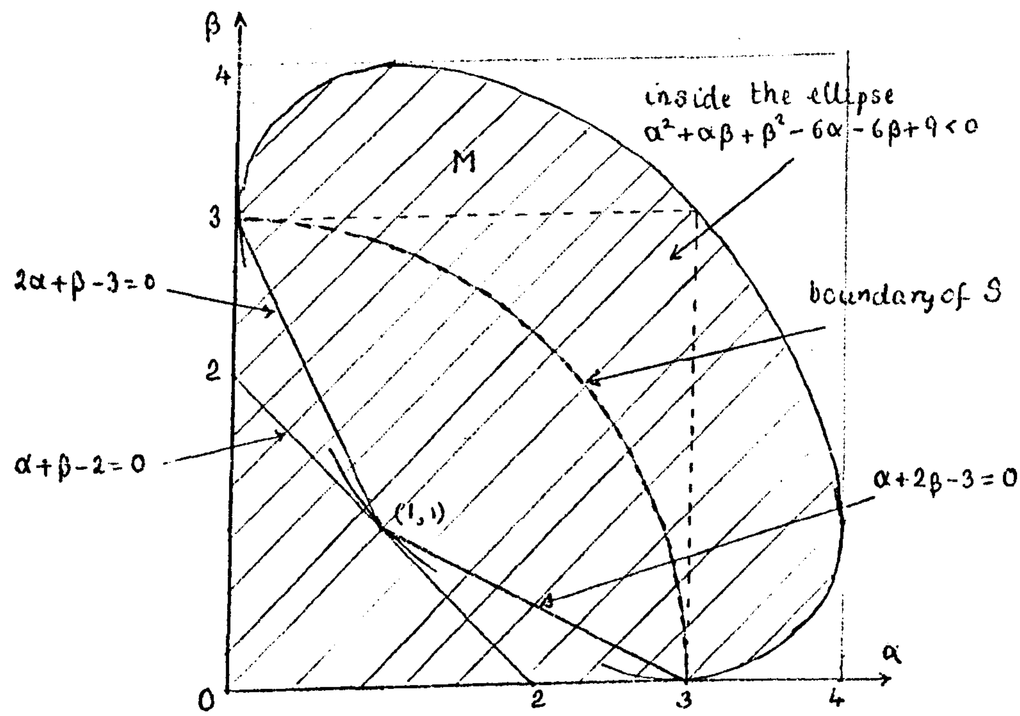


Fig.2.3.1.

Each interval gives rise to a pair (α_i, β_i), but we note that consecutive pairs are related, since $\alpha_{i+1} = \beta_i(\Delta_i/\Delta_{i+1})$.

By reducing systematically all the derivatives $d_i (i=1,\ldots,n)$ it is clear that all points (α_i, β_i) will enter the monotonicity region M. In one of the schemes, the authors suggest working with the subregion of M given by

$$S = \{(\alpha_i, \beta_i) : \alpha_i, \beta_i \geq 0,\ (\alpha_i^2 + \beta_i^2)^{\frac{1}{2}} \leq 3\}.$$

After an initialisation of values, the final pairs (α_i, β_i) are adjusted so as to fall in S, or on its boundary.
Another suggestion still isto use the square subregion of M defined by

$$\{(\alpha_i, \beta_i) : 0 \leq \alpha_i < 3\ ,\ \ 0 \leq \beta_i < 3\}$$

in place of S.
Butland's modification of this procedure (see [5]) is to set the d_i directly. At interior points i= 2,...,n-1 the values are taken as

$$d_i = \frac{1}{\dfrac{\alpha}{\Delta_{i-1}} + \dfrac{1-\alpha}{\Delta_i}} \qquad (\Delta_{i-1}, \Delta_i > 0)$$

where $\alpha = \frac{1}{3}\left(1 + \dfrac{h_i}{h_{i-1}+h_i}\right)$.
We note that, since $1/3 < \alpha < 2/3$, we will have $0 \leq d_i/\Delta_{i-1} < 3$ and $0 \leq d_i/\Delta_i < 3$, and so monotonicity will be guaranteed.
The Fritsch-Carlson-Butland formula, to be found in [12] , is similar, except that α is treated as a parameter with $1/3 \leq \alpha \leq 1$.
The value $\alpha = 1/3$ seems to give good interpolants.
Also, the value $\alpha = 1/2$ in the case of equal intervals gives Butland's settings.

Remark

The Fritsch-Carlson procedure which employs subregion S will be used to test some of the data, and the graphs obtained can then be compared with those given by our piecewise rational schemes. The first comparisons are made in Chapter 3, but at the end of the present chapter we show for the data the Fritsch-Carlson curves, which we take as a reference in all our comparisons.

2.4 Ellis and McLain's interpolatory cubics

The method proposed in Ellis & McLain [11] is a local one, similar to Akima's in its object, but offering some practical advantages because it is algebraic and not designed from considerations of geometry.

The authors show how to construct a C^1 interpolant through the data using piecewise defined cubics. A gradient at an interior point x_i, $i=3,\ldots,n-2$, is calculated as follows. The cubic polynomial through $(x_{i-1},f_{i-1}),(x_i,f_i),(x_{i+1},f_{i+1})$ which gives a least squares fit to the neighbouring points (x_{i-2},f_{i-2}), (x_{i+2},f_{i+2}) is taken. The weights given to the last two points are not chosen equal, but are taken inversely proportional to the squares of the distances from x_{i-1} and x_{i+1}. Thus, much smaller discontinuities in the second derivatives arise at the data points. The gradients at x_1,x_2 and at x_{n-1},x_n are calculated simply by passing two cubics, one through the first four points, the other through the last four. If a cubic polynomial f(x) passes through all the data, the method would produce an exact fit. In any case, on the data actually tested, the method produced much smaller second derivative discontinuities than those obtained by the Akima method.

The authors present an algorithm for the computation of the coefficients in their interpolant. They test the method on five mathematical functions ($x^3,x^4,e^{-x^2/2}$, tanh x, sin x), for 18 unequally spaced values of x ranging from -2.95 to 3.0. No graphs are given, but the authors indicate from tables that their algorithm performs well. One table makes a comparison between average deviations, for various interpolation methods. The other compares the average second derivative discontinuity at data points. The method was found to be useful also in two-dimensional problems of interpolation. We must note that the method does not necessarily preserve the shape of the data.

2.5 Mettke's method for convex cubic Hermite-spline interpolation

Mettke [17] presents a method of constructing convex cubic C^1 splines which interpolate a given convex data set. The interpolant $s(x)$ is convex if and only if $s^{(2)}(x) \geq 0$ on (x_i, x_{i+1}), $i=1,\ldots,n-1$, giving

$$2d_i + d_{i+1} \leq 3\Delta_i \quad \text{and} \quad -d_i - 2d_{i+1} \leq -3\Delta_i \ , \quad i=1,\ldots,n-1. \qquad (2.5.1)$$

Thus the problem is reduced to the solution of a system of linear inequalities.

Not every convex data set possesses a convex interpolation spline, but under the assumptions

$$2(\Delta_{i+1} - \Delta_i) - (\Delta_i - \Delta_{i-1}) \geq 0 \ , \quad i=2,\ldots,n-2 \ , \qquad (2.5.2)$$

the derivatives d_i giving the convex spline may be obtained through the following algorithm:

$d_i \in [\underline{d}_i, \bar{d}_i]$ are chosen with

$$\underline{d}_1 = 3\Delta_1 - 2\Delta_2 \ , \quad \bar{d}_1 = \Delta_1 \ ,$$

for $i=2,\ldots,n-2$,

$$\underline{d}_i = \max\{\tfrac{1}{2}(3\Delta_{i-1} - d_{i-1}), 3\Delta_i - 2\Delta_{i+1}\}, \quad \bar{d}_i = \min\{3\Delta_{i-1} - 2d_{i-1}, \Delta_i\} \ ,$$

and

$$\underline{d}_{n-1} = \tfrac{1}{2}(3\Delta_{n-2} - d_{n-2}), \quad \bar{d}_{n-1} = \min\{3\Delta_{n-2} - 2d_{n-2}, \Delta_{n-1}\} \ ,$$

$$\underline{d}_n = \tfrac{1}{2}(3\Delta_{n-1} - d_{n-1}) \ , \quad \bar{d}_n = 3\Delta_{n-1} - 2d_{n-1} \ .$$

Mettke further proves that under (2.5.2)

(a) the solution of the difference equation $d_i + 2d_{i+1} = 3\Delta_i$ $(i=1,\ldots n-1)$ with $d_1 = 3\Delta_1 - 2\Delta_2$, is also a solution of (2.5.1); and

(b) if $\Delta_1 \geq 0$ and $3\Delta_i - 2\Delta_{i+1} \leq 0$, $i=1,\ldots,n-2$, then the solution of the difference equation $d_i + 2d_{i+1} = 3\Delta_i$ $(i=1,\ldots,n-1), d_1 = 0$ is also a solution of (2.5.1).

The paper gives error estimates. If $f \in C[x_1, x_n]$ and $f_i = f(x_i)$, then it is proved that

$$\|f - s\| \leq \frac{58}{27}(1+K)\omega(f,h) \ ,$$

where $$K = \max\{\max(h_i/h_{i-1}), 3+3h_1/h_2\}, \quad \text{and}$$

$$\omega(f,h) = \sup\{\ |f(x)-f(y)|\ : x,y \in [x_1,x_n]\ ,\ |x-y| \le h\}$$

is the modulus of continuity.

The theory is extended to a discussion of monotonic and convex cubic-Hermite interpolation splines, and an error analysis is given for this case too.

2.6 Schumaker's shape preserving quadratic spline

The paper by Schumaker [21] considers the use of quadratic splines for solving the interpolation problem $s(x_i)=f_i$, $i=1,\dots,n$, and is concerned with methods which preserve the shape of the data. It shows when it is necessary to add knots to a subinterval and where they can be placed.

The basis of the paper rests on the results of a number of lemmas which help solve the problem of finding $s(x)$ such that

$$\text{in } I_i=[x_i,x_{i+1}], \quad s(x_i)=f_i \ , \quad s(x_{i+1}) = f_{i+1} \ ,$$
$$s^{(1)}(x_i)=d_i \ , \ s^{(1)}(x_{i+1})=d_{i+1} \ .$$

The problem can be solved by a quadratic polynomial only if $d_i+d_{i+1}=2\Delta_i$. Assuming this condition, if $d_i, d_{i+1} \geq 0$, then s is monotonic increasing and if $0 < d_i < d_{i+1}$ then s is convex in I_i .

If $d_i+d_{i+1} \neq 2\Delta_i$, the problem is shown to be solved uniquely using a quadratic spline with a simple knot at $\xi=x_i+\theta h_i$, $0<\theta<1$, where the slope is chosen as

$$\bar{d} = s^{(1)}(\xi) = 2\Delta_i - \{\theta d_i + (1-\theta)d_{i+1}\} \ ,$$

and the quadratic functions in the intervals $x_i<x<\xi$, $\xi \le x < x_{i+1}$ can be constructed easily.

The spline is monotonic increasing only if $d_i, d_{i+1} \geq 0$ and $\bar{d} \geq 0$.

If $d_i < d_{i+1}$ it will be convex if and only if $d_i \le \bar{d} \le d_{i+1}$.

The question then arises as to what knot locations lead to convex splines.

Now, s has no inflexion in I_i if $(d_{i+1}-\Delta_i)(d_i-\Delta_i) < 0$. Assuming this condition and $d_i < d_{i+1}$, then

(a) if $|d_{i+1}-\Delta_i| < |d_i-\Delta_i|$, then for all ξ such that $x_i < \xi \leq \bar{\xi} \equiv x_i + 2h_i(d_{i+1}-\Delta_i)/(d_{i+1}-d_i)$, s is convex on I_i; and further if d_i, $d_{i+1} \geq 0$, s is also monotonic increasing.

(b) if $|d_{i+1}-\Delta_i| > |d_i-\Delta_i|$, then for all ξ such that

$x_{i+1} > \xi \geq \bar{\xi} \equiv x_{i+1} + 2h_i(d_i-\Delta_i)/(d_{i+1}-d_i)$, s is convex on I_i ; and further if d_i, $d_{i+1} \geq 0$, s is also monotonic increasing.

These results form the basis of an algorithm which Schumaker describes to permit the construction of quadratic splines. The user can adjust the knot positions, as required, interactively, and hence can adjust the shape of the spline. In the algorithm, the slope d_i at x_i is computed as a special weighted average of Δ_{i-1} and Δ_i . The choice of knots in those intervals where they are required is made to assure local monotonicity and local convexity. When $(d_i-\Delta_i)(d_{i+1}-\Delta_i) \geq 0$, ξ is chosen at the mid-point of I_i; otherwise it is chosen at the mid-point of the allowed interval. If the condition guaranteeing local monotonicity fails, a correction is made, interactively.

The algorithm is an alternative to McAllister and Roulier's method (see next section). The main difference between these methods is the manner in which slopes and knots are selected.

2.7 McAllister and Roulier's shape-preserving osculatory quadratic spline

In this section, we describe the basis of the McAllister and Roulier method for determining a shape-preserving quadratic spline when the data is monotonic increasing and convex. A detailed account is given in [16] .

Assuming monotonic increasing data, and $\Delta_{i-1} < \Delta_i$ for i=2,...,n-1, the slope values d_i are set as follows:

At interior points (i=2,...,n-1)

$$d_i = \frac{f_i - f_{i-1}}{x_i - \hat{x}}$$, where $\hat{x} = \frac{1}{2}(x_{i-1} + \bar{x})$ and $\bar{x}$ is determined

from the intersection of the horizontal line through (x_{i-1}, f_{i-1}) and the extension of the line of slope Δ_i through (x_i, f_i). See Figure 2.7.1.

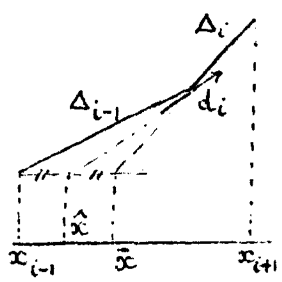


Fig.2.7.1

It is easy to see that d_i satisfies the inequalities

$$\Delta_{i-1} < d_i < \min(\Delta_i, 2\Delta_{i-1}) .$$

At the end points x_1, x_n, essential use is made of d_2, d_{n-1} in obtaining d_1, d_n . See Figure 2.7.2.

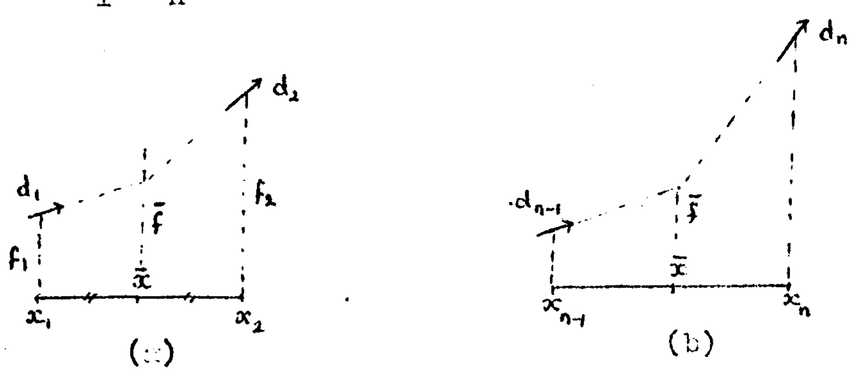


Fig.2.7.2

Thus, for example,

$$d_1 = \frac{\bar{f}-f_1}{\bar{x}-x_1} \quad \text{where } \bar{x} = \tfrac{1}{2}(x_1+x_2) \text{ and } \bar{f} \text{ is determined from}$$

the intersection of the vertical line $x=\bar{x}$ and the line of slope d_2 through (x_2,f_2). Consider, now, an arbitrary interval $I_i=[x_i,x_{i+1}]$ as in Figure 2.7.3.

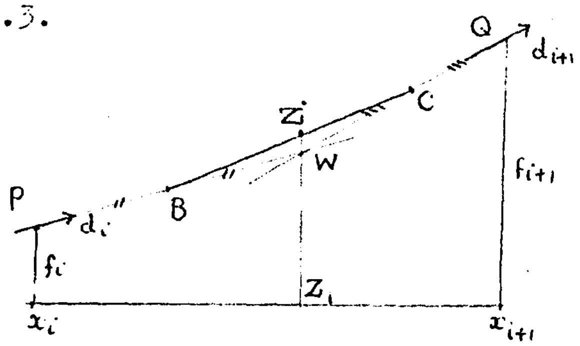


Fig.2.7.3

Using the lines of slopes d_i, d_{i+1} through $P(x_i,f_i), Q(x_{i+1},f_{i+1})$ respectively, we obtain:

the points W, Z_1 ;

mid-points B, C of PW, QW respectively, hence the line segment BC;

and the point Z.

If $(\bar{x},\bar{f})$ are the coordinates of Z, and f_B, f_C are the ordinates at B,C, then in I_i we construct $s(x)$ thus:

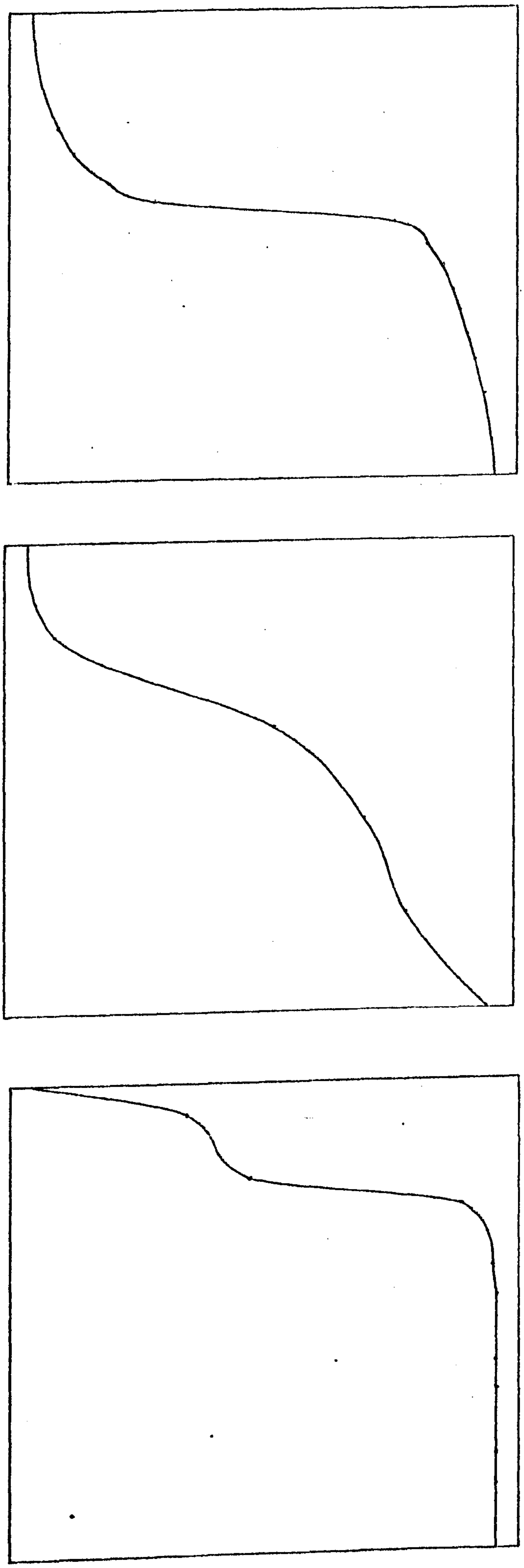
$$s(x) = \begin{cases} \dfrac{1}{(\bar{x}-x_i)^2}\{f_i(\bar{x}-x)^2+2f_B(x-x_i)(\bar{x}-x)+\bar{f}(x-x_i)^2\}, & \text{on } [x_i,\bar{x}] ; \\[2ex] \dfrac{1}{(x_{i+1}-\bar{x})^2}\{\bar{f}(x_{i+1}-x)^2+2f_C(x-\bar{x})(x_{i+1}-x)+f_{i+1}(x-\bar{x})^2\}, & \text{on } [\bar{x},x_{i+1}] \end{cases}$$

These Bernstein piecewise defined quadratic polynomials are such that

$$s(x_i)=f_i, \; s(x_{i+1})=f_{i+1}, \quad s^{(1)}(x_i)=d_i \; , \; s^{(1)}(x_{i+1})=d_{i+1} \; ,$$

and it is assured that $s^{(2)}(x) \geq 0$ throughout $[x_i,x_{i+1}]$.

Thus, a quadratic spline can be constructed on the entire domain $[x_1,x_n]$, and, like the data, is both monotonic and convex.

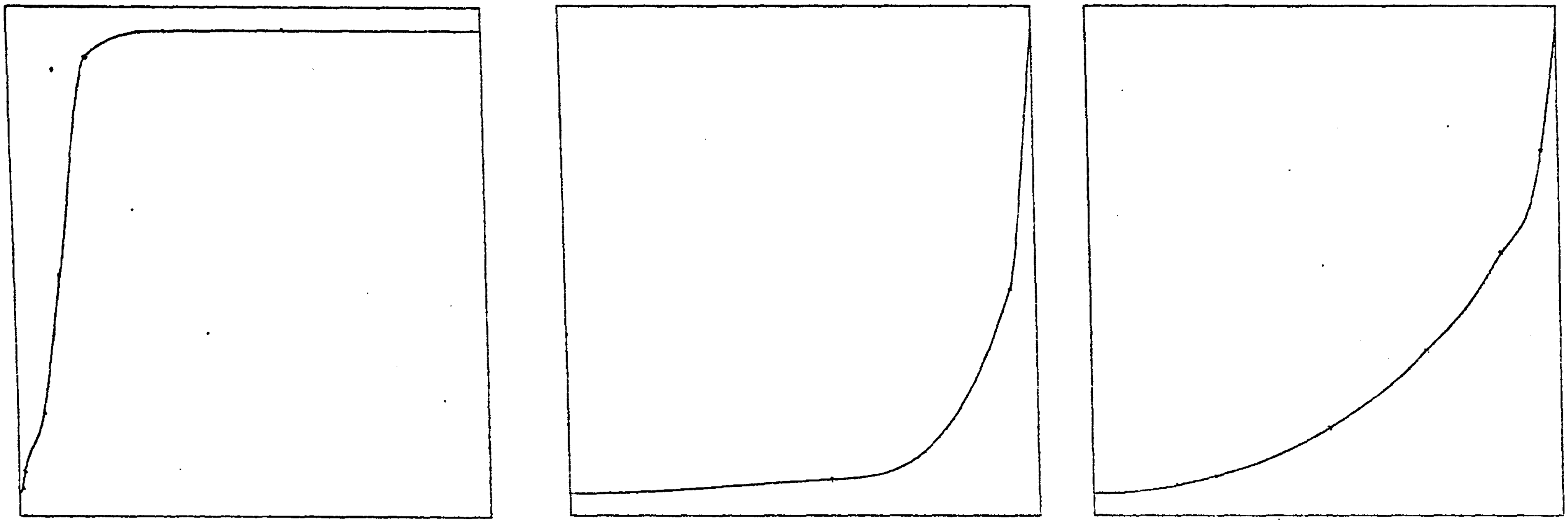


(M1) data

(M2) data

(M3) data

FIG. 2.3.2 FRITSCH-CARLSON program



(M4) data (MC1) data (MC3) data

FIG. 2.3.3 FRITSCH-CARLSON program

Chapter 3

PIECEWISE RATIONAL QUADRATIC INTERPOLATION TO MONOTONIC DATA

3.1 The Rational Quadratic Interpolant

Throughout this chapter the data are monotonic. Without loss of generality, they are considered to be increasing:

$$\left.\begin{array}{r} f_i \leq f_{i+1}\ , \\ \text{or, equivalently, } \Delta_i \geq 0 \end{array}\right\} i = 1,2,\dots,n-1 \ . \tag{3.1.1}$$

A monotonic increasing function $s \in C^1[x_1, x_n]$ is constructed such that

$$s(x_i) = f_i \qquad i = 1,2,\dots,n \tag{3.1.2}$$

$$s^{(1)}(x_i) = d_i \geq 0 \qquad i = 1,2,\dots,n \tag{3.1.3}$$

This is achieved as follows.

On any interval $[x_i, x_{i+1}]$ in which $\Delta_i = 0$, s will have the constant value $f_i = f_{i+1}$, and the derivative values at the ends of the interval must be set at $d_i = 0$, $d_{i+1} = 0$.

On an interval $[x_i, x_{i+1}]$ where $\Delta_i > 0$, s is taken as a rational function, given by

$$s(x) = \frac{f_i(x_{i+1}-x)^2 + a^{(i)}(x_{i+1}-x)(x-x_i) + f_{i+1}(x-x_i)^2}{(x_{i+1}-x)^2 + b^{(i)}(x_{i+1}-x)(x-x_i) + (x-x_i)^2}\ , \quad x \in [x_i, x_{i+1}]$$

to satisfy (3.1.2).

Writing

$$x = x_i + \theta h_i\ , \tag{3.1.4}$$

then $0 \leq \theta \leq 1$, and

$$s(x) = S(\theta) = \frac{f_i(1-\theta)^2 + a^{(i)}\theta(1-\theta) + f_{i+1}\theta^2}{(1-\theta)^2 + b^{(i)}\theta(1-\theta) + \theta^2}\ .$$

To satisfy (3.1.3), we take

$$h_i d_i = S^{(1)}(0) = a^{(i)} - b^{(i)} f_i\ , \quad h_i d_{i+1} = S^{(1)}(1) = -a^{(i)} + b^{(i)} f_{i+1}\ .$$

Solving for $a^{(i)}$, $b^{(i)}$ gives

$$a^{(i)} = (f_i d_{i+1} + f_{i+1} d_i)/\Delta_i ,$$

$$b^{(i)} = (d_{i+1} + d_i)/\Delta_i ,$$

and we note that $b^{(i)} \geq 0$ shows that $S(\theta)$ is never singular.

Thus in $[x_i, x_{i+1}]$, where $\Delta_i > 0$,

$$s(x) = S(\theta) = \frac{\Delta_i f_i (1-\theta)^2 + (f_i d_{i+1} + f_{i+1} d_i)\theta(1-\theta) + \Delta_i f_{i+1}\theta^2}{\Delta_i (1-\theta)^2 + (d_{i+1} + d_i)\theta(1-\theta) + \Delta_i \theta^2}, \quad (3.1.5a)$$

$$= f_i + \frac{(f_{i+1} - f_i)[\Delta_i \theta^2 + d_i \theta(1-\theta)]}{\Delta_i + (d_i + d_{i+1} - 2\Delta_i)\theta(1-\theta)} . \quad (3.1.5b)$$

The latter is a better form for numerical computation, particularly when Δ_i is small.

It is still to be shown that on the assumption that $d_i \geq 0$, $d_{i+1} \geq 0$, then $s^{(1)}(x) \geq 0$ for all $x \in [x_i, x_{i+1}]$, and hence that s is monotonic increasing. A differentiation of either of the formulae (3.1.5) leads, after a little calculation, to

$$s^{(1)}(x) = h_i^{-1} S^{(1)}(\theta)$$

$$= \frac{\Delta_i^2 \{d_{i+1}\theta^2 + 2\Delta_i \theta(1-\theta) + d_i (1-\theta)^2\}}{\{\Delta_i + (d_{i+1} + d_i - 2\Delta_i)\theta(1-\theta)\}^2} , \quad (3.1.6)$$

Monotonicity of the interpolant now follows immediately, given only the necessary conditions $d_i \geq 0$, $d_{i+1} \geq 0$. Hence these conditions are also sufficient. We would like to draw the reader's attention to this, the most important property of the interpolant s. An $O(h^2)$ family of explicit derivative settings for the d_i, $i=1,\ldots,n$ is given in section 3.3. The convergence analysis of the next section shows then that an $O(h^3)$ convergence can be obtained, but that if more accurate derivative data are available an $O(h^4)$ convergence result is achievable. (Chapter 5 considers the problem of high order explicit settings for the d_i.)

3.2 Error bound analysis

Given a monotonic increasing function $f \in C^1[x_1,x_n]$ such that $f(x_i)= f_i$ $(i= 1,\ldots,n)$ and s the piecewise rational quadratic interpolant defined in section 3.1, then if $x\in[x_i,x_{i+1}]$, we have $|f(x)-s(x)| \leq f(x_{i+1})-f(x_i) = h_i f^{(1)}(\xi_i)$, for some $\xi_i \in (x_i,x_{i+1})$.

A higher order convergence result may be obtained. To discuss this, we introduce the following notation, and write (3.1.5a) as

$$s(x)= S(\theta)= P_i(\theta)/Q_i(\theta) \quad (x\in[x_i,x_{i+1}]), \qquad (3.2.1)$$

where

$$\begin{aligned} P_i(\theta) &= \Delta_i f_i(1-\theta)^2 +(f_i d_{i+1}+f_{i+1}d_i)\theta(1-\theta) +\Delta_i f_{i+1}\theta^2 \\ Q_i(\theta) &= \Delta_i(1-\theta)^2 +(d_{i+1}+ d_i)\theta(1-\theta) +\Delta_i\theta^2 . \end{aligned} \qquad (3.2.2)$$

Theorem 3.2.1

Let s be the monotonic increasing interpolant of section 3.1; let $f \in C^4[x_1,x_n]$, and suppose $f^{(1)}(x)>0$ for all x on a compact set $K\subset[x_1,x_n]$ (thus f is strictly monotonic increasing on K). Then for $x\in[x_i,x_{i+1}]$ and $[x_i,x_{i+1}]\subset K$,

$$|f(x)-s(x)| \leq \frac{h_i}{4c_i}\| f^{(1)}\|.\max\{|\lambda_i|,|\lambda_{i+1}|\}$$
$$+\frac{h_i^4}{384c_i}\left[\| f^{(4)}\|.\| f^{(1)}\| + \frac{2h_i}{3}\| f^{(3)}\|^2 + 2\| f^{(2)}\|\,\| f^{(3)}\|\right] \qquad (3.2.3)$$

where
$$c_i \geq \tfrac{1}{2}\min_{[x_i,x_{i+1}]} f^{(1)}(x), \qquad (3.2.4)$$
$$\lambda_i= d_i- f_i^{(1)}, \quad \lambda_{i+1}= d_{i+1} - f_{i+1}^{(1)} \qquad (3.2.5)$$

and all norms are uniform norms on $[x_1,x_n]$.

Proof:

Let $x\in[x_i,x_{i+1}]$, so $x= x_i+ \theta h_i$, $0\leq \theta \leq 1$; and write $f(x)=F_i(\theta)$. We require an upper bound for $|F_i(\theta)-S(\theta)|$, i.e. for

$$\frac{|F_i(\theta)Q_i(\theta) - P_i(\theta)|}{Q_i(\theta)} . \qquad (3.2.6)$$

Now $Q_i(\theta) = \Delta_i(1-\theta)^2 + (d_{i+1}+d_i)\theta(1-\theta) + \Delta_i\theta^2$

$\geq \min\{\Delta_i , \frac{1}{2}\Delta_i + \frac{1}{4}(d_{i+1}+d_i)\}$, for all $0 \leq \theta \leq 1$,

$\geq \frac{1}{2}\Delta_i$, for all choices of d_i, $d_{i+1} \geq 0$,

$$\geq \frac{1}{2} \min_{[x_i, x_{i+1}]} f^{(1)}(x) \qquad (3.2.7)$$

A lower bound for the denominator has thus been found.

The numerator

$$F_i(\theta)Q_i(\theta) - P_i(\theta) = F_i(\theta)\{Q_i(\theta) - Q_i^*(\theta)\} + P_i^*(\theta) - P_i(\theta) + \{F_i(\theta)Q_i^*(\theta) - P_i^*(\theta)\} \qquad (3.2.8)$$

where

$$P_i^*(\theta) = \Delta_i f_i(1-\theta)^2 + (f_i f_{i+1}^{(1)} + f_{i+1} f_i^{(1)})\theta(1-\theta) + \Delta_i f_{i+1}\theta^2 ,$$

$$Q_i^*(\theta) = \Delta_i(1-\theta)^2 + (f_{i+1}^{(1)} + f_i^{(1)})\theta(1-\theta) + \Delta_i\theta^2 \qquad (3.2.9)$$

In (3.2.8),

$$|F_i(\theta)\{Q_i(\theta) - Q_i^*(\theta)\} + P_i^*(\theta) - P_i(\theta)|$$

$$= |\theta(1-\theta)\{(d_{i+1} - f_{i+1}^{(1)})(F_i(\theta) - f_i) + (d_i - f_i^{(1)})(F_i(\theta) - f_{i+1})\}|$$

$= \theta(1-\theta)|\lambda_{i+1}(F_i(\theta) - f_i) + \lambda_i(F_i(\theta) - f_{i+1})|$, using definition(3.2.5)

$= \theta(1-\theta)|\lambda_{i+1}.\theta h_i f^{(1)}(\xi_1) + \lambda_i(1-\theta)h_i f^{(1)}(\xi_2)|$,

for some $\xi_1, \xi_2 \in [x_1, x_2]$,

$$\leq \tfrac{1}{4}h_i \|f^{(1)}\|.\max\{|\lambda_i|, |\lambda_{i+1}|\} \qquad (3.2.10)$$

Also, in (3.2.8), $P_i^*(\theta)$ is the cubic Hermite interpolant to $F_i(\theta)Q_i^*(\theta)$ on $0 \leq \theta \leq 1$ (in fact degenerating to a quadratic function). This is because

$P_i(0) = F_i(0)Q_i^*(0)$, $\qquad P_i^{(1)}(0) = \frac{d}{d\theta}(F_i(\theta)Q^*(\theta))\big|_{\theta=0}$

and $P_i(1) = F_i(1)Q_i^*(1)$, $\qquad P_i^{(1)}(1) = \frac{d}{d\theta}(F_i(\theta)Q^*(\theta))\big|_{\theta=1}$.

Hence, for the last term in (3.2.8), the error bound for cubic Hermite interpolation gives

$$\left|F_i(\theta)Q_i^*(\theta) - P_i^*(\theta)\right| \leq \frac{1}{384} \max_{[0,1]} \left| \frac{d^4}{d\theta^4}(F_i(\theta)Q_i^*(\theta))\right| \qquad (3.2.11)$$

Here,

$$\frac{d^4}{d\theta^4}(F_i(\theta)Q_i^*(\theta)) = F_i^{(4)}(\theta)Q_i^*(\theta) + 4F_i^{(3)}(\theta)Q_i^{*(1)}(\theta) + 6F_i^{(2)}(\theta)Q_i^{*(2)}(\theta),$$

since $Q_i^*(\theta)$ is quadratic,

$$= h_i^4 f^{(4)}(x(\theta))Q_i^*(\theta)$$
$$+ 4h_i^3 f_i^{(3)}(x(\theta)).[f_{i+1}^{(1)} + f_i^{(1)} - 2\Delta_i].(1-2\theta)$$
$$+ 6h_i^2 f^{(2)}(x(\theta)).[f_{i+1}^{(1)} + f_i^{(1)} - 2\Delta_i].(-2)$$

and $Q_i^*(\theta) = f^{(1)}(\xi(\theta))$, since $Q_i^*(\theta)$ is a convex combination of derivatives $f_i^{(1)}$, $f_{i+1}^{(1)}$ and $\Delta_i = f_\alpha^{(1)}$ $(\alpha \in (x_i, x_{i+1}))$;

also, $f_{i+1}^{(1)} + f_i^{(1)} - 2\Delta_i = \frac{1}{6}h_i^2 f^{(3)}(\eta)$, by a Peano Kernel argument (see Appendix A.2)

Hence (3.2.11) gives

$$\left|F_i(\theta)Q_i^*(\theta) - P_i^*(\theta)\right| \leq \frac{1}{384}\left\{ h_i^4\|f^{(4)}\|.\|f^{(1)}\| + \frac{2}{3}h_i^5 \|f^{(3)}\|^2 + 2h_i^4\|f^{(2)}\|\,\|f^{(3)}\|\right\}$$

(3.2.12)

Using (3.2.10) and (3.2.12) in (3.2.8) gives

$$\left|F_i(\theta)Q_i(\theta) - P_i(\theta)\right| \leq \tfrac{1}{4}h_i \|f^{(1)}\|.\max\{|\lambda_i|, |\lambda_{i+1}|\}$$
$$+ \frac{h_i^4}{384}\{\|f^{(4)}\|\,\|f^{(1)}\| + \frac{2}{3}h_i\|f^{(3)}\|^2 + 2\|f^{(2)}\|\,\|f^{(3)}\|\}$$

(3.2.13)

Thus an upper bound for the numerator of (3.2.6) has been found. The result (3.2.3) now follows from (3.2.13) and (3.2.7).

<u>Remark 1</u>

By following the above method, we can find upper bounds for

$$|f^{(r)}(x) - s^{(r)}(x)| \quad , r=1,2,3 .$$

For example,

$$|f^{(1)}(x)-s^{(1)}(x)| = \frac{1}{h_i}\left|\frac{1}{Q_i(\theta)}\cdot\frac{d}{d\theta}(F_i(\theta)Q_i(\theta)-P_i(\theta)) - \frac{Q_i^{(1)}(\theta)}{\{Q_i(\theta)\}^2}(F_i(\theta)Q_i(\theta)-P_i(\theta))\right| ,$$

where we can write $F_i(\theta)Q_i(\theta)-P_i(\theta)$ in the form (3.2.8). Upper bounds for $\frac{d}{d\theta}(F_i(\theta)Q_i^*(\theta)-P_i^*(\theta))$ can be found using the results of Birkhoff and Priver, [3] , and the remaining terms can also be bounded. However, the calculations are rather involved and so are not quoted here.

Remark 2

Equation (3.2.3) in Theorem 3.2.1 shows that the order of convergence depends on the accuracy of d_i, d_{i+1} as approximations to $f_i^{(1)}$, $f_{i+1}^{(1)}$ respectively. If one chooses $d_i = f_i^{(1)}$ and $d_{i+1}=f_{i+1}^{(1)}$ then (since $\lambda_i=0=\lambda_{i+1}$) a best order of bound, $O(h_i^4)$ is possible. In practice, however, derivatives have to be estimated at the points x_i . A family of suitable choices for calculating the d_i explicitly is discussed next.

3.3 Explicit $O(h^2)$ settings for derivatives d_i: a family of choices

When the piecewise rational quadratic interpolation scheme is applied to a monotonic increasing data set (such as one from section 1.3), a method for choosing the derivatives d_i appropriately $(i=1,2,\ldots,n)$ has to be found.

For the purposes of this section (only), let us use the abbreviations

$$\left.\begin{array}{l} \alpha_1 = 1+h_1/h_2 , \quad \beta_1 = -h_1/h_2 \\ \gamma_i = h_i/(h_{i-1}+h_i), \quad \delta_i = h_{i-1}/(h_{i-1}+h_i) , \text{ for } i=2,\ldots,n-1 \\ \alpha_n = 1+h_{n-1}/h_{n-2} , \quad \beta_n = -h_{n-1}/h_{n-2} \end{array}\right\} \quad (3.3.1$$

In terms of these constants, we further define

$$\left.\begin{array}{ll}
\text{(i)} & A_1 = \alpha_1\Delta_1 + \beta_1\Delta_{1,3} \\
& G_1 = \Delta_1^{\alpha_1}.\Delta_{1,3}^{\beta_1} \quad \text{when} \quad \Delta_{1,3} > 0 \\
& H_1 = 1/\{\alpha_1/\Delta_1 + \beta_1/\Delta_{1,3}\}\ , \ \text{when} \ \Delta_2 > 0 \\
\text{(ii)} & A_i = \gamma_i\Delta_{i-1} + \delta_i\Delta_i \\
& G_i = \Delta_{i-1}^{\gamma_i}.\Delta_i^{\delta_i} \qquad (i=2,\ldots,n-1) \\
& H_i = 1/\{\gamma_i/\Delta_{i-1} + \delta_i/\Delta_i\}\ , \ \text{when} \ \Delta_{i-1,i+1} > 0 \\
\text{(iii)} & A_n = \alpha_n\Delta_{n-1} + \beta_n\Delta_{n-2,n} \\
& G_n = \Delta_{n-1}^{\alpha_n}.\Delta_{n-2,n}^{\beta_n} \quad \text{when} \quad \Delta_{n-2,n} > 0 \\
& H_n = 1/\{\alpha_n/\Delta_{n-1} + \beta_n/\Delta_{n-2,n}\}\ , \ \text{when} \ \Delta_{n-2,n} > 0
\end{array}\right\} \qquad (3.3.2)$$

We propose now three possible choices of methods for estimating the d_i, and analyse their usefulness.

Method 1. ("A" settings)

We set

$$d_1 = \begin{cases} A_1, & \text{if } A_1 > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$d_i = A_i\ , \ \text{for } i=2,\ldots,n-1 \qquad (3.3.3)$$

$$d_n = \begin{cases} A_n, & \text{if } A_n > 0 \\ 0 & \text{otherwise} \end{cases}$$

Method 2. ("G" settings)

We set

$$d_1 = \begin{cases} G_1, & \text{if } \Delta_{1,3} > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$d_i = G_i\ , \ \text{for } i=2,\ldots,n-1 \qquad (3.3.4)$$

$$d_n = \begin{cases} G_n, & \text{if } \Delta_{n-2,n} > 0 \\ 0 & \text{otherwise} \end{cases}$$

Method 3. ("H" settings)

We set

$$d_1 = \begin{cases} H_1, & \text{if } \Delta_2 > 0 \\ 2\Delta_1 & \text{otherwise} \end{cases}$$

$$d_i = \begin{cases} H_i & \text{if } \Delta_{i-1,i+1} > 0 \\ 0 & \text{otherwise} \end{cases} \qquad (3.3.5)$$

$$d_n = \begin{cases} H_n, & \text{if } \Delta_{n-2} > 0 \\ 2\Delta_{n-1} & \text{otherwise} \end{cases}$$

The quantities A,G,H are generalised arithmetic, geometric and harmonic means, respectively, of pairs of slopes. The numbers $\alpha_1, \ldots, \beta_n$ are weights, and all but two of these (β_1 and β_n) are positive. When, as happens only at the interior points, the weights concerned are positive, we have the standard result

$$0 < H_i \leq G_i \leq A_i \qquad (i=2,\ldots,n-1)$$

For a given $f \in C^3[x_1, x_n]$, Taylor expansions show that these give $O(h^2)$ approximations for the derivatives. An example of this result can be followed in Appendix A.3. For a more systematic mode of proceeding, see Chapter 5, where the above $O(h^2)$ settings are justified and higher order settings are worked out.

Further remarks in connexion with the choice of method must be made.

Remark 1. (On interior settings d_i, $i=2,\ldots,n-1$)

A_i, G_i, H_i satisfy the inequalities

$$\min\{\Delta_{i-1}, \Delta_i\} \leq H_i \leq G_i \leq A_i \leq \max\{\Delta_{i-1}, \Delta_i\}$$

The fact that $A_i \not\leq 0$ if only one of Δ_{i-1}, Δ_i is zero shows that the setting $d_i = A_i$ is undesirable. (If either $\Delta_{i-1} = 0$ or $\Delta_i = 0$, then $d_i = 0$ for monotonicity).

G_i, however, is well-behaved, although it involves exponentiation. H_i has a simplified form, namely

$$H_i = \Delta_{i-1}\Delta_i / \Delta_{i-1,i+1}$$

which is suitable whenever Δ_{i-1} and Δ_i are not both zero.

Remark 2 (On end conditions d_1, d_n)

The expressions A_1, G_1, H_1 for d_1 all have the desirable property that

$$d_1 \leq \Delta_1 \quad \text{if } \Delta_1 \leq \Delta_{1,3} \quad \text{, and} \quad d_1 \geq \Delta_1 \quad \text{if } \Delta_1 \geq \Delta_{1,3} \text{ .}$$

A_1 has the more familiar, equivalent form

$$A_1 = \Delta_1 + (\Delta_1 - \Delta_2)h_1/(h_1+h_2)$$

but it may yield a negative value. (When this occurs we set $d_1=0$).

G_1 may be written as

$$G_1 = \Delta_1(\Delta_1/\Delta_{1,3})^{h_1/h_2} \text{ .}$$

Here, we may prove that $G_1 \rightarrow 0$ as $\Delta_{1,3} \rightarrow 0$.

H_1 may be written as

$$H_1 = \Delta_1 \Delta_{1,3}/\Delta_2 \text{ .}$$

When $\Delta_2 > 0$, H_1 provides a satisfactory setting for d_1. In the case $\Delta_2 = 0$, a large but finite positive value is used for d_1 (e.g., $2\Delta_1$ is suggested as a possibility).

A similar argument can be given for d_n.

3.4 Test results and discussion

Our first set of results concerns the order of convergence of the rational quadratic interpolation schemes discussed above. We apply our schemes to the exponential test function, data set (MC2), for various choices of the derivative parameters d_i. These choices are for exact derivative settings and the approximations corresponding to using the "A", "G" and "H" settings of methods 1,2 and 3, respectively, except that the exact end conditions are taken, namely $d_1=1.0$, $d_n=\exp(1.0)$. Table 3.4.1 shows the interpolation errors $E=\|f - s\|_\infty$ associated with equal interval lengths h=0.2,0.1,0.05, 0.025. The ratios of the errors E_j/E_{j+1} confirm the expected convergence rate, namely $O(h^4)$ for the exact derivative settings and $O(h^3)$ for each of methods 1,2,3.

Method	Error E_1 (h=0.2)	Error E_2 (h=0.1)	Error E_3 (h=0.05)	Error E_4 (h=0.025)	E_1/E_2	E_2/E_3	E_3/E_4
d_i Arith. "A" settings	$.4620\times10^{-3}$	$.6266\times10^{-4}$	$.8081\times10^{-5}$	$.1029\times10^{-5}$	7.42	7.70	7.85
d_i Geom. "G" settings	$.1217\times10^{-3}$	$.1597\times10^{-4}$	$.2046\times10^{-5}$	$.2529\times10^{-6}$	7.62	7.81	7.90
d_i Harm. "H" settings	$.2178\times10^{-3}$	$.3030\times10^{-4}$	$.3988\times10^{-5}$	$.5113\times10^{-6}$	7.19	7.60	7.80
d_i exact	$.1023\times10^{-4}$	$.6731\times10^{-6}$	$.4315\times10^{-7}$	$.2731\times10^{-8}$	15.20	15.60	15.80

Table 3.4.1 Rational quadratic interpolation errors $E=\|f - s\|_\infty$ on exponential data (MC2)

Our second set of results are graphical. The piecewise rational quadratic schemes using arithmetic, geometric and harmonic derivative approximations, are tested on the data sets (M1), (M2), (M3), (M4), (MC1) and (MC3), and the graphs are shown in Figures 3.4.1 to 3.4.6.

We comment on these, briefly, in relation to the Fritsch-Carlson graphs in Chapter 2.

<u>(M1) data</u>

The first graph seems to be closest to the Fritsch-Carlson curve, while the best of the set of 3 appears to be the middle one.

<u>(M2) data</u>

It is the middle graph which appears comparable to the Fritsch-Carlson curve. Population figures for the year 1965 work out to 202.3,201.8,201.3 million and compare well with the Fritsch-Carlson estimate of 202.5 million.

(M3) data

Our graphs using the "G" and "H" settings for the derivatives seem to smooth out the irregularities in the data.

(M4) data

The last graph shows most similarity with the one of Fritsch-Carlson. The middle one would be our choice.

(MC1) data

The nearest to the Fritsch-Carlson graph is probably the middle one, but the last seems to be the smoothest.

(MC3) data

All three graphs show an improvement on the Fritsch-Carlson curve: the inflexion is not so pronounced in our figures.

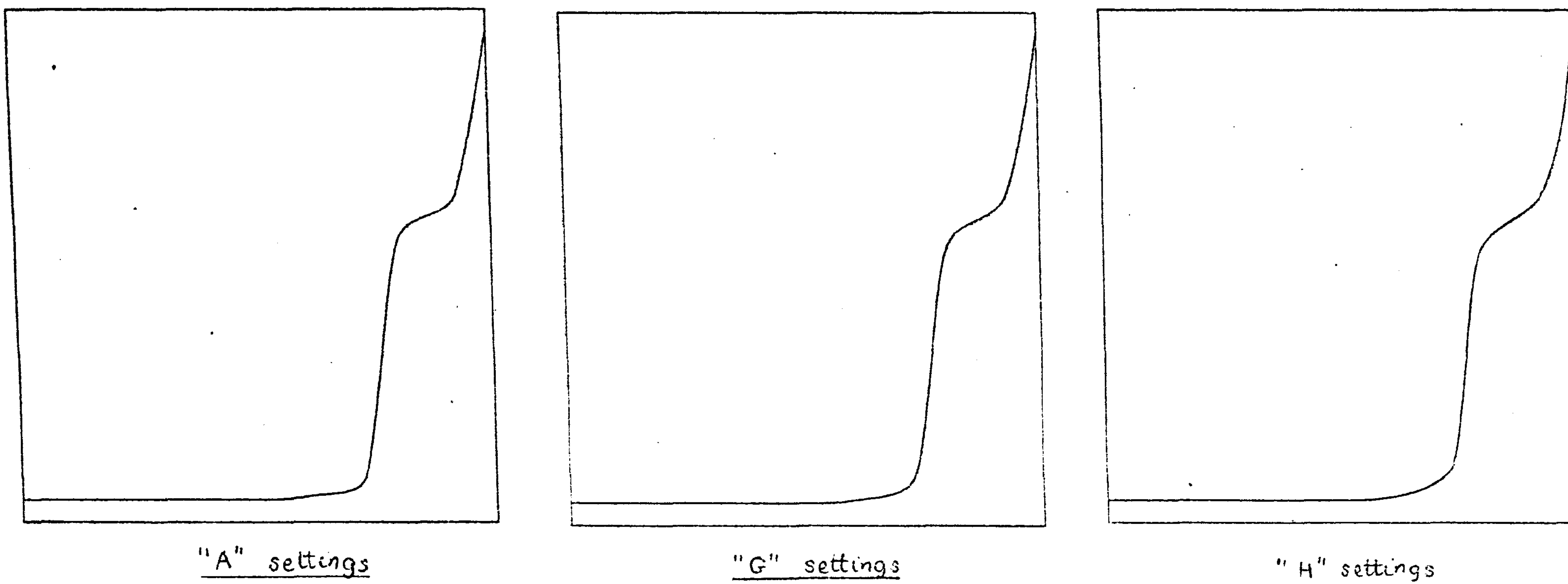
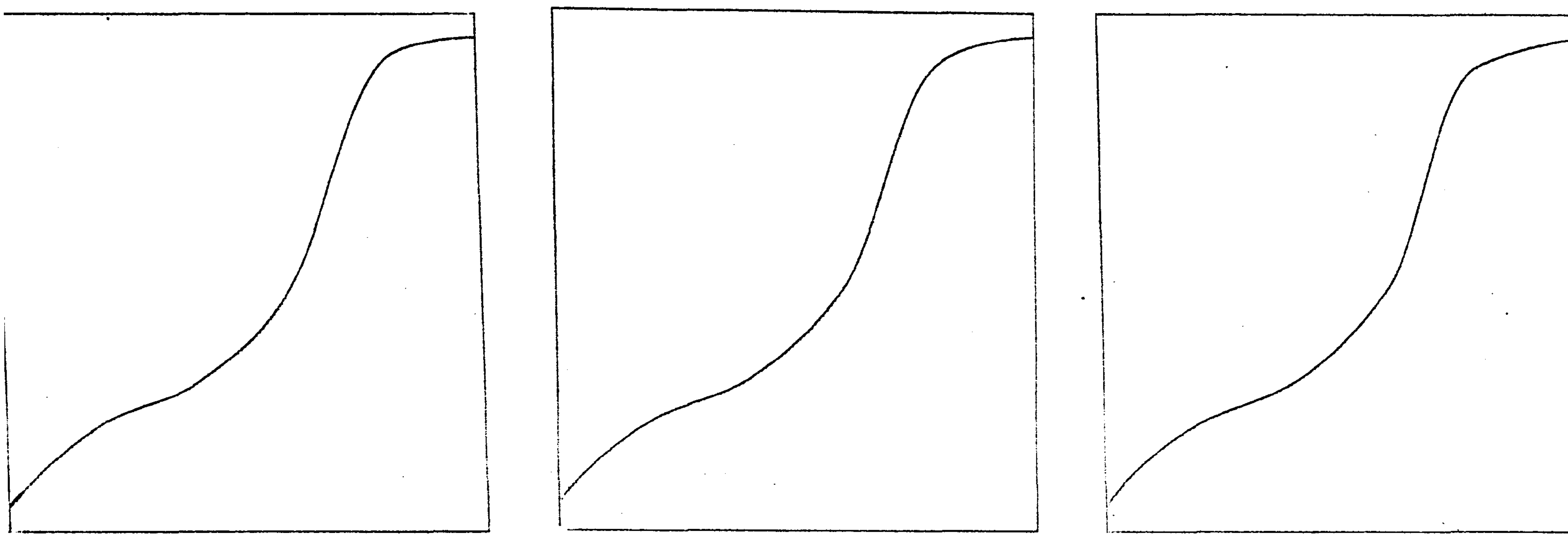


FIG. 3.4.1. C' RATIONAL QUADRATIC (M1) data



"A" settings

"G" settings

"H" settings

FIG. 3.4.2

C' RATIONAL QUADRATIC

(M2) data

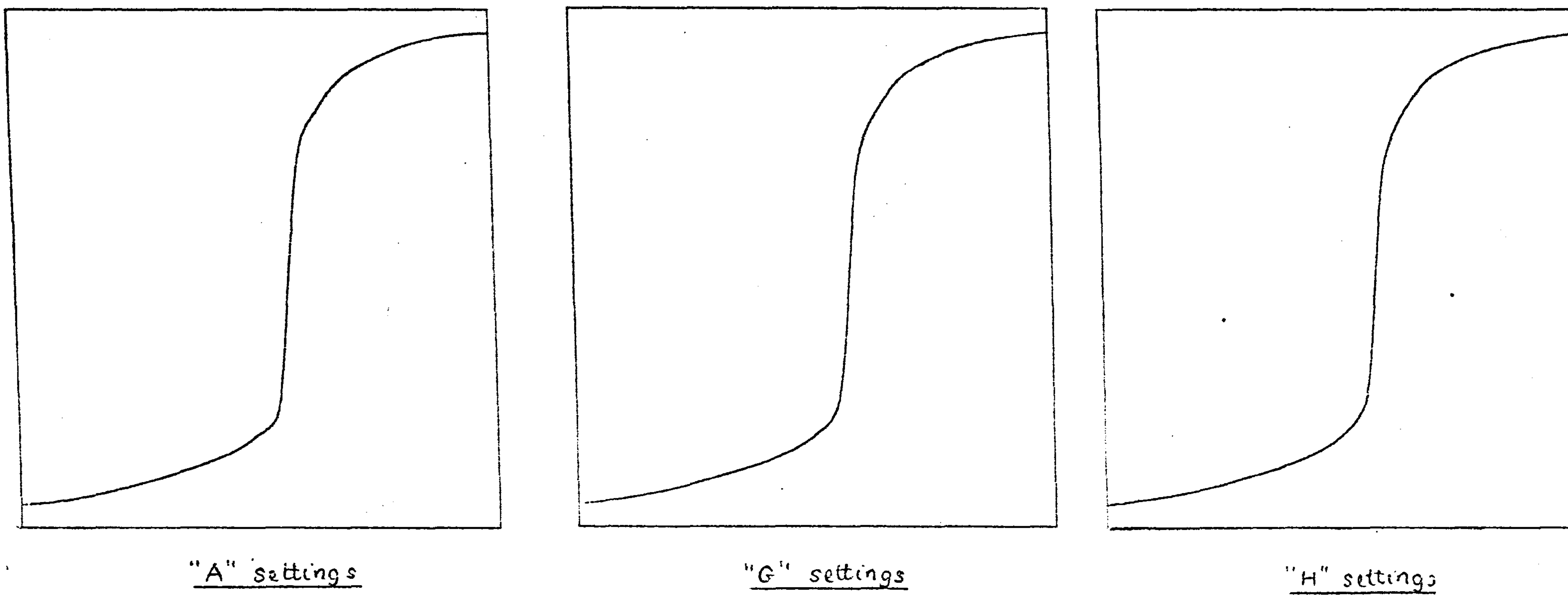
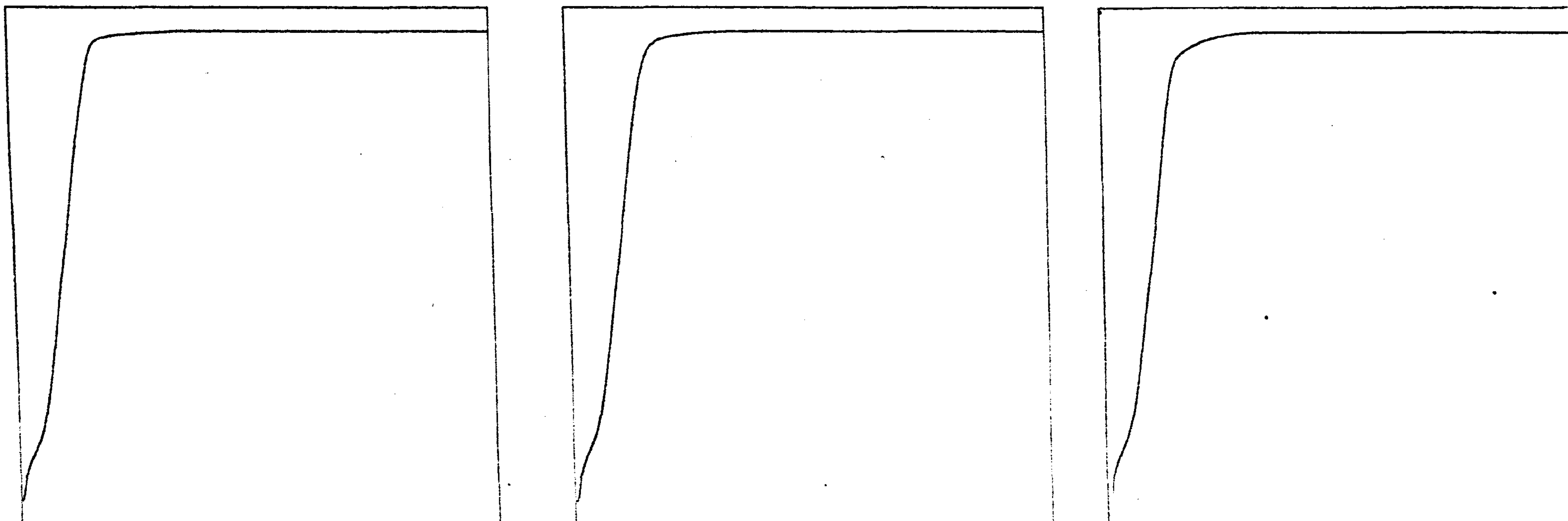


FIG. 3.4.3

C' RATIONAL QUADRATIC (M3) data



"A" settings

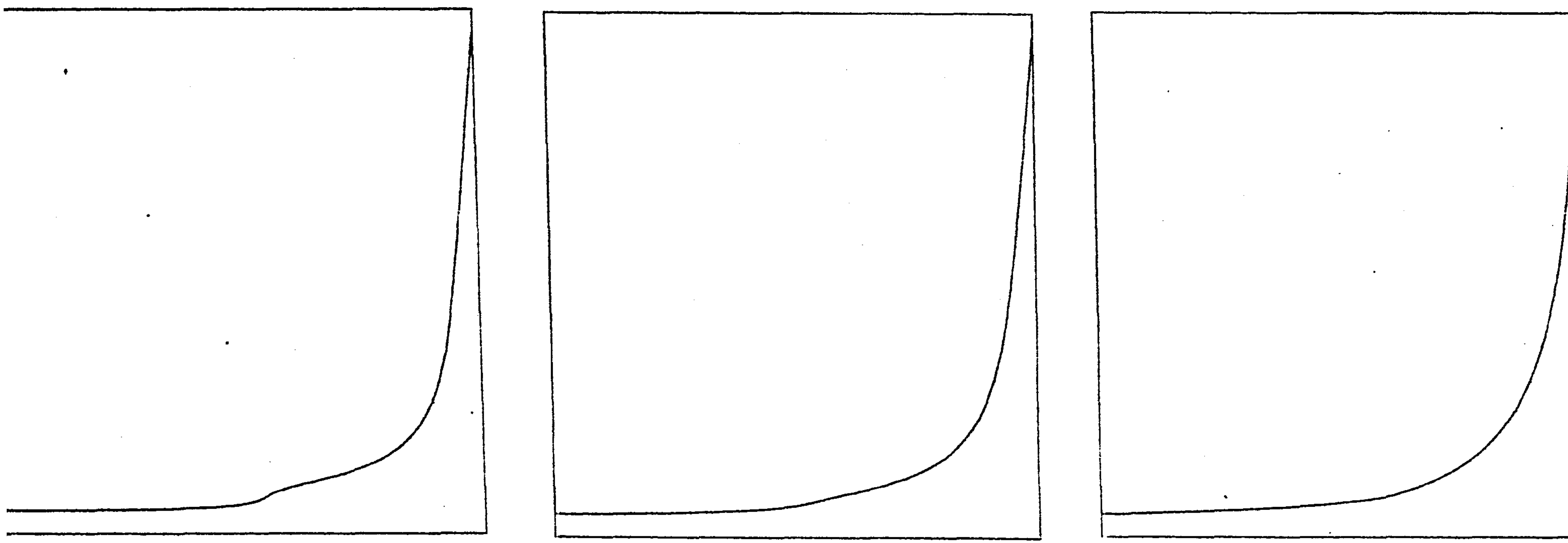
"G" settings

"H" settings

FIG. 3.4.4

C' RATIONAL QUADRATIC

(M4) data



"A" settings

"G" settings

"H" settings

FIG. 3.4.5

C' RATIONAL QUADRATIC

(MCI) data

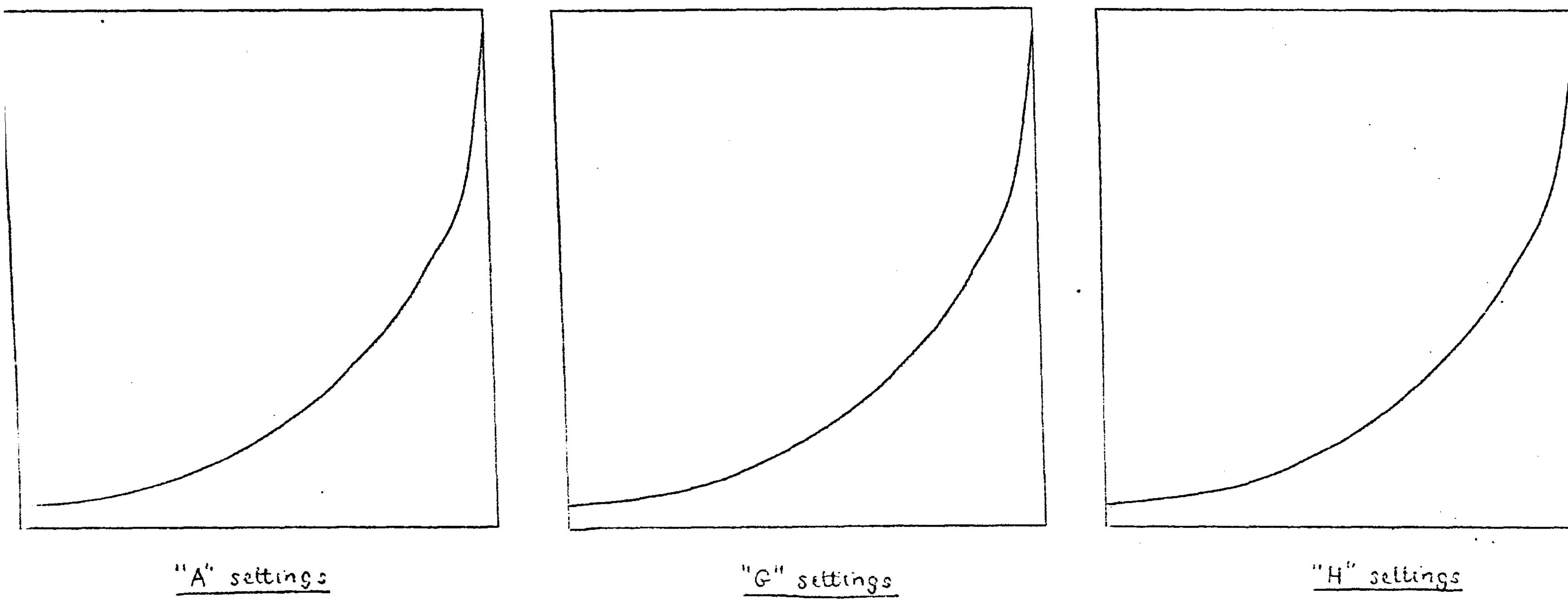


FIG. 3.4.6 C' RATIONAL QUADRATIC

(MC3) data

Chapter 4

C^2 RATIONAL QUADRATIC SPLINE INTERPOLATION TO MONOTONIC DATA

4.1 The Monotonic Rational Spline

In the previous chapter, a piecewise rational quadratic function was developed which produced a monotonic interpolant to monotonic data. The interpolant gives acceptably good curves and is of continuity class C^1.

In this chapter, we restrict the data to be strictly monotonic increasing and show that it is possible to obtain a monotonic rational spline interpolant, with continuity class C^2. Whereas, earlier, the derivatives d_i were determined by local approximation methods, here we need to determine positive values of d_i to make $s \in C^2[x_1,x_n]$. The continuity requirement means that at all interior knots x_i $(i=2,\dots,n-1)$,

$$s^{(2)}(x_i+) - s^{(2)}(x_i-) = 0 .$$

Using the derivative of equation (3.1.6) for the value $\theta = 0$, and the derivative of the equation corresponding to (3.1.6) for the interval $[x_{i-1},x_i]$, for the value $\theta = 1$, we find that

$$\frac{2}{h_i}\left[\Delta_i + d_i\left(1- \frac{d_i + d_{i+1}}{\Delta_i}\right)\right] + \frac{2}{h_{i-1}}\left[\Delta_{i-1} + d_i\left(1- \frac{d_{i-1}+d_i}{\Delta_{i-1}}\right)\right] = 0 \quad (4.1.1)$$

$$(i=2,\dots,n-1)$$

These equations become

$$d_i[-c_i + a_{i-1}d_{i-1} + (a_{i-1}+a_i)d_i + a_i d_{i+1}] = b_i \quad (i=2,\dots,n-1) \quad (4.1.2)$$

where

$$\begin{aligned} a_i &= 1/(h_i\Delta_i) \\ b_i &= \Delta_{i-1}/h_{i-1} + \Delta_i/h_i \\ c_i &= 1/h_{i-1} + 1/h_i \end{aligned} \quad (4.1.3)$$

We note that $c_i > 0$, and for data which are strictly monotonic increasing, $a_i > 0$, $b_i > 0$ for all $i=2,\dots,n-1$.

The equations (4.1.2) are non-linear equations in the derivatives. The sections to follow show that there exists a unique solution to these equations, with all $d_i > 0$, and show also how the solution can be obtained by a convergent iteration.

4.2 Existence and Uniqueness of a solution

Each equation (4.1.2) is a quadratic in the variable d_i. Solving for the positive root gives

$$d_i = \frac{1}{2(a_{i-1}+a_i)}[c_i - a_{i-1}d_{i-1} - a_i d_{i+1} + \{(c_i - a_{i-1}d_{i-1} - a_i d_{i+1})^2 + 4(a_{i-1}+a_i)b_i\}^{\frac{1}{2}}], i=2,\ldots,n-1 \quad (4.2.1)$$

A Jacobi iteration may be defined by the equations

$$d_i^{(k+1)} = \frac{1}{2(a_{i-1}+a_i)}[c_i - a_{i-1}d_{i-1}^{(k)} - a_i d_{i+1}^{(k)} + \{(c_i - a_{i-1}d_{i-1}^{(k)} - a_i d_{i+1}^{(k)})^2 + 4(a_{i-1}+a_i)b_i\}^{\frac{1}{2}}], \ i=2,\ldots,n-1 \quad (4.2.2)$$

where $d_1^{(k+1)} = d_1^{(k)} = d_1$ and $d_n^{(k+1)} = d_n^{(k)} = d_n$ are given end conditions.

In the theorem below the existence and uniqueness of a positive solution (with $d_2,\ldots,d_{n-1} > 0$) is proved by analysing the Jacobi type of iteration (4.2.2). To simplify the presentation of the proof we substitute

$$a_{i-1}+a_i = A_i \quad , \quad 4(a_{i-1}+a_i)b_i = K_i \qquad (4.2.3)$$

<u>Theorem 4.2.1</u> (Existence and Uniqueness)

For strictly increasing data and given end conditions $d_1 \geq 0$, $d_n \geq 0$, there exists a unique solution $d_2,\ldots,d_{n-1}$ satisfying the non-linear consistency equations (4.1.2) and the monotonicity conditions $d_i > 0$, $i=2,\ldots,n-1$.

<u>Proof:</u>

A set of functions G_i, $i=1,\ldots,n$ is defined initially on the domain R^n by

$$G_1(\underline{\xi}) = d_1$$

$$G_i(\underline{\xi}) = \frac{1}{2A_i}[c_i - a_{i-1}\xi_{i-1} - a_i\xi_{i+1} + \{(c_i - a_{i-1}\xi_{i-1} - a_i\xi_{i+1})^2 + K_i\}^{\frac{1}{2}}] , \quad i=2,\ldots,n-1 ,$$

$$G_n(\underline{\xi}) = d_n ,$$

where $\underline{\xi} = (\xi_1,\ldots,\xi_n) \in R^n$. Let $\underline{G} = (G_1,\ldots,G_n)$ and $\underline{d} = (d_1,\ldots,d_n)$. Then the Jacobi iteration (4.2.2) assumes the form

$$\underline{d}^{(k+1)} = \underline{G}(\underline{d}^{(k)}) .$$

Now $c_i - \zeta_i + \{(c_i - \zeta_i)^2 + K_i\}^{\frac{1}{2}}$ is a monotonic decreasing function of $\zeta_i = a_{i-1}\xi_{i-1} + a_i\xi_{i+1}$. Thus, for $\xi_i \geq 0$, $i=1,\ldots,n$ and hence $\zeta_i \geq 0$, we have

$$G_i(\underline{\xi}) \leq \beta_i , \quad i=1,\ldots,n$$

where $\beta_i = \frac{1}{2A_i}[c_i + (c_i^2 + K_i)^{\frac{1}{2}}]$, $i=2,\ldots,n-1$

and $\beta_1 = d_1$, $\beta_n = d_n$.

Furthermore, for $\xi_i \leq \beta_i$, $i=1,\ldots,n$, and hence $\zeta_i \leq a_{i-1}\beta_{i-1} + a_i\beta_{i+1}$, we have $\alpha_i \leq G_i(\underline{\xi})$, $i=1,\ldots,n$,

where $\alpha_i = \frac{1}{2A_i}[c_i - a_{i-1}\beta_{i-1} - a_i\beta_{i+1} + \{(c_i - a_{i-1}\beta_{i-1} - a_i\beta_{i+1})^2 + K_i\}^{\frac{1}{2}}]$, $i=2,\ldots,n-1$,

and $\alpha_1 = d_1$, $\alpha_n = d_n$.

Hence $\alpha_i \leq G_i(\underline{\xi}) \leq \beta_i$ for all $\alpha_i \leq \xi_i \leq \beta_i$, where $\alpha_i > 0$, $i=2,\ldots,n-1$. Thus $\underline{G}: I \rightarrow I$, where $I \in R^n$ is the n-dimensional interval $I = [\alpha_1,\beta_1] \times \ldots \times [\alpha_n,\beta_n]$. Moreover, since $G_i(\underline{\xi}) \geq 0$ for all $\underline{\xi} \in R^n$, the analysis shows that the Jacobi iteration will become restricted to I for any initial vector $\underline{d}^{(0)} \in R^n$.

Next, we show $\underline{G}$ is a contraction mapping on I. For this, let $\underline{\xi}$, $\underline{\eta} \in I$, and

$$X_i = c_i - a_{i-1}\xi_{i-1} - a_i\xi_{i+1}, \quad Y_i = c_i - a_{i-1}\eta_{i-1} - a_i\eta_{i+1} .$$

Then, for $i=2,\ldots,n-1$,

$$G_i(\underline{\xi})-G_i(\underline{\eta}) = \frac{1}{2A_i}[X_i - Y_i + (X_i^2+K_i)^{\frac{1}{2}} - (Y_i^2+K_i)^{\frac{1}{2}}]$$

$$= \frac{X_i - Y_i}{2A_i}\left[1 + \frac{X_i+Y_i}{(X_i^2+K_i)^{\frac{1}{2}}+(Y_i^2+K_i)^{\frac{1}{2}}}\right] ,$$

and $G_1(\underline{\xi}) - G_1(\underline{\eta}) = 0, \quad G_n(\underline{\xi}) - G_n(\underline{\eta}) = 0$.

Now

$$|X_i - Y_i|/A_i \leq \|\underline{\xi} - \underline{\eta}\|_\infty \text{ , and}$$

$$\frac{|X_i + Y_i|}{(X_i^2+K_i)^{\frac{1}{2}}+(Y_i^2+K_i)^{\frac{1}{2}}} \leq \frac{|X_i|+|Y_i|}{\{(|X_i| + |Y_i|)^2+2K_i\}^{\frac{1}{2}}}$$

$$= \frac{1}{\{1 + 2K_i/(|X_i|^2+|Y_i|)^2\}^{\frac{1}{2}}}$$

$$\leq \frac{1}{(1+L)^{\frac{1}{2}}}$$

where, since $|X_i|$, $|Y_i| \leq c_i+a_{i-1}\beta_{i-1}+a_i\beta_{i+1}$,

$$L = 2 \min_{2\leq i\leq n-1} (a_{i-1}+a_i)b_i \Big/ \max_{2\leq i\leq n-1} (c_i+a_{i-1}\beta_{i-1}+a_i\beta_{i+1})^2 > 0 .$$

Hence

$$\| \underline{G}(\underline{\xi}) - \underline{G}(\underline{\eta})\|_\infty \leq \tfrac{1}{2}[1 + 1/(1+L)^{\frac{1}{2}}].\| \underline{\xi} - \underline{\eta}\|_\infty ,$$

and so $\underline{G}$ is a contraction mapping on I. Thus the Jacobi iteration converges to a unique fixed point $\underline{d} \in I$, i.e. $\underline{d} = \underline{G}(\underline{d})$, and hence (4.2.1) has a unique solution. It follows that $\underline{d}$ is the unique solution of (4.1.2) satisfying $d_i > 0$, $i=2,\ldots,n-1$, since the positive root formulation of (4.2.1) is the only one which can give a positive solution.

4.3 Solution by iteration

In practice a Gauss-Seidel type of iteration can be used to solve (4.2.1). This is defined by

$$d_i^{(k+1)} = \frac{1}{2(a_{i-1}+a_i)} [c_i - a_{i-1}d_{i-1}^{(k+1)} - a_i d_{i+1}^{(k)} + \{(c_i - a_{i-1}d_{i-1}^{(k+1)} - a_i d_{i+1}^{(k)})^2 + 4(a_{i-1}+a_i)b_i\}^{\frac{1}{2}}], \quad i=2,..,n-1 \tag{4.3.1}$$

where $d_1^{(k+1)} = d_1^{(k)} = d_1$ and $d_n^{(k+1)} = d_n^{(k)} = d_n$ are given end conditions.

These equations (4.3.1) should be carefully distinguished from equations (4.2.2).

The next theorem shows that this iteration will converge for any choice of the initial values $d_i^{(0)}$, $i=2,...,n-1$.

Theorem 4.3.1

The Gauss-Seidel iteration (4.3.1) converges to the unique positive solution of the non-linear consistency equations (4.1.2).

Proof:

By Theorem 4.2.1., there exist unique $d_i > 0$ satisfying

$$b_i/d_i + c_i - a_{i-1}d_{i-1} - (a_{i-1}+a_i)d_i - a_i d_{i+1} = 0, \quad i=2,...,n-1.$$

Also, the Gauss-Seidel iterates satisfy

$$b_i/d_i^{(k+1)} + c_i - a_{i-1}d_{i-1}^{(k+1)} - (a_{i-1}+a_i)d_i^{(k+1)} - a_i d_{i+1}^{(k)} = 0, \quad i=2,...,n-1.$$

Subtract and write $d_i^{(k)} = d_i + \varepsilon_i^{(k)}$. Then

$$[b_i/\{d_i(d_i+\varepsilon_i^{(k+1)})\} + a_{i-1}+a_i]\, \varepsilon_i^{(k+1)} = -a_{i-1}\varepsilon_{i-1}^{(k+1)} - a_i \varepsilon_{i+1}^{(k)} :$$

since $d_i + \varepsilon_i^{(k+1)} = d_i^{(k+1)} > 0$, on taking moduli, we obtain

$$[b_i/\{d_i(d_i+|\varepsilon_i^{(k+1)}|)\} + a_{i-1}+a_i]|\varepsilon_i^{(k+1)}| \leq a_{i-1}|\varepsilon_{i-1}^{(k+1)}| + a_i|\varepsilon_{i+1}^{(k)}|.$$

Consider the j th inequality, where j is chosen so that

$$|\varepsilon_j^{(k+1)}| = \max_{2\leq i\leq n-1} |\varepsilon_i^{(k+1)}| = \|\underline{\varepsilon}^{(k+1)}\|_\infty .$$

Then

$$[b_j/\{d_j(d_j+\|\underline{\varepsilon}^{(k+1)}\|_\infty)\} + a_{j-1}+a_j]\, \|\underline{\varepsilon}^{(k+1)}\|_\infty \leq a_{j-1}\|\underline{\varepsilon}^{(k+1)}\|_\infty + a_j\|\underline{\varepsilon}^{(k)}\|_\infty ,$$

which reduces to

$$\|\underline{\varepsilon}^{(k+1)}\|_\infty \le \frac{a_j \|\underline{\varepsilon}^{(k)}\|_\infty}{a_j + b_j/\{d_j(d_j + \|\underline{\varepsilon}^{(k+1)}\|_\infty)\}} .$$

It follows that

$$\|\underline{\varepsilon}^{(k+1)}\|_\infty \le \beta \|\underline{\varepsilon}^{(k)}\|_\infty ,$$

where

$$\beta = \frac{a_j}{a_j + b_j/\{d_j(d_j + \|\underline{\varepsilon}^{(0)}\|_\infty)\}}$$

and $0 < \beta < 1$.

Thus $\|\underline{\varepsilon}^{(k)}\|_\infty \to 0$ as $k \to \infty$ and hence $d_i^{(k+1)} \to d_i$, $i=2,\dots,n-1$.

4.4 Error bound analysis

We begin by recalling Theorem 3.2.1 which gives an upper bound for $|f(x) - s(x)|$ in any interval $[x_i, x_{i+1}]$ $(i=1,\dots,n-1)$. The dependence of this bound on the values of $|\lambda_i| = |d_i - f_i^{(1)}|$ and $|\lambda_{i+1}| = |d_{i+1} - f_{i+1}^{(1)}|$ should, in particular, be noted. The next theorem establishes an upper bound for $\max_{2\le i\le n-1} |\lambda_i| = \max_{2\le i\le n-1} |d_i - f_i^{(1)}|$ when the derivatives d_i are the solutions of the non-linear consistency equations (4.1.2).

Theorem 4.4.1

Let $d_1 = f_1^{(1)}$ and $d_n = f_n^{(1)}$ in the rational quadratic spline interpolant. Then with $f \in C^4[x_1, x_n]$, $f^{(1)}(x) > 0$ on $[x_1, x_n]$ and h sufficiently small, $h = \max\{h_i\}$,

$$\max_{2\le i\le n-1} |d_i - f_i^{(1)}| \le \frac{h^3 K(h) \|f^{(1)}\|}{2m^3/\|f^{(1)}\| - h^3 K(h)} \qquad (4.4.1)$$

where

$$K(h) = \frac{1}{12}\{7\|f^{(1)}\|\|f^{(4)}\| + \|f^{(2)}\|\|f^{(3)}\|\} + O(h) \ , \qquad (4.4.2)$$

and

$$m = \min_{[x_1, x_n]} f^{(1)}(x) > 0 . \qquad (4.4.3)$$

Thus $\max_{2\le i\le n-1} |\lambda_i| = O(h^3)$.

Proof:

Consider the consistency equations

$$b_i/d_i + c_i - a_{i-1}d_{i-1} - (a_{i-1}+a_i)d_i - a_i d_{i+1} = 0$$

and let

$$b_i/f_i^{(1)} + c_i - a_{i-1}f_{i-1}^{(1)} - (a_{i-1}+a_i)f_i^{(1)} - a_i f_{i+1}^{(1)} = E_i \,, \qquad (4.4.4)$$

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

On subtracting we find

$$b_i\lambda_i/\{f_i^{(1)}(f_i^{(1)}+\lambda_i)\} + a_{i-1}\lambda_{i-1} + (a_{i-1}+a_i)\lambda_i + a_i\lambda_{i+1} = E_i \,, \qquad (4.4.5)$$

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

An upper bound on $\max\limits_{2\le i\le n-1} |\lambda_i|$ is required.

From (4.4.4) and the definitions (4.1.3) it follows that

$$E_i h_{i-1}h_i\Delta_{i-1}\Delta_i = \{h_i\Delta_{i-1}^2\Delta_i + h_{i-1}\Delta_{i-1}\Delta_i^2\}/f_i^{(1)} + (h_i+h_{i-1})\Delta_{i-1}\Delta_i$$

$$- h_i\Delta_i(f_{i-1}^{(1)}+f_i^{(1)}) - h_{i-1}\Delta_{i-1}(f_i^{(1)}+f_{i+1}^{(1)}) \,.$$

On the right the following Taylor expansions are made (see Appendix A.3):

$$\Delta_{i-1} = f_i^{(1)} - \tfrac{1}{2}h_{i-1}f_i^{(2)} + \tfrac{1}{6}h_{i-1}^2 f_i^{(3)} - \tfrac{1}{24}h_{i-1}^3 f_{i-\alpha}^{(4)} \,,$$

$$\Delta_i = f_i^{(1)} + \tfrac{1}{2}h_i f_i^{(2)} + \tfrac{1}{6}h_i^2 f_i^{(3)} + \tfrac{1}{24}h_i^3 f_{i+\beta}^{(4)} \,,$$

$$f_{i-1}^{(1)} = f_i^{(1)} - h_{i-1}f_i^{(2)} + \tfrac{1}{2}h_{i-1}^2 f_i^{(3)} - \tfrac{1}{6}h_{i-1}^3 f_{i-\gamma}^{(4)} \,,$$

$$f_{i+1}^{(1)} = f_i^{(1)} + h_i f_i^{(2)} + \tfrac{1}{2}h_i^2 f_i^{(3)} + \tfrac{1}{6}h_i^3 f_{i+\delta}^{(4)} \,,$$

where $f_{i-\alpha}^{(4)}$ means $f^{(4)}(x_i - \alpha h_{i-1})$, $0<\alpha<1$, etc.

After much lengthy algebra (which we do not reproduce here), the result of these substitutions is

$$E_i\Delta_{i-1}\Delta_i = f_i^{(1)}\{\tfrac{1}{8}(h_i^2 f_{i+\beta}^{(4)} - h_{i-1}^2 f_{i-\alpha}^{(4)}) - \tfrac{1}{6}(h_i^2 f_{i+\delta}^{(4)} - h_{i-1}^2 f_{i-\gamma}^{(4)})\}$$

$$+ \tfrac{1}{12}(h_i^2 - h_{i-1}^2) f_i^{(2)} f_i^{(3)} + O(h^3) \quad . \qquad (4.4.6)$$

Now $\Delta_{i-1}\Delta_i \geqslant m^2$, where m has been defined in (4.4.3).

Thus,

$$m^2|E_i| \leqslant \tfrac{1}{12}h^2\{\, 7\| f^{(1)}\|\|f^{(4)}\| + \| f^{(2)}\|\| f^{(3)}\|\} + O(h^3)$$

and so

$$|E_i| \leqslant m^{-2}h^2K(h) \ , \qquad (4.4.7)$$

using the definition of K(h) in (4.4.2).

Now consider equation (4.4.5) with index i=j taken so that $|\lambda_j| = \max\limits_{2\leqslant i\leqslant n-1} |\lambda_i|$. Then

$$[b_j/\{f_j^{(1)}(f_j^{(1)}+\lambda_j)\} + a_{j-1} + a_j]\, \lambda_j = E_j - a_{j-1}\lambda_{j-1} - a_j\lambda_{j+1} \ ,$$

where $|\lambda_j| = \|\underline{\lambda}\|_\infty$, because $\lambda_1 = 0 = \lambda_n$. Taking moduli and noting that $0 < f_j^{(1)}+\lambda_j \leqslant f_j^{(1)} + \|\underline{\lambda}\|_\infty$, gives

$$[b_j/\{f_j^{(1)}(f_j^{(1)} + \|\underline{\lambda}\|_\infty)\} + a_{j-1} + a_j]\, \|\underline{\lambda}\|_\infty \leqslant |E_j| + (a_{j-1}+a_j)\|\underline{\lambda}\|_\infty \, .$$

This inequality reduces to

$$\|\underline{\lambda}\|_\infty \leqslant f_j^{(1)}|E_j|/\{b_j/f_j^{(1)} - |E_j|\} \ , \qquad (4.4.8)$$

under the assumption that the denominator is positive.

Now

$$b_j/f_j^{(1)} = (\Delta_{j-1}/h_{j-1} + \Delta_j/h_j)/f_j^{(1)} \ ,$$

$$= (f_{j-\theta}^{(1)}/h_{j-1} + f_{j+\phi}^{(1)}/h_j)/f_j^{(1)}, \text{ for some } 0<\theta,\phi<1 \ ,$$

$$\geqslant 2m/\{h\, \| f^{(1)}\|\} \ ,$$

using the definition of m.

Thus, from (4.4.7),

$$b_j/f_j^{(1)} - |E_j| \geqslant 2m/\{h\,\| f^{(1)}\|\} - m^{-2}h^2K(h) \qquad (4.4.9)$$

which is positive for h sufficiently small.

Finally, substituting (4.4.9) and (4.4.7) in (4.4.8) gives the bound of the theorem.

For equal intervals and under the further assumption that $f \in C^5[x_1,x_n]$, the above result can be improved.

Theorem 4.4.2

Let $f \in C^5[x_1,x_n]$, $f^{(1)}(x) > 0$ for x in $[x_1,x_n]$ and assume $h_i = h$, $i=1,\dots,n-1$ (equal intervals). Let $d_1 = f_1^{(1)}$ and $d_n = f_n^{(1)}$ in the rational quadratic spline interpolant. Then for sufficiently small h,

$$\max_{2 \le i \le n-1} |d_i - f_i^{(1)}| \le \frac{h^4 K(h) \| f^{(1)} \|}{2m^3/\|f^{(1)}\| - h^4 K(h)} \qquad (4.4.10)$$

where

$$K(h) = \frac{1}{12}\|f^{(2)}\|\|f^{(1)}\| + \frac{1}{18}\|f^{(3)}\|^2 + \frac{1}{12}\|f^{(2)}\|^2\|f^{(3)}\|/m + \frac{2}{15}\|f^{(1)}\|\|f^{(5)}\| + O(h) \qquad (4.4.11)$$

Thus $\max_{2 \le i \le n-1} |d_i - f_i^{(1)}| = O(h^4)$.

Proof:

All that it is necessary to do is to consider higher order Taylor expansions. It will result in a cancellation of $O(h^2)$ terms in the equality corresponding to equation (4.4.6). The lengthy details are omitted.

Remark

Taken together, Theorems 3.2.1 and 4.4.1 show that $f(x)-s(x)=O(h^4)$ when $d_1 = f_1^{(1)}$ and $d_n = f_n^{(1)}$ are given end conditions. The last theorem, concerning equal intervals, shows that the first term in the bound on $f(x)-s(x)$ is now $O(h^5)$, and so, for small h, the bound is dominated by the second term.

4.5 $O(h^2)$ end derivative settings

In the application of the C^2 rational quadratic spline scheme to practical data we must set the end derivatives d_1 and d_n to suitable non-negative values. The settings for d_1, d_n defined by the formulae in section 3.3, being $O(h^2)$ accurate, are adequate, although $O(h^4)$ accuracy of the spline is sacrificed near the ends.

Written out explicitly, the choices are as follows:

Method 1 ("A" settings)

$$d_1 = \begin{cases} A_1 \equiv \Delta_1+(h_1/h_2)(\Delta_1-\Delta_{1,3}) & \text{if } A_1>0 \\ 0 & \text{otherwise} \end{cases}$$

$$d_n = \begin{cases} A_n \equiv \Delta_{n-1}+(h_{n-1}/h_{n-2})(\Delta_{n-1}-\Delta_{n-2,n}) & \text{if } A_n>0 \\ 0 & \text{otherwise} \end{cases}$$

Method 2 ("G" settings)

$$d_1 = G_1 \equiv \Delta_1(\Delta_1/\Delta_{1,3})^{h_1/h_2}$$

$$d_n = G_n \equiv \Delta_{n-1}(\Delta_{n-1}/\Delta_{n-2,n})^{h_{n-1}/h_{n-2}}$$

Method 3 ("H" settings)

$$d_1 = H_1 \equiv \Delta_1\Delta_{1,3}/\Delta_2$$

$$d_n = H_n \equiv \Delta_{n-1}\Delta_{n-2,n}/\Delta_{n-2}$$

4.6 Test results and discussion

We have carried out tests on the effectiveness of the Gauss-Seidel iterative method, on the order of error of the interpolation scheme and on the practical data sets listed in Chapter 1.

The Gauss-Seidel Iteration. The Gauss-Seidel type of iteration is easily implemented, where the convergence test $\max_i |d_i^{(k+1)}-d_i^{(k)}| \leq \varepsilon$ is used as the stopping condition. No case was found where convergence presented a problem. In the worst case, of the (M3) data, 13 iterations are required to obtain convergence with $\varepsilon=\frac{1}{2}\times10^{-5}$ and 19 iterations with $\varepsilon=\frac{1}{2}\times10^{-10}$. In other examples, 6-14 iterations suffice to satisfy the convergence test with $\varepsilon=\frac{1}{2}\times10^{-10}$. Our experiments also indicate that the Gauss-Seidel iteration converges in little more than half the number of iterations required for the Jacobi iteration.

In order to implement the iterative method, we have taken

$$d_i^{(0)} = \{b_i/(a_{i-1}+a_i)\}^{\frac{1}{2}}\ , \ i=2,\dots,n-1 \qquad (4.6.1)$$

as initial values. These arise by assuming $c_i - a_{i-1}d_{i-1} - a_i d_{i+1} = 0$ in (4.3.1) and from (4.1.3) it can be seen that this choice corresponds to a local approximation to the d_i. However, numerical experiments show that the choice of initial values is not significant. Indeed, the choice $d_i^{(0)} = 0$ or $d_i^{(0)} = 100$, $i=2,\dots, n-1$, does not substantially increase the number of iterations in all the numerical examples that were considered.

Order of Error. Our first set of results is concerned with the order of error of the interpolation scheme. Tables 4.6.1, 4.6.2 show the errors which arise from the application of the rational quadratic spline scheme to the exponential data, (MC2), when the choice of exact end conditions $d_1=1.0$, $d_n=\exp(1.0)$ is made. For the four choices of interval length $h=0.2, 0.1, 0.05, 0.025$, the non-linear spline equations involve 4, 9, 19 and 39 unknowns, respectively and the Gauss-Seidel iteration converges in 12, 14, 13 and 12 steps respectively with a convergence test of $\varepsilon = \frac{1}{2}\times10^{-10}$.

Error e_1 ($h=0.2$)	Error e_2 ($h=0.1$)	Error e_3 ($h=0.05$)	Error e_4 ($h=0.025$)	e_1/e_2	e_2/e_3	e_3/e_4
$.1697\times10^{-4}$	$.1166\times10^{-4}$	$.7625\times10^{-7}$	$.4844\times10^{-8}$	14.55	15.29	15.74

Table 4.6.1

Rational quadratic spline interpolation errors

$e = \max_i \left| f_i^{(1)} - d_i \right|$ on exponential data (MC2)

Error E_1 (h=0.2)	Error E_2 (h=0.1)	Error E_3 (h=0.05)	Error E_4 (h=0.025)	E_1/E_2	E_2/E_3	E_3/E_4
$.1067\times10^{-4}$	$.6880\times10^{-6}$	$.4363\times10^{-7}$	$.2746\times10^{-8}$	15.51	15.77	15.89

Table 4.6.2

Rational quadratic spline interpolation errors

$E = \| f - s \|_\infty$ on exponential data (MC2)

The first table gives values of $\max_{2 \le i \le n-1} |f_i^{(1)} - d_i|$ and the expected $O(h^4)$ result of Theorem 4.4.2 is confirmed by the ratio of the errors which approach 2^4. The second table shows the uniform norm $\| f - s \|_\infty$ on $[0,1]$, and the ratio of these errors also confirms the expected $O(h^4)$ result given by the theory of section 4.4. The bounds of section 4.4. give overestimates of the errors. This is not surprising since the analysis of the errors involves a non-linear approximation method. Thus with h= 0.025, Theorem 4.4.2 gives

$$\max_i |f^{(1)} - d_i| \le 0.53 \times 10^{-5}$$

which substantially overestimates the true error of 0.4844×10^{-8}. Theorem 3.2.1 gives the bound

$$\| f - s \|_\infty \le 0.46 \times 10^{-7} ,$$

and Theorem 3.2.1, with the bound of Theorem 4.4.2, gives

$$\| f - s \|_\infty \le 0.23 \times 10^{-6} ,$$

compared with the true error $\| f - s \|_\infty = 0.2746 \times 10^{-8}$.

Practical examples

Our second set of results concerns the application of the rational spline scheme to the data sets (M1),(M2),(M3),(M4),(M5) ,(MC1), (MC3).

As end conditions for the C^2 spline scheme we have taken values of d_1, d_n given be the "G" derivative settings of section 4.5. The graphs are appended. See Figures 4.6.1 to 4.6.7. Graphs obtained using the "A" and "H" settings are not shown here, as they are very similar to the present ones. For each data set we have drawn two figures. In the one on the right, we show, additionally, the variation in the derivative. The smoothness of the derivative curves is expected, since the original curves are C^2 continuous.

We comment, very briefly on the individual graphs:

<u>(M1) data</u>

We have used only the strictly increasing portion of this data set, setting the left hand end derivative to zero. On the interval [8,15] ,the constructed curve is C^2.

<u>(M2) data</u>

Using the rational spline scheme, we can predict a population of 201.9 million for the year 1965. This value compares well with our earlier ones using the C^1 schemes.

<u>(M3) data</u>

The C^2 constraint has led to more variation in the curve than that given by the C^1 schemes.

<u>(M4) data</u>

Eleven iterations were necessary to give the d_i. The resulting curve is at least as good as our earlier ones.

<u>(M5) data</u>

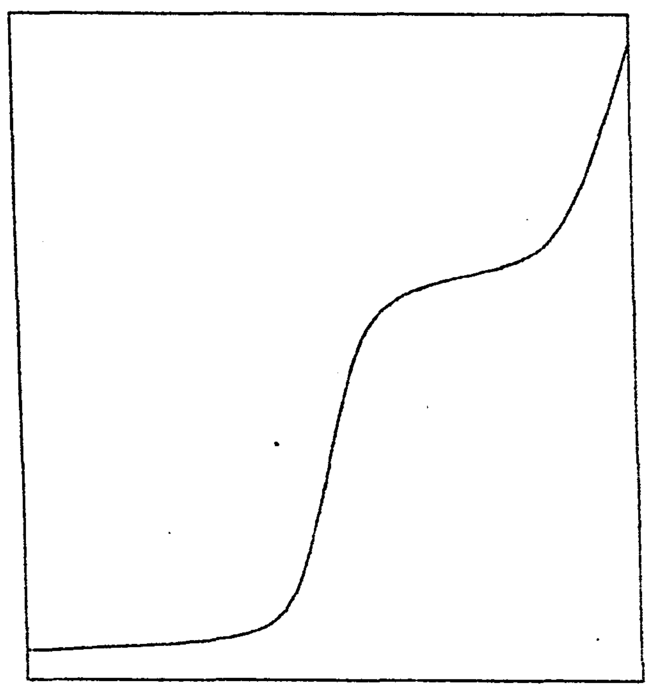
The derivative of this function is the familiar Gaussian distribution function in Statistics.

(MC1) data

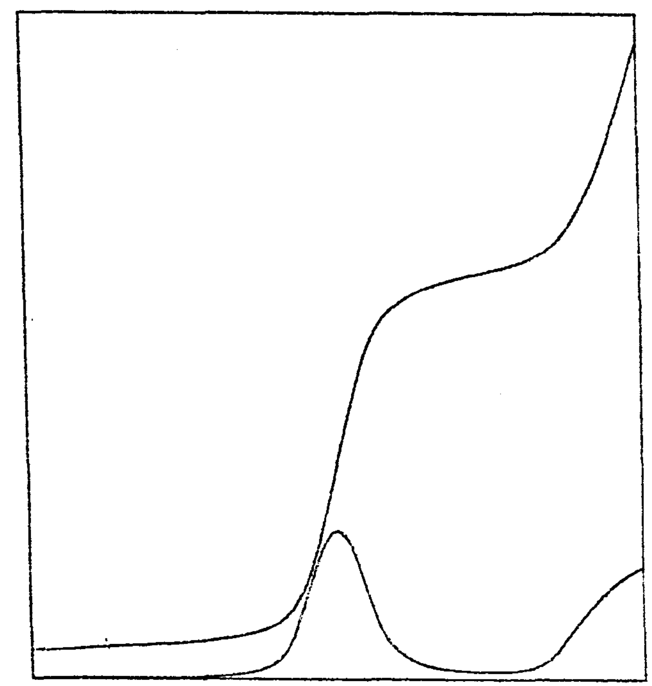
Only 6 Gauss-Seidel iterations establish the interior derivative values for this four-point data set. The curve compares favourably with the true graph of $f(x) = 1/x^2$ through the knots.

(MC3) data

An inflexion point appears in the penultimate interval. A change in the end conditions does not succeed in removing it, however. For example, setting $d_1=0$ and $d_n=25$ produces a similar graph with an inflexion in the same interval as before.



"G" settings



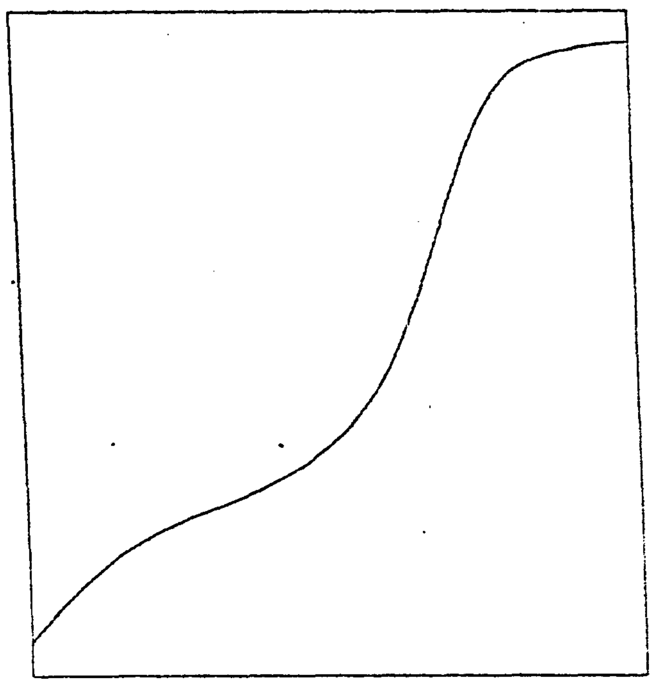
Derivative graph superposed

FIG. 4.6.1

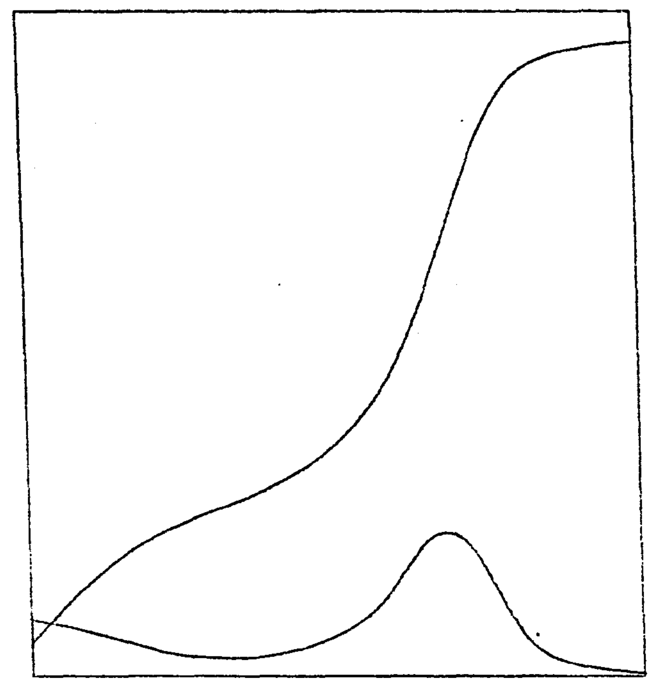
C^2 RATIONAL QUADRATIC SPLINE

(M1) data

[13 iterations]



"G" settings

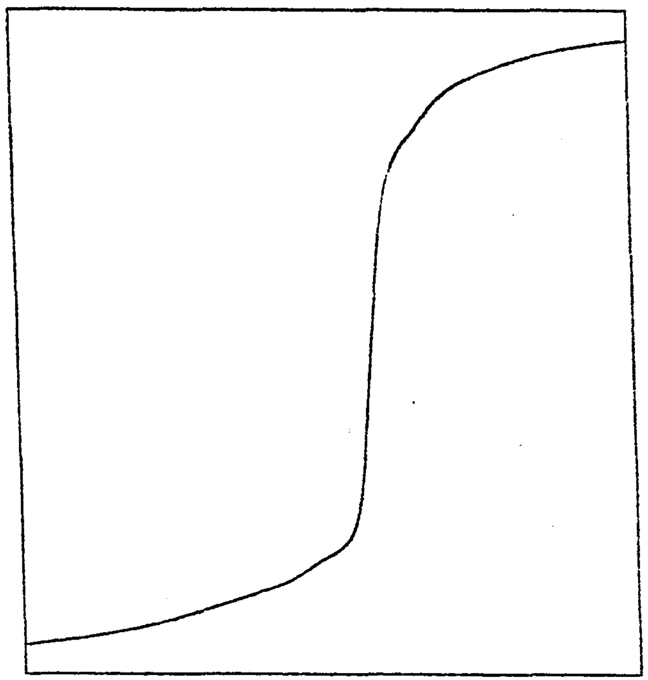


Derivative graph superposed

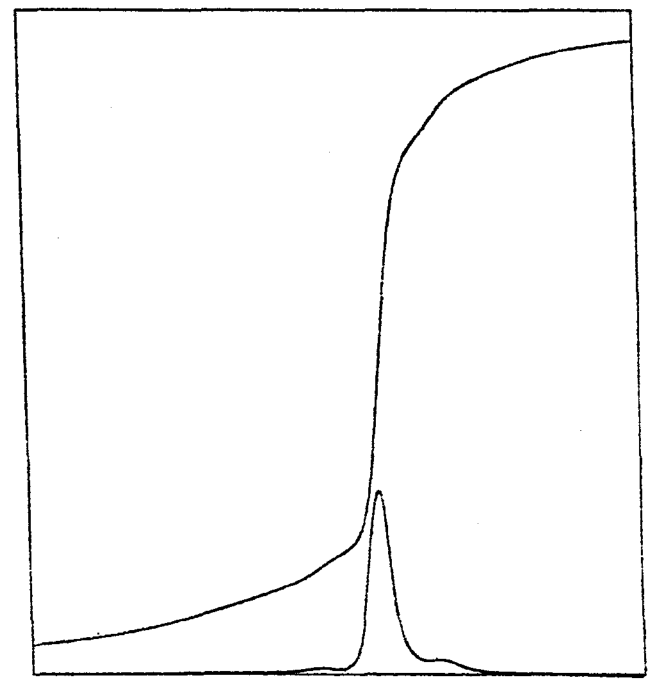
FIG. 4.6.2 C^2 RATIONAL QUADRATIC SPLINE

(M2) data

[10 iterations]



"G" settings

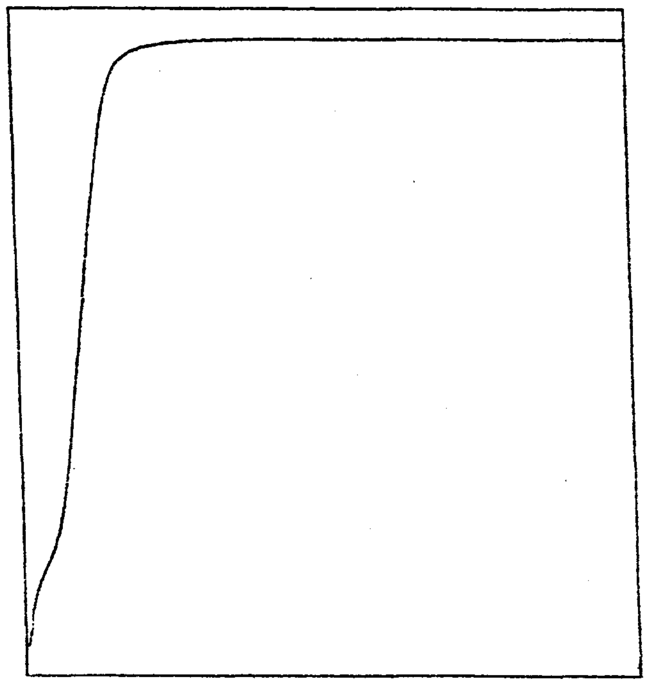


Derivative graph superposed

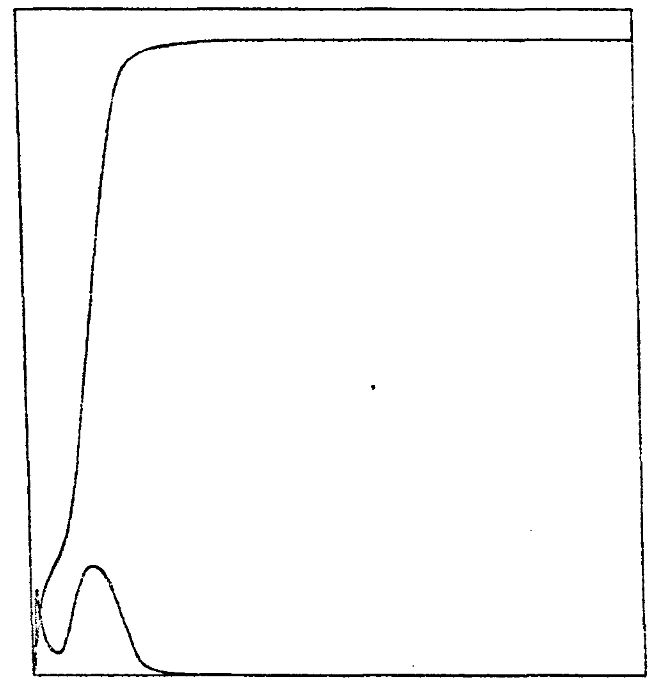
FIG. 4.6.3 C^2 RATIONAL QUADRATIC SPLINE

(M3) data

[19 iterations]



"G" settings

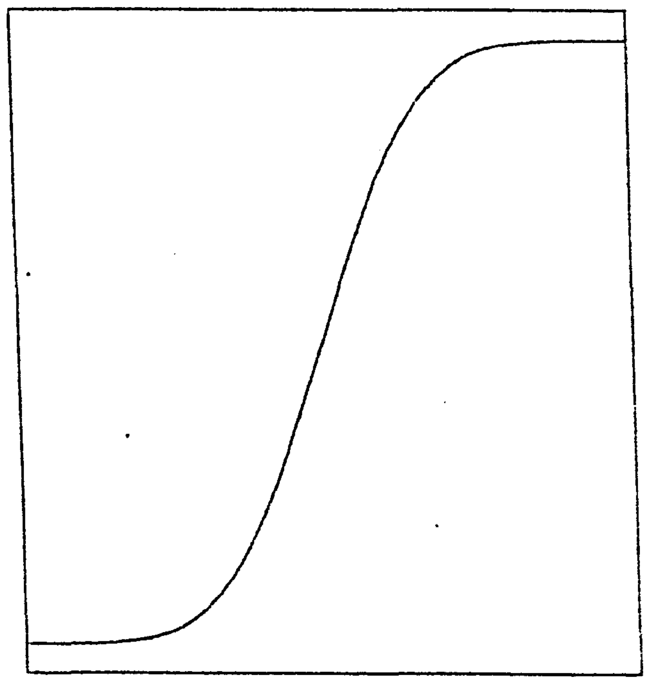


Derivative graph superposed

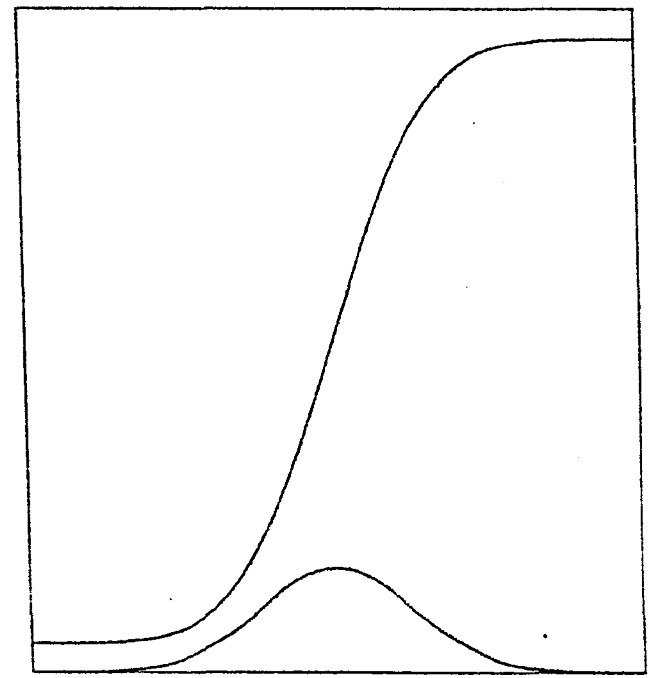
FIG. 4.6.4 C^2 RATIONAL QUADRATIC SPLINE

(M4) data)

[11 iterations]



"G" settings

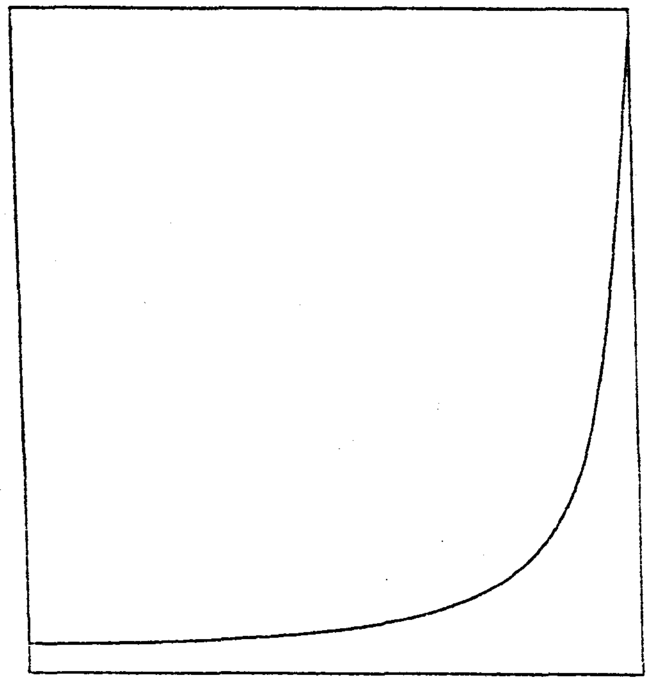


Derivative graph superposed

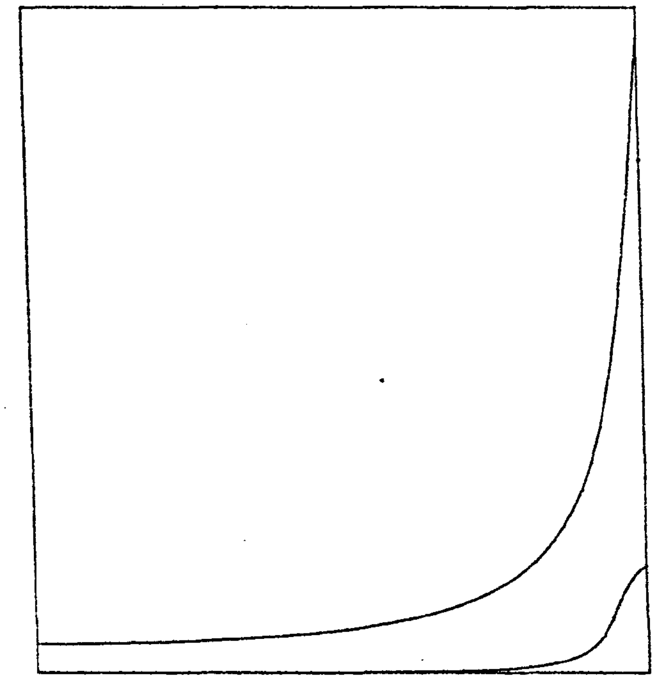
FIG. 4.6.5 C^2 RATIONAL QUADRATIC SPLINE

(M5) data

[11 iterations]



"G" settings

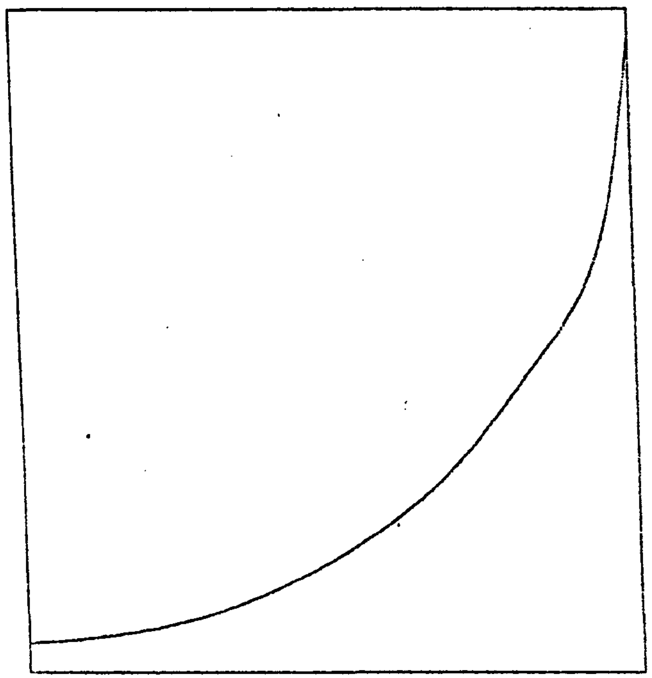


Derivative graph superposed

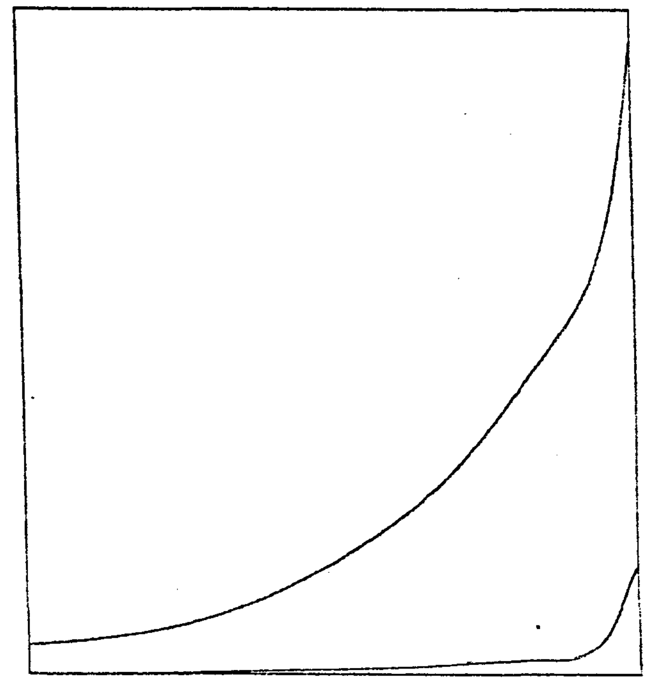
FIG. 4.6.6 C^2 RATIONAL QUADRATIC SPLINE

(MC1) data

[6 iterations]



"G" settings



Derivative graph superposed

FIG. 4.6.7 C^2 RATIONAL QUADRATIC SPLINE

(MC3) data

[13 iterations]

Chapter 5

ACCURATE DERIVATIVE ESTIMATION FOR MONOTONIC RATIONAL QUADRATIC INTERPOLATION: EXPLICIT METHODS

This chapter expands on a few of the ideas of Chapter 3. There, a rational quadratic interpolant was defined and applied with explicit $O(h^2)$ approximations for the derivative parameters. The result was a piecewise defined C^1 monotonic interpolant with a $O(h^3)$ error bound. Following that, in Chapter 4, a monotonic rational quadratic C^2 spline method was developed. The derivative parameters, determined as solutions of a non-linear system of equations, are $O(h^3)$ approximations when exact derivative end conditions are known, so that an $O(h^4)$ error bound for the spline was achieved.

Here we proceed to show how to develop a class of methods for determining explicit high order derivative estimates for the rational quadratic scheme. The end conditions given by the methods, when exact conditions are not known, will be suitable also for the C^2 rational quadratic spline.

Data will be assumed to be monotonic increasing. The monotonicity requirement necessitates that the derivative parameters should be all non-negative, and this leads us to study high order finite difference approximation formulae which will include those arithmetic, geometric and harmonic $O(h^2)$ approximations of section 3.3 as very special cases.

5.1 Preliminary lemmas

Explicit derivative approximations which are $O(h^p)$, $p=2,3,\ldots$, will be based on the use of the methods derived in the next two lemmas.

Lemma 5.1.1

Let ε_j, $j=1,\ldots,N$ be distinct non-zero constants. Then the linear system in the unknowns α_j , $j=1,\ldots,N$ defined by

$$\sum_{j=1}^{N} \alpha_j = 1, \quad \sum_{j=1}^{N} \alpha_j \varepsilon_j^k = 0 \quad \text{for } k=1,\ldots,N-1 \tag{5.1.1}$$

has the unique solution

$$\alpha_j = \prod_{\substack{i=1 \\ i\neq j}}^{N} \frac{\varepsilon_i}{\varepsilon_i - \varepsilon_j} \quad , \quad j=1,\ldots,N \tag{5.1.2}$$

Proof:

The system (5.1.1) is

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_N \\ \vdots & \vdots & \cdots & \\ \varepsilon_1^{N-1} & \varepsilon_2^{N-1} & & \varepsilon_N^{N-1} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \\ \\ \alpha_N \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The coefficient matrix has the non-zero Vandermonde determinant

$$V(\varepsilon_1,\ldots,\varepsilon_j,\ldots,\varepsilon_N) = \prod_{p>q}^{N} (\varepsilon_p - \varepsilon_q) \ .$$

Therefore (5.1.1) has the unique solution

$$\alpha_j = \frac{V(\varepsilon_1,\ldots,0,\ldots,\varepsilon_N)}{V(\varepsilon_1,\ldots,\varepsilon_j,\ldots,\varepsilon_N)} = \prod_{\substack{i=1 \\ i\neq j}}^{N} \frac{\varepsilon_i}{\varepsilon_i - \varepsilon_j}$$

Lemma 5.1.2

Let ε_j , $j=1,\ldots,N$ be distinct non-zero constants, and let

$$\lambda_j = a + \sum_{k=1}^{N-1} b_k \varepsilon_j^k + O(\varepsilon^N), \quad j=1,\ldots,N \tag{5.1.3}$$

where $\varepsilon = \max|\varepsilon_j|$, $a > 0$, $\lambda_j > 0$ and b_k , $k=1,\ldots,N-1$ are constants. Then

$$A \equiv \sum_{j=1}^{N} \alpha_j \lambda_j = a + O(\varepsilon^N) \tag{5.1.4}$$

$$G \equiv \prod_{j=1}^{N} \lambda_j^{\alpha_j} = a + O(\varepsilon^N) \tag{5.1.5}$$

$$H \equiv 1/(\sum_{j=1}^{N} \alpha_j/\lambda_j) = a + O(\varepsilon^N) \tag{5.1.6}$$

where the α_j are defined by (5.1.2).

Proof:

(i) $$A \equiv \sum_{j=1}^{N}\alpha_j\lambda_j = a\sum_{j=1}^{N}\alpha_j + \sum_{k=1}^{N-1} b_k \sum_{j=1}^{N}\alpha_j\varepsilon_j^k + O(\varepsilon^N),$$

on using (5.1.3) in the definition of A.

But the α_j satisfy (5.1.1). Hence $A = a + O(\varepsilon^N)$, which is (5.1.4)

(ii) $$\log G = \sum_{j=1}^{N}\alpha_j\Big[\log a + \log\Big\{1+\sum_{k=1}^{N-1} a^{-1}b_k\varepsilon_j^k + O(\varepsilon^N)\Big\}\Big],$$

on noting $a > 0$, $\lambda_i > 0$,

$$= \sum_{j=1}^{N}\alpha_j \log a + \sum_{j=1}^{N}\alpha_j\Big[\sum_{i=1}^{N-1} c_i\varepsilon_j^{i-1} + O(\varepsilon^N)\Big],$$

the c_i being some constants obtained by the power series expansion of $\log(1+\theta)$.

Thus, $\log G = \log a + O(\varepsilon^N)$, since the α_j satisfy (5.1.1).

Hence $G = a + O(\varepsilon^N)$, which is (5.1.5).

(iii) $$1/H = \sum_{j=1}^{N}\alpha_j/\lambda_j$$

$$= a^{-1}\sum_{j=1}^{N}\alpha_j\Big[1+\sum_{k=1}^{N-1} a^{-1}b_k\varepsilon_j^k + O(\varepsilon^N)\Big]^{-1},$$ noting $a \neq 0$,

$$= a^{-1}\sum_{j=1}^{N}\alpha_j\Big[1+\sum_{i=1}^{N-1}\gamma_i\varepsilon_j^i + O(\varepsilon^N)\Big].$$

the γ_i being some constants obtained by the power series expansion of $(1+\theta)^{-1}$.

Thus, $1/H = a^{-1}[1+O(\varepsilon^N)]$, on using (5.1.1).

Hence $H = a + O(\varepsilon^N)$, which is (5.1.6).

The lemma is proved.

To see how the lemmas are applied, we examine the simplest case first: N=2. Consider possible $O(h^2)$ settings for an end derivative, say d_1.

A Taylor expansion analysis shows that

$$\lambda_1 \equiv \Delta_1 = a + b_1\varepsilon_1 + O(\varepsilon_1^2) > 0$$

$$\lambda_2 \equiv \Delta_{1,3} = a + b_1\varepsilon_2 + O(\varepsilon_2^2) > 0$$

where $a \equiv f_1^{(1)} > 0$, $b_1 \equiv f_1^{(2)}/2$ and $\varepsilon_1 \equiv h_1$, $\varepsilon_2 \equiv h_1+h_2$.
Then taking

$$\alpha_1 = \varepsilon_2/(\varepsilon_2-\varepsilon_1) = 1+ h_1/h_2 \,., \quad \alpha_2 = \varepsilon_1/(\varepsilon_1-\varepsilon_2) = -h_1/h_2$$

in equations (5.1.4) to (5.1.6), we recover the d_1 settings A_1, G_1, H_1 in equations (3.3.2).

For an interior derivative d_i $(i=2,\ldots,n-1)$

$$\lambda_1 \equiv \Delta_{i-1} = a + b_1\varepsilon_1 + O(\varepsilon_1^2) > 0$$

$$\lambda_2 \equiv \Delta_i = a + b_1\varepsilon_2 + O(\varepsilon_2^2) > 0$$

where $a \equiv f_i^{(1)} > 0$, $b_1 \equiv f_i^{(2)}/2$ and $\varepsilon_1 \equiv -h_{i-1}$, $\varepsilon_2 \equiv h_i$.

Then

$$\alpha_1 = h_i/(h_{i-1}+h_i) \,, \quad \alpha_2 = h_{i-1}/(h_{i-1}+h_i) \,,$$

which yield the d_i settings A_i, G_i, H_i previously encountered in equations (3.3.2).

In the next section the technique is used to obtain higher order approximations.

5.2 Explicit $O(h^3)$ conditions

(i) End conditions d_1 and d_n

We use the slopes $\Delta_1, \Delta_{1,3}, \Delta_{1,4}$ and establish that

$$d_1 = \begin{cases} A_1 \equiv \alpha_{1,1}\Delta_1 + \alpha_{2,1}\Delta_{1,3} + \alpha_{3,1}\Delta_{1,4} \\ G_1 \equiv \Delta_1^{\alpha_{1,1}} \cdot \Delta_{1,3}^{\alpha_{2,1}} \cdot \Delta_{1,4}^{\alpha_{3,1}} \\ H_1 \equiv 1/[\alpha_{1,1}/\Delta_1 + \alpha_{2,1}/\Delta_{1,3} + \alpha_{3,1}/\Delta_{1,4}] \end{cases} ,$$

where
$$\alpha_{1,1} = (h_1+h_2)(h_1+h_2+h_3)/[h_2(h_2+h_3)] \ ,$$
$$\alpha_{2,1} = -h_1(h_1+h_2+h_3)/[h_2h_3] \ ,$$
$$\alpha_{3,1} = h_1(h_1+h_2)/[h_3(h_2+h_3)] \ .$$

The approximations for d_n are the duals of these.

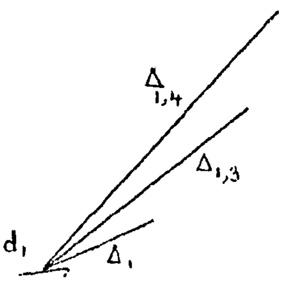


Fig.5.2.1

Proof:

We have, corresponding to Taylor expansions of $\Delta_1, \Delta_{1,3}, \Delta_{1,4}$ in Figure 5.2.1, with $O(h^3)$ remainders, the values

$$\varepsilon_1 = h_1 \ , \qquad \varepsilon_2 = h_1+h_2 \ , \qquad \varepsilon_3 = h_1+h_2+h_3 \ ,$$

hence
$$\alpha_{1,1} = \frac{\varepsilon_2}{\varepsilon_2-\varepsilon_1} \cdot \frac{\varepsilon_3}{\varepsilon_3-\varepsilon_1} = \frac{(h_1+h_2)}{h_2} \cdot \frac{(h_1+h_2+h_3)}{(h_2+h_3)} \ ,$$
$$\alpha_{2,1} = \frac{\varepsilon_1}{\varepsilon_1-\varepsilon_2} \cdot \frac{\varepsilon_3}{\varepsilon_3-\varepsilon_2} = \frac{h_1}{(-h_2)} \cdot \frac{(h_1+h_2+h_3)}{h_3} \ ,$$
$$\alpha_{3,1} = \frac{\varepsilon_1}{\varepsilon_1-\varepsilon_3} \cdot \frac{\varepsilon_2}{\varepsilon_2-\varepsilon_3} = \frac{h_1}{-(h_2+h_3)} \cdot \frac{(h_1+h_2)}{(-h_3)} \ .$$

(ii) Interior conditions d_i

(a) If we use the slopes $\Delta_{i-1}, \Delta_i, \Delta_{i,i+2}$ as in Figure 5.2.2(a), we have, for $i=2,\dots,n-2$,

$$d_i = \begin{cases} A_i \equiv \alpha_{1,i}\Delta_{i-1} + \alpha_{2,i}\Delta_i + \alpha_{3,i}\Delta_{i,i+2} \ , \\ G_i \equiv \Delta_{i-1}^{\alpha_{1,i}} \cdot \Delta_i^{\alpha_{2,i}} \cdot \Delta_{i,i+2}^{\alpha_{3,i}} \ , \\ H_i \equiv 1/[\alpha_{1,i}/\Delta_{i-1} + \alpha_{2,i}/\Delta_i + \alpha_{3,i}/\Delta_{i,i+2}] \ , \end{cases}$$

where
$$\alpha_{1,i} = h_i(h_i+h_{i+1})/[(h_{i-1}+h_i)(h_{i-1}+h_i+h_{i+1})]$$
$$\alpha_{2,i} = h_{i-1}(h_i+h_{i+1})/[(h_{i-1}+h_i)h_{i+1}]$$
$$\alpha_{3,i} = -h_{i-1}h_i/[(h_{i-1}+h_i+h_{i+1})h_{i+1}] .$$

Proof:

Corresponding to Taylor expansions of $\Delta_{i-1}, \Delta_i, \Delta_{i,i+2}$ in Figure 5.2.2(a), we have

$$\varepsilon_1 = -h_{i-1} \ , \ \varepsilon_2 = h_i \ , \ \varepsilon_3 = h_i+h_{i+1} \ ,$$

and these give the results for the $\alpha_{j,i}$.

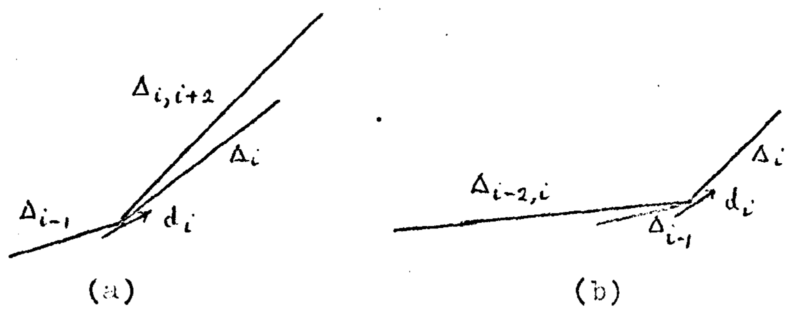


Fig.5.2.2

(b) If, alternatively, we use the slopes $\Delta_i, \Delta_{i-1}, \Delta_{i-2,i}$, we have, for i=3,...,n-1, the duals to those in (a). See Figure 5.2.2(b). We list these for reference.

$$d_i = \begin{cases} \bar{A}_i \equiv \bar{\alpha}_{1,i}\Delta_i + \bar{\alpha}_{2,i}\Delta_{i-1} + \bar{\alpha}_{3,i}\Delta_{i,i-2} \\ \bar{G}_i \equiv \Delta_i^{\bar{\alpha}_{1,i}} \cdot \Delta_{i-1}^{\bar{\alpha}_{2,i}} \cdot \Delta_{i,i-2}^{\bar{\alpha}_{3,i}} \\ \bar{H}_i \equiv 1/[\bar{\alpha}_{1,i}/\Delta_i + \bar{\alpha}_{2,i}/\Delta_{i-1} + \bar{\alpha}_{3,i}/\Delta_{i,i-2}] , \end{cases}$$

where

$$\bar{\alpha}_{1,i} = h_{i-1}(h_{i-2}+h_{i-1})/[(h_{i-1}+h_i)(h_{i-2}+h_{i-1}+h_i)]$$
$$\bar{\alpha}_{2,i} = h_i(h_{i-2}+h_{i-1})/[(h_{i-1}+h_i)h_{i-2}]$$
$$\bar{\alpha}_{3,i} = -h_{i-1}h_i/[(h_{i-2}+h_{i-1}+h_i)h_{i-2}] .$$

5.3 Explicit $O(h^4)$ conditions

In the rational quadratic, $O(h^4)$ settings of the d_i are not required for $O(h^4)$ accuracy, but their use is to be preferred to the $O(h^3)$ settings in the previous section, since they are, in fact, symmetrical forms about the points x_i, for $i=3,\ldots,n-2$. We obtain, for $i=3,\ldots,n-2$:

$$d_i = \begin{cases} A_i \equiv \alpha_{1,i}\Delta_{i,i-2} + \alpha_{2,i}\Delta_{i-1} + \alpha_{3,i}\Delta_i + \alpha_{4,i}\Delta_{i,i+2} \\ G_i \equiv \Delta_{i,i-2}^{\alpha_{1,i}} \cdot \Delta_{i-1}^{\alpha_{2,i}} \cdot \Delta_i^{\alpha_{3,i}} \cdot \Delta_{i,i+2}^{\alpha_{4,i}} \\ H_i \equiv 1/[\alpha_{1,i}/\Delta_{i,i-2} + \alpha_{2,i}/\Delta_{i-1} + \alpha_{3,i}/\Delta_i + \alpha_{4,i}/\Delta_{i,i+2}] \end{cases},$$

where

$$\alpha_{1,i} = -h_{i-1}h_i(h_i+h_{i+1})/[(h_{i-2}+h_{i-1}+h_i)(h_{i-2}+h_{i-1}+h_i+h_{i+1})h_{i-2}]$$
$$\alpha_{2,i} = h_i(h_i+h_{i+1})(h_{i-2}+h_{i-1})/[(h_{i-1}+h_i)(h_{i-1}+h_i+h_{i+1})h_{i-2}]$$
$$\alpha_{3,i} = h_{i-1}(h_i+h_{i+1})(h_{i-2}+h_{i-1})/[(h_{i-1}+h_i)(h_{i-2}+h_{i-1}+h_i)h_{i+1}]$$
$$\alpha_{4,i} = -h_{i-1}h_i(h_{i-2}+h_{i-1})/[(h_{i-1}+h_i+h_{i+1})(h_{i-2}+h_{i-1}+h_i+h_{i+1})h_{i+1}].$$

Proof:

The values $\varepsilon_1 = -(h_{i-2}+h_{i-1})$, $\varepsilon_2 = -h_{i-1}$, $\varepsilon_3 = h_i$, $\varepsilon_4=(h_i+h_{i+1})$ give the four values of the $\alpha_{j,i}$, by application of equation (5.1.2).

5.4 Example. High order accuracy in the case of equal intervals

The arithmetic, geometric and harmonic approximations of sections 5.2 and 5.3 give results which, in the special case of equal intervals, are worth stating separately. Again, we assume monotonic increasing data throughout. In the three methods below, the end conditions for d_1 and d_n are the $O(h^3)$ settings of section 5.2(i). The settings for d_2 and d_{n-1} are the $O(h^3)$ settings of section 5.2.(ii). If any of the approximations obtained from the arithmetic or harmonic formulae are negative they are reset to zero.

We also note that there is the possibility of zero or near zero denominators on the harmonic settings. For d_i, $i=3,\ldots,n-2$, the use of symmetric $O(h^4)$ approximations is preferred, as given in section 5.3.

Thus we have, for equal intervals:

Method 1 ("A" settings)

We set
$$d_1 = \begin{cases} A_1 \equiv 3(\Delta_1 - \Delta_{1,3}) + \Delta_{1,4}, & \text{if } A_1 > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$d_2 = \begin{cases} A_2 \equiv \frac{1}{3}(\Delta_1 - \Delta_{2,4}) + \Delta_2, & \text{if } A_2 > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$d_i = \begin{cases} A_i \equiv \frac{2}{3}(\Delta_{i-1} + \Delta_i) - \frac{1}{6}(\Delta_{i,i-2} + \Delta_{i,i+2}), & \text{if } A_i > 0 \\ 0 & \text{otherwise} \end{cases} \quad (i=3,\ldots,n-2)$$

$$d_{n-1} = \begin{cases} A_{n-1} \equiv \frac{1}{3}(\Delta_{n-1} - \Delta_{n-1,n-3}) + \Delta_{n-2}, & \text{if } A_{n-1} > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$d_n = \begin{cases} A_n \equiv 3(\Delta_{n-1} - \Delta_{n,n-2}) + \Delta_{n,n-3}, & \text{if } A_n > 0 \\ 0 & \text{otherwise} \end{cases}$$

Method 2 ("G" settings)

We set
$$d_1 = \begin{cases} G_1 \equiv (\Delta_1/\Delta_{1,3})^3 \Delta_{1,4}, & \text{if } \Delta_{1,3} > 0 \\ 0 & \text{if } \Delta_{1,3} = 0 \end{cases}$$

$$d_2 = \begin{cases} G_2 \equiv (\Delta_1/\Delta_{2,4})^{1/3} \cdot \Delta_2, & \text{if } \Delta_{2,4} > 0 \\ 0 & \text{if } \Delta_{2,4} = 0 \end{cases}$$

$$d_i = \begin{cases} G_i \equiv (\Delta_{i-1}\Delta_i)^{2/3} / (\Delta_{i,i-2} \cdot \Delta_{i,i+2})^{1/6} & \text{if } \Delta_{i,i-2} \cdot \Delta_{i,i+2} \neq 0 \\ 0 & \text{if } \Delta_{i,i-2} \cdot \Delta_{i,i+2} = 0 \end{cases} \quad (i=3,\ldots,n-2)$$

$$d_{n-1} = \begin{cases} G_{n-1} \equiv (\Delta_{n-1}/\Delta_{n-1,n-3})^{1/3} \Delta_{n-2}, & \text{if } \Delta_{n-1,n-3} > 0 \\ 0 & \text{if } \Delta_{n-1,n-3} = 0 \end{cases}$$

$$d_n = \begin{cases} G_n \equiv (\Delta_{n-1}/\Delta_{n,n-2})^3 \Delta_{n,n-3} & \text{if } \Delta_{n,n-2} > 0 \\ 0 & \text{if } \Delta_{n,n-2} = 0. \end{cases}$$

Method 3 ("H" settings)

We set $d_1 = \begin{cases} H_1 \equiv 1/[3(\Delta_1^{-1} - \Delta_{1,3}^{-1}) + \Delta_{1,4}^{-1}] \text{ , if } H_1 > 0 \\ 0 \text{ otherwise} \end{cases}$

$$d_2 = \begin{cases} H_2 \equiv 1/[\tfrac{1}{3}(\Delta_1^{-1} - \Delta_{2,4}^{-1}) + \Delta_2^{-1}], \text{ if } H_2 > 0 \\ 0 \text{ otherwise} \end{cases}$$

$$d_i = \begin{cases} H_i \equiv 1/[\tfrac{2}{3}(\Delta_{i-1}^{-1} + \Delta_i^{-1}) - \tfrac{1}{6}(\Delta_{i,i-2}^{-1} + \Delta_{i,i+2}^{-1})] \text{ if } H_i > 0 \\ 0 \text{ otherwise} \end{cases} \quad (i=3,\ldots,n-2)$$

$$d_{n-1} = \begin{cases} H_{n-1} \equiv 1/[\tfrac{1}{3}(\Delta_{n-1}^{-1} - \Delta_{n-1,n-3}^{-1}) + \Delta_{n-2}^{-1}] \text{ if } H_{n-1} > 0 \\ 0 \text{ otherwise} \end{cases}$$

$$d_n = \begin{cases} H_n \equiv 1/[3(\Delta_{n-1}^{-1} - \Delta_{n,n-2}^{-1}) + \Delta_{n,n-3}^{-1}] \text{ if } H_n > 0 \\ 0 \text{ otherwise} \end{cases}$$

5.5 Test results and discussion

Using $O(h^3)$ end conditions for d_1, d_n , $O(h^3)$ conditions for d_2, d_{n-1} , and $O(h^4)$ conditions for d_i , $i=3,\ldots,n-2$, we carry out tests on the exponential function (MC2), as in previous chapters, to determine the order of error of the schemes. Table 5.5.1 shows the results for the arithmetic, geometric and harmonic explicit settings.

Method	Error E_1 (h=0.2)	Error E_2 (h=0.1)	Error E_3 (h=0.05)	Error E_3 (h=0.025)	E_1/E_2	E_2/E_3	$E_3/$[illegible]
d_i "A" settings	$.5058\times10^{-4}$	$.3528\times10^{-5}$	$.2331\times10^{-6}$	$.1498\times10^{-7}$	14.34	15.14	15.[illegible]
d_i "G" settings	$.1036\times10^{-4}$	$.6774\times10^{-6}$	$.4329\times10^{-6}$	$.2756\times10^{-8}$	15.29	15.65	15.[illegible]
d_i "H" settings	$.9724\times10^{-5}$	$.6557\times10^{-6}$	$.4258\times10^{-7}$	$.2713\times10^{-8}$	14.83	15.40	15.[illegible]

Table 5.5.1 Rational quadratic interpolation errors
$E = \|f - s\|_\infty$; $O(h^4)$ derivative approximations
on exponential data (MC2)

The table gives the uniform norm error $\|f - s\|_{\infty}$ on $[0,1]$ for each of the choices of h, and the ratios of the errors confirm the expected $O(h^4)$ order of convergence. We should note that both the geometric and harmonic settings have given smaller error norms than those with the arithmetic setting. Our experiments on the practical data sets below would seem to confirm that the geometric and harmonic choices of derivative settings are to be generally preferred.

We have used the five data sets (M2), (M3), (M4), (M5), (MC3), and now comment , briefly, on the graphs, individually.

(M2) data

Using the "A", "G" and "H" settings, we can predict populations of 202.5, 202.3 and 201.5 million for the year 1965. These values are in closed agreement with our previous figures.

(M3) data

The "A" settings for the d_i give poor results; the "G" settings give only slightly better results. Only the "H" settings for the d_i produce an acceptable curve.

(M4) data

Results using the "G" and "H" approximations for d_i are preferred.

(M5) data

Derivative graphs of this Normal distribution function show that the "H" settings provide the best curve.

(MC3) data

Only the "H" settings give an acceptable curve, but even here an unwanted inflexion point has occurred.

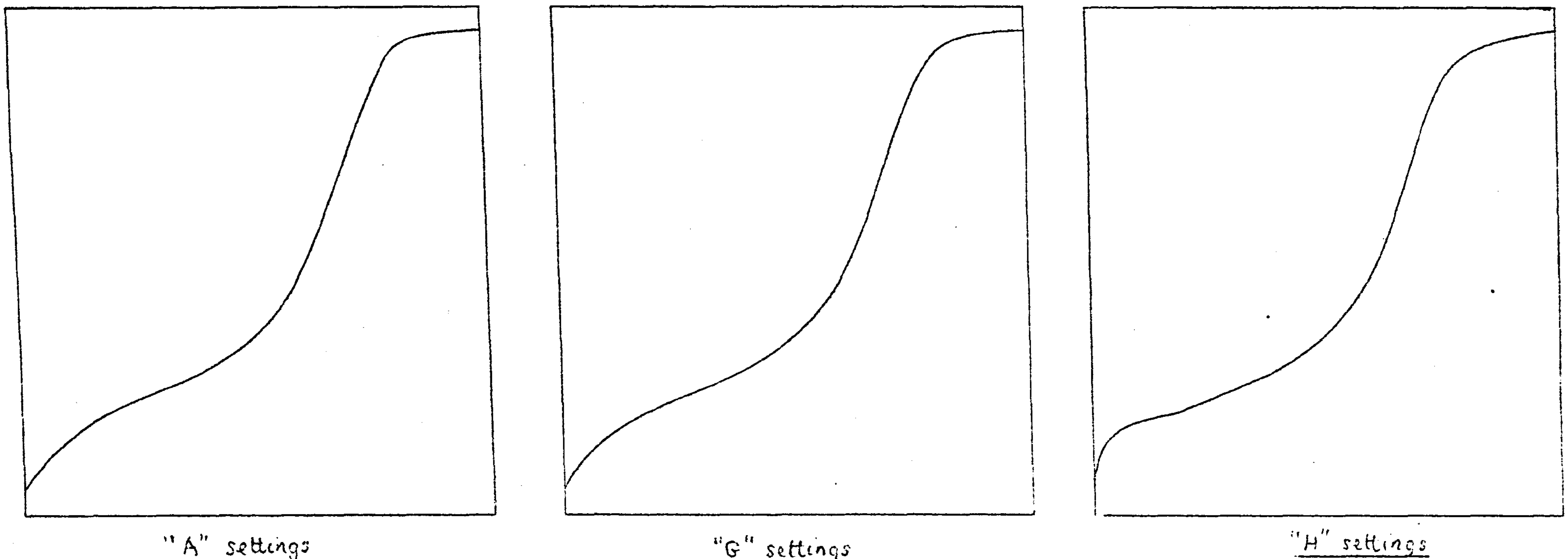


FIG. 5.5.1

$O(h^4)$ Explicit derivative approximations
RATIONAL QUADRATIC
(M2) data

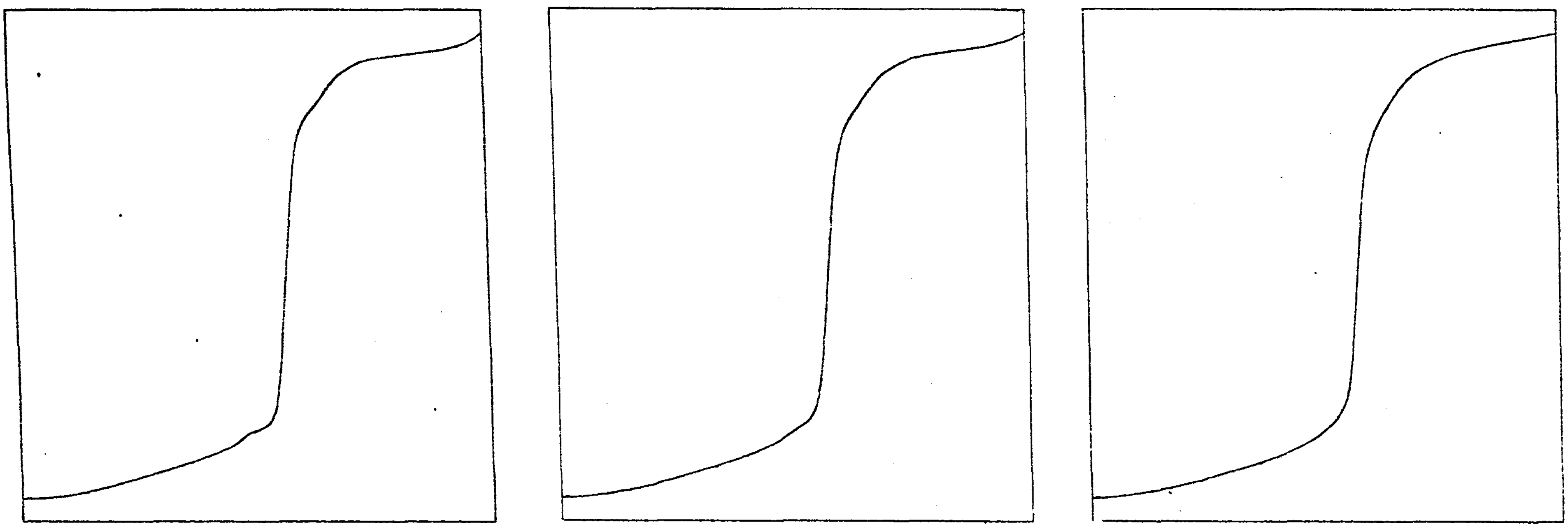
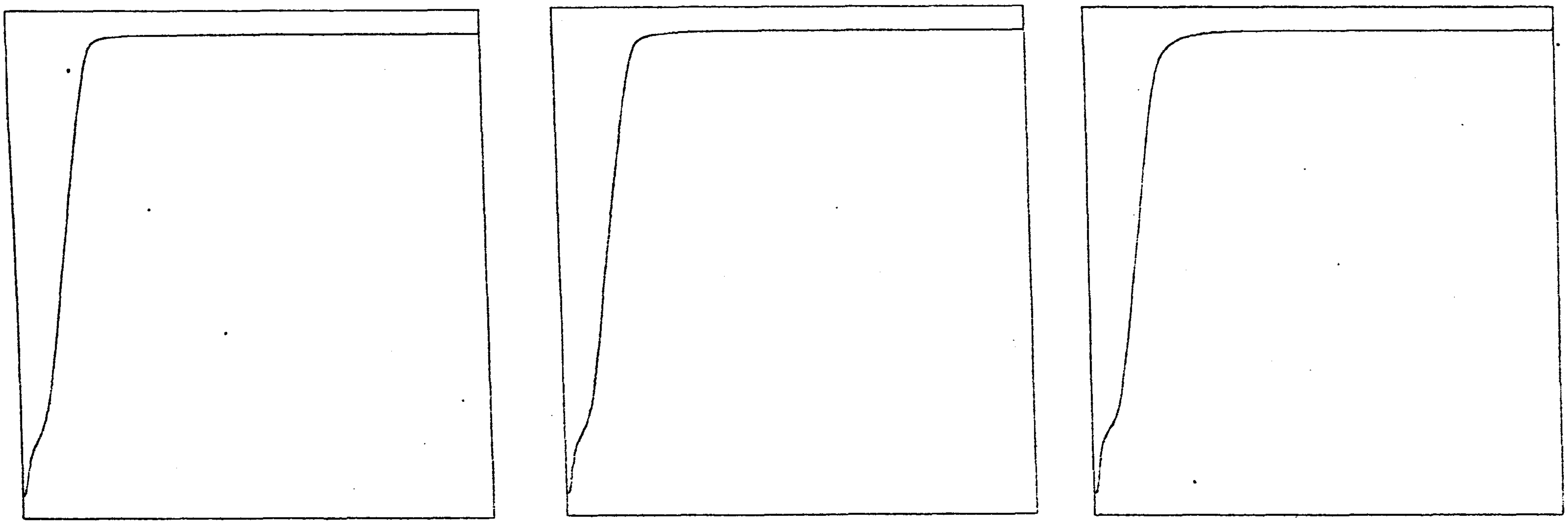


FIG. 5.5.2

$O(h^4)$ Explicit derivative approximations

RATIONAL QUADRATIC

(M3) data



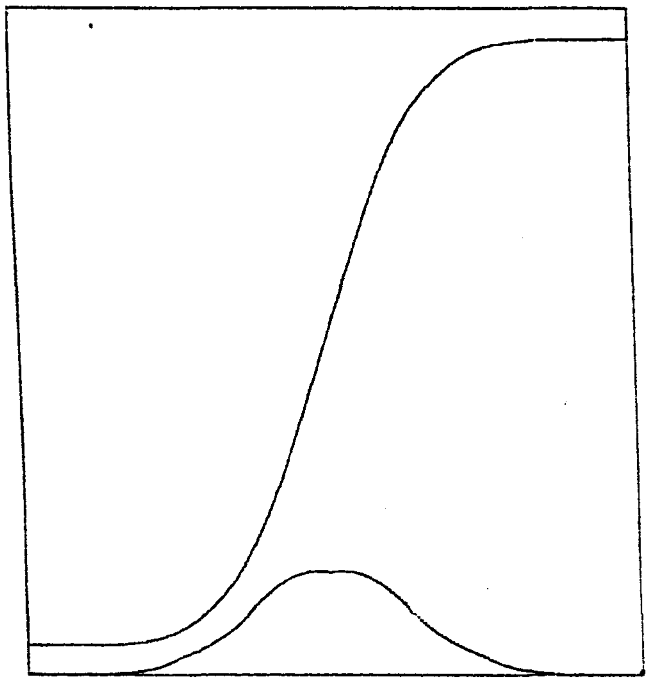
"A" settings "G" settings "H" settings

FIG. 5.5.3

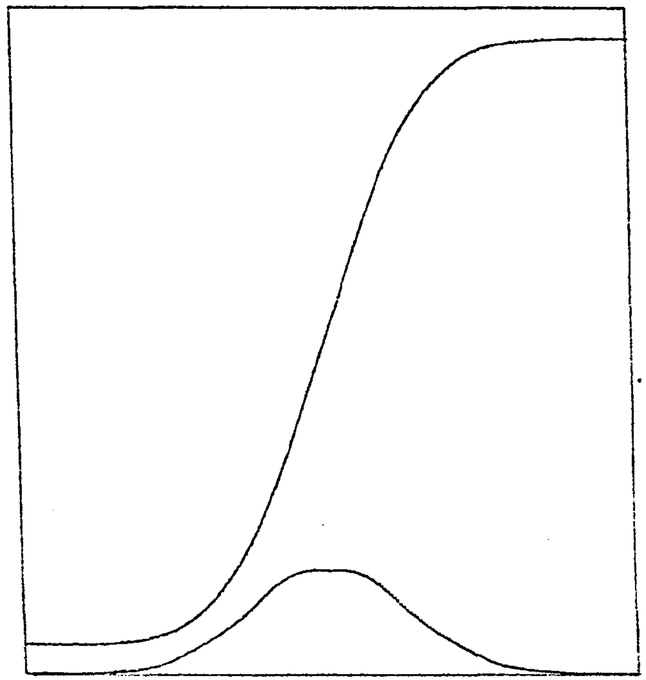
$O(h^4)$ Explicit derivative approximations

RATIONAL QUADRATIC

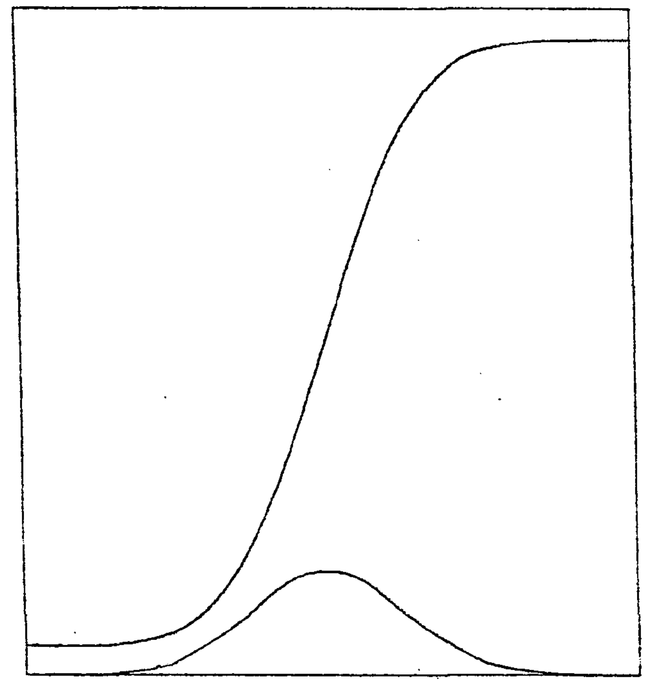
(M4) data



"A" settings



"G" settings



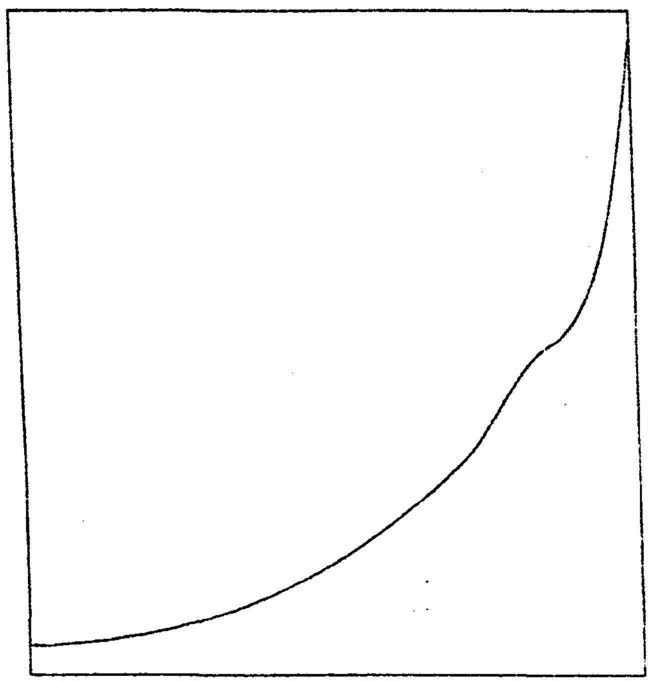
"H" settings

FIG. 5.5.4

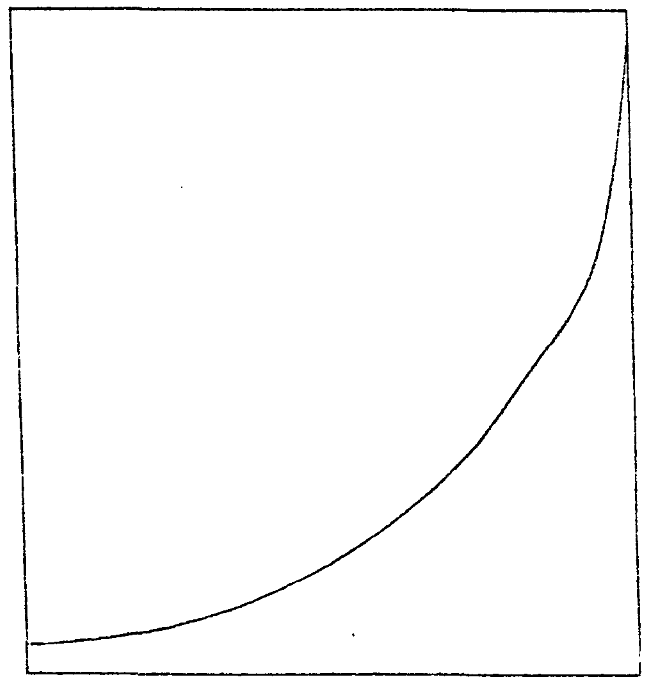
$O(h^4)$ Explicit derivative approximations

RATIONAL QUADRATIC

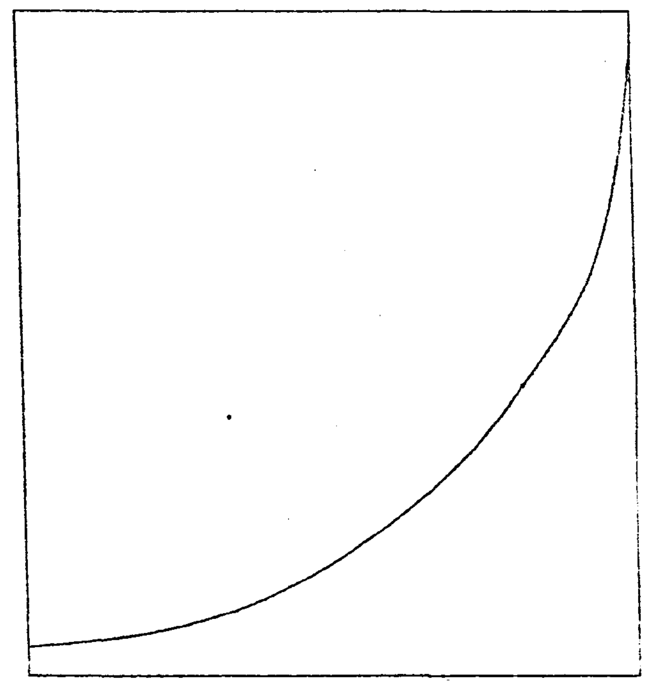
(M5) data



"A" settings



"G" settings



"H" settings

FIG. 5.5.5

$O(h^4)$ Explicit derivative approximations

RATIONAL QUADRATIC

(MC3) data

Chapter 6

ACCURATE DERIVATIVE ESTIMATION FOR MONOTONIC RATIONAL QUADRATIC INTERPOLATION: AN IMPLICIT METHOD

Our purpose in this chapter is to consider again data which is strictly increasing, and the possibility of generating a system of equations for the d_i which will lead to accurate positive derivative estimates. We attempt to produce two-term recurrence schemes by first considering a generalisation of the C^2 consistency equations (4.1.2)

6.1 Generalised consistency equations

We generalise equations (4.1.2) to the form

$$d_i[-c_i+p_i d_{i-1}+(p_i+q_i)d_i+q_i d_{i+1}] = b_i \quad , \quad i=2,\ldots,n-1, \qquad (6.1.1)$$

where

$$b_i = \Delta_{i-1}/h_{i-1} + \Delta_i/h_i \quad ,$$

$$c_i = \hat{c}_i(1/h_{i-1}+1/h_i) \quad ,$$

$$p_i = \hat{p}_i/(h_{i-1}\Delta_{i-1}) \quad ,$$

$$q_i = \hat{q}_i/(h_i\Delta_i) \quad . \qquad (6.1.2)$$

Since the data are increasing, $\Delta_{i-1} > 0$, $\Delta_i > 0$, and we take $\hat{c}_i$, $\hat{p}_i$, $\hat{q}_i$ as free, non-dimensional parameters. The special case $\hat{c}_i = \hat{p}_i = \hat{q}_i = 1$ $(i=2,\ldots,n-1)$ gives the original C^2 consistency equations of Chapter 4. Solved by iteration, they give positive $O(h^3)$ estimates d_i. More generally, however, we seek to examine different choices of $\hat{c}_i$, $\hat{p}_i$, $\hat{q}_i$ which avoid the iterative solution of the equations while still preserving the order of accuracy of the d_i. The error analysis of (6.1.1) is carried out in a similar manner to that for the C^2 equations.

We define E_i by

$$f_i^{(1)}[-c_i+p_i f_{i-1}^{(1)}+(p_i+q_i)f_i^{(1)}+q_i f_{i+1}^{(1)} + E_i] = b_i \quad ,i=2,\ldots,n-1. \qquad (6.1.3)$$

Then we have the following two lemmas, leading to the main result, Theorem 6.1.1.

Lemma 6.1.1

Let $f \in C^1[x_1,x_n]$ and $f^{(1)}(x) > 0$ on $[x_1,x_n]$. Let $d_1 = f_1^{(1)}, d_n = f_n^{(1)}$, p_i, $q_i \geq 0$ and assume there exist d_i, $i=2,\ldots,n-1$ satisfying (6.1.1). Then for some j, $2 \leq j \leq n-1$,

$$\max_{2\leq i\leq n-1} |d_i - f_i^{(1)}| \leq \frac{f_j^{(1)}|E_j| h}{2m/\|f^{(1)}\|_\infty - |E_j| h} , \qquad (6.1.4)$$

where $m = \min_{[x_1,x_n]} f^{(1)}(x) > 0$, $h = \max h_i$ and the denominator in (6.1.4) is assumed positive.

Proof:

Divide (6.1.1) by d_i, and (6.1.3) by $f_i^{(1)}$, and subtract to give

$$[b_i/\{f_i^{(1)}(f_i^{(1)}+\lambda_i)\} + p_i + q_i]\ \lambda_i = E_i - p_i\lambda_{i-1} - q_i\lambda_{i+1}; i=2,\ldots,n-1 \qquad (6.1.5)$$

where $\lambda_i = d_i - f_i^{(1)}$. (6.1.6)

Let $i=j$ be chosen in (6.1.5) so that

$$|\lambda_j| = \max_{2\leq i\leq n-1} |\lambda_i| = \max_{1\leq i\leq n} |\lambda_i| ,$$

since $\lambda_1 = \lambda_n = 0$.

Then, since $0 < f_j^{(1)} + \lambda_j \leq f_j^{(1)} + |\lambda_j|$,

$$[b_j/\{f_j^{(1)}(f_j^{(1)}+|\lambda_j|)\} + p_j + q_j].|\lambda_j| \leq E_j + p_j|\lambda_{j-1}| + q_j|\lambda_{j+1}|$$
$$\leq E_j + (p_j+q_j)|\lambda_j| ,$$

reducing to

$$|\lambda_j| \leq \frac{f_j^{(1)}|E_j|}{b_j/f_j^{(1)} - |E_j|} ,$$

assuming the denominator is positive.

Now $b_j/f_j^{(1)} = (\Delta_{j-1}/h_{j-1} + \Delta_j/h_j)/f_j^{(1)}$
$\geq 2m/(h\,\|f^{(1)}\|_\infty)$.

Combining the last two inequalities gives the result (6.1.4)

Lemma 6.1.2

With the notation and assumptions of the last lemma, and for $f \in C^5[x_1, x_n]$,

$$E_i h_{i-1} h_i \Delta_{i-1} \Delta_i = \{f_i^{(1)}\}^2 r_i + \{f_i^{(1)} f_i^{(2)}/2\}(h_i - h_{i-1}) r_i$$
$$- \{(f_i^{(2)})^2/4\} h_{i-1} h_i r_i$$
$$+ \{f_i^{(1)} f_i^{(3)}/6\}.[(h_i^2 + h_{i-1}^2) r_i + h_i h_{i-1}(h_{i-1} + h_i - h_{i-1}\hat{p}_i - h_i \hat{q}_i)]$$
$$+ \{f_i^{(1)} f_i^{(4)}/24\}.[(h_i^3 - h_{i-1}^3) r_i + h_i h_{i-1}(-h_{i-1}^2 + h_i^2 + 2\hat{p}_i h_{i-1}^2 - 2\hat{q}_i h_i^2)]$$
$$+ \{f_i^{(2)} f_i^{(3)}/12\}.h_{i-1} h_i [(h_{i-1} - h_i) r_i - h_{i-1}^2 + h_i^2 - (\hat{p}_i - \hat{q}_i) h_{i-1} h_i]$$
$$+ O(h^5), \qquad (6.1.7)$$

where
$$r_i = (h_{i-1} + h_i)(1 + \hat{c}_i) - 2h_i \hat{p}_i - 2h_{i-1}\hat{q}_i \qquad (6.1.8)$$

Proof:

From equations (6.1.2), (6.1.3) we find

$$E_i h_{i-1} h_i \Delta_{i-1} \Delta_i = (h_i \Delta_{i-1}^2 \Delta_i + h_{i-1} \Delta_{i-1} \Delta_i^2)/f_i^{(1)}$$
$$+ (h_i + h_{i-1})\hat{c}_i \Delta_{i-1} \Delta_i - h_i \hat{p}_i \Delta_i (f_{i-1}^{(1)} + f_i^{(1)}) - h_i \hat{q}_i \Delta_{i-1}(f_i^{(1)} + f_{i+1}^{(1)}) .$$

On the right we make the Taylor expansions

$$\Delta_{i-1} = f_i^{(1)} - \tfrac{1}{2} h_{i-1} f_i^{(2)} + \tfrac{1}{6} h_{i-1}^2 f_i^{(3)} - \tfrac{1}{24} h_{i-1}^3 f_i^{(4)} + \tfrac{1}{120} h_{i-1}^4 f_{i-\alpha}^{(5)} ,$$
$$\Delta_i = f_i^{(1)} + \tfrac{1}{2} h_i f_i^{(2)} + \tfrac{1}{6} h_i^2 f_i^{(3)} + \tfrac{1}{24} h_i^3 f_i^{(4)} + \tfrac{1}{120} h_i^4 f_{i+\beta}^{(5)} ,$$
$$f_{i-1}^{(1)} = f_i^{(1)} - h_{i-1} f_i^{(2)} + \tfrac{1}{2} h_{i-1}^2 f_i^{(3)} - \tfrac{1}{6} h_{i-1}^3 f_i^{(4)} + \tfrac{1}{24} h_{i-1}^4 f_{i-\gamma}^{(5)}$$
$$f_{i+1}^{(1)} = f_i^{(1)} + h_i f_i^{(2)} + \tfrac{1}{2} h_i^2 f_i^{(3)} + \tfrac{1}{6} h_i^3 f_i^{(4)} + \tfrac{1}{24} h_i^4 f_{i+\delta}^{(5)} ,$$

where $f_{i-\alpha}^{(5)} = f^{(5)}(x_i - \alpha h_{i-1})$, $0 < \alpha < 1$, etc.

The result of these substitutions is (6.1.7).

Theorem 6.1.1

With the assumptions of the two lemmas, if

$$h_i \hat{p}_i + h_{i-1}\hat{q}_i = \tfrac{1}{2}(h_{i-1} + h_i)(1 + \hat{c}_i) \qquad (6.1.9)$$
$$h_{i-1}\hat{p}_i + h_i \hat{q}_i = h_{i-1} + h_i \qquad (6.1.10)$$

then
$$\max_{2 \le i \le n-1} |d_i - f_i^{(1)}| \le \frac{h^3 K(h) \| f^{(1)} \|_\infty}{2m^3/\| f^{(1)} \|_\infty - h^3 K(h)} \qquad (6.1.11)$$

where

$$K(h) = \frac{1}{12}\left[\left(\frac{1}{2}+\max\{\hat{p}_i,\hat{q}_i\}\right)\|f^{(1)}\|_\infty\|f^{(4)}\|_\infty+(1+|\hat{p}_i-\hat{q}_i|)\|f^{(2)}\|_\infty\|f^{(3)}\|_\infty\right] + O(h) \qquad (6.1.12)$$

<u>Proof:</u>

Equations (6.1.9) and (6.1.10) eliminate many of the terms in (6.1.7). Indeed, from (6.1.8), $r_i=0$ and then from (6.1.7) we find

$$E_i\Delta_{i-1}\Delta_i = (f_i^{(1)}f_i^{(4)}/24)(h_i^2-h_{i-1}^2+2\hat{p}_ih_{i-1}^2-2\hat{q}_ih_i^2)$$
$$+(f_i^{(2)}f_i^{(3)}/12)(h_i^2-h_{i-1}^2-(\hat{p}_i-\hat{q}_i)h_{i-1}h_i)+O(h^3).$$

Taking absolute values and using the fact that $\Delta_{i-1}\Delta_i \geq m^2$, then gives

$$|E_i| \leq h^2K(h)/m^2 + O(h^3) ,$$

where K(h) is given in (6.1.12).

Substituting in inequality (6.1.4) of the first lemma then gives (6.1.11), assuming the denominator in (6.1.4) is positive, h being sufficiently small.

<u>Remark 1</u>

If exact end conditions are not given ($d_1 \neq f_1^{(1)}, d_n \neq f_n^{(1)}$) then either

$$\max_{1\leq i\leq n}|\lambda_i| = \max_{2\leq i\leq n-1}|\lambda_i|$$

and the bound (6.1.11) holds for all i=1,...,n; or

$$\max_{1\leq i\leq n}|\lambda_i| = \max\{|\lambda_1|,|\lambda_n|\} .$$

Then d_1, d_n must be chosen as $O(h^3)$ approximations so as to maintain an $O(h^3)$ bound on $\max_{1\leq i\leq n}|d_i-f_i^{(1)}|$. The explicit $O(h^3)$ end conditions of Chapter 5 provide such approximations.

<u>Remark 2</u>

A similar $O(h^3)$ bound to (6.1.11) holds for $f\in C^4[x_1,x_n]$ where the Taylor expansions in the proof of lemma 6.1.2 are taken with $O(h^3)$ remainders. The use of $O(h^4)$ remainders show that if

$$\begin{cases} 2\hat{p}_ih_{i-1}^2 - 2\hat{q}_ih_i^2 = h_{i-1}^2 - h_i^2 & (6.1.13) \\ (\hat{p}_i-\hat{q}_i)h_ih_{i-1} = -h_{i-1}^2 + h_i^2 & (6.1.14) \end{cases}$$

in addition to (6.1.9) and (6.1.10) then the equivalent bound to (6.1.11) becomes an $O(h^4)$ expression. This $O(h^4)$ result is attainable only for the C^2 spline scheme ($\hat{c}_i=\hat{p}_i=\hat{q}_i=1$) and then only for equal intervals.

6.2 Implicit derivative estimates

Theorem 6.1.1 gives an $O(h^3)$ bound on the derivative approximations d_i provided equations (6.1.9), (6.1.10) relating $\hat{c}_i, \hat{p}_i, \hat{q}_i$ are satisfied. Apart from the suitability of the choice $\hat{c}_i=\hat{p}_i=\hat{q}_i=1$ already mentioned, we have the following further choices:

(i) C^1 forward recurrence scheme

Let $\hat{c}_i = 2h_i/h_{i-1} - 1$, $\hat{q}_i = 0$, $\hat{p}_i = 1 + h_i/h_{i-1}$ (6.2.1)

Then solving for the positive root in (6.1.1) gives the non-linear recurrence scheme

$$\left.\begin{aligned} & d_i = -B_i + (B_i^2 + C_i)^{\frac{1}{2}} , \\ \text{where} \quad & B_i = d_{i-1}/2 - \Delta_{i-1}\{1 - h_{i-1}/(2h_i)\} , \\ & C_i = h_{i-1}^2 \Delta_{i-1}(\Delta_{i-1}/h_{i-1} + \Delta_i/h_i)/(h_{i-1}+h_i) . \end{aligned}\right\} \qquad (6.2.2)$$

(ii) C^1 backward recurrence scheme

Let $\hat{c}_i = 2h_{i-1}/h_i - 1$, $\hat{p}_i = 0$, $\hat{q}_i = 1 + h_{i-1}/h_i$ (6.2.3)

Inserting these in (6.1.1) gives the recurrence scheme

$$\left.\begin{aligned} & d_i = -B_i + (B_i^2 + C_i)^{\frac{1}{2}} , \\ \text{where} \quad & B_i = d_{i+1}/2 - \Delta_i\{1 - h_i/(2h_{i-1})\} , \\ & C_i = h_i^2 \Delta_i (\Delta_{i-1}/h_{i-1} + \Delta_i/h_i)/(h_{i-1}+h_i) . \end{aligned}\right\} \qquad (6.2.4)$$

Both (i) and (ii) are new and avoid the use of iteration. They clearly have unique positive solutions, and the methods are stable. The proof of stability is as follows.

Stability of recurrence schemes

Consider, for example, the forward recurrence (6.2.2). When a perturbation δd_{i-1} is made in d_{i-1}, let δd_i be the consequent change in d_i. Since $\delta B_i = \delta d_{i-1}/2$ and $\delta C_i = 0$,

$$\delta d_i = -\delta B_i + \{(B_i+\delta B_i)^2+C_i\}^{\frac{1}{2}} - \{B_i^2+C_i\}^{\frac{1}{2}}$$

$$= \delta B_i\left[-1 + \frac{2B_i+\delta B_i}{\{(B_i+\delta B_i)^2+C_i\}^{\frac{1}{2}}+\{B_i^2+C_i\}^{\frac{1}{2}}}\right]$$

Hence

$$|\delta d_i| \le \frac{1}{2}|\delta d_{i-1}|\cdot\left[1 + \frac{|B_i+\delta B_i| + |B_i|}{\{(B_i+\delta B_i)^2+C_i\}^{\frac{1}{2}}+\{B_i^2+C_i\}^{\frac{1}{2}}}\right]$$

$$\le |\delta d_{i-1}| \quad . \qquad (6.2.5)$$

Thus, a perturbation in the initial setting d_1 is not magnified as we progress through the recurrence to d_{n-1}. Similarly, for the backward recurrence scheme a perturbation in d_n is not magnified through to d_2.

6.3 Test results and discussion

The above discussion on stability suggests that the direction of the recurrence scheme should be chosen to start from the end with the smaller derivative value.

It suggests also that the forward recurrence might be unstable in regions where the chord gradients Δ_i decrease, since, when $d_{i-1} \gg d_i$, large relative perturbations $|\delta d_i/d_i|$ are possible for small relative changes $|\delta d_{i-1}/d_{i-1}|$. Similar considerations apply to the backward recurrence for increasing gradients Δ_i. These remarks are confirmed in practice with the $O(h^3)$ geometric end conditions of section 5.2.

The methods were applied to the data sets (M2),(M3),(M4) and (MC3). Graphs (a) and (b) pertain to the forward and backward recurrence schemes, respectively. For (MC3), graph (b) is unsatisfactory; for (M3), which has both increasing and decreasing regions of Δ_i, (a) is unsatisfactory in the upper region and (b) in the lower region.

A symmetrical approach to the problem, to produce graphs (c), is then adopted which alleviates this behaviour. A weighted average of the two schemes taking into account gradient changes is

$$d_i = d_i^{(f)} \cdot \frac{\Delta_i^p}{\Delta_{i-1}^p + \Delta_i^p} + d_i^{(b)} \cdot \frac{\Delta_{i-1}^p}{\Delta_{i-1}^p + \Delta_i^p}, \quad p \geq 1, \quad (6.3.1)$$

where $d_i^{(f)}$, $d_i^{(b)}$ represent the forward and backward derivative values (at x_i) as given by (6.1.16), (6.1.18).

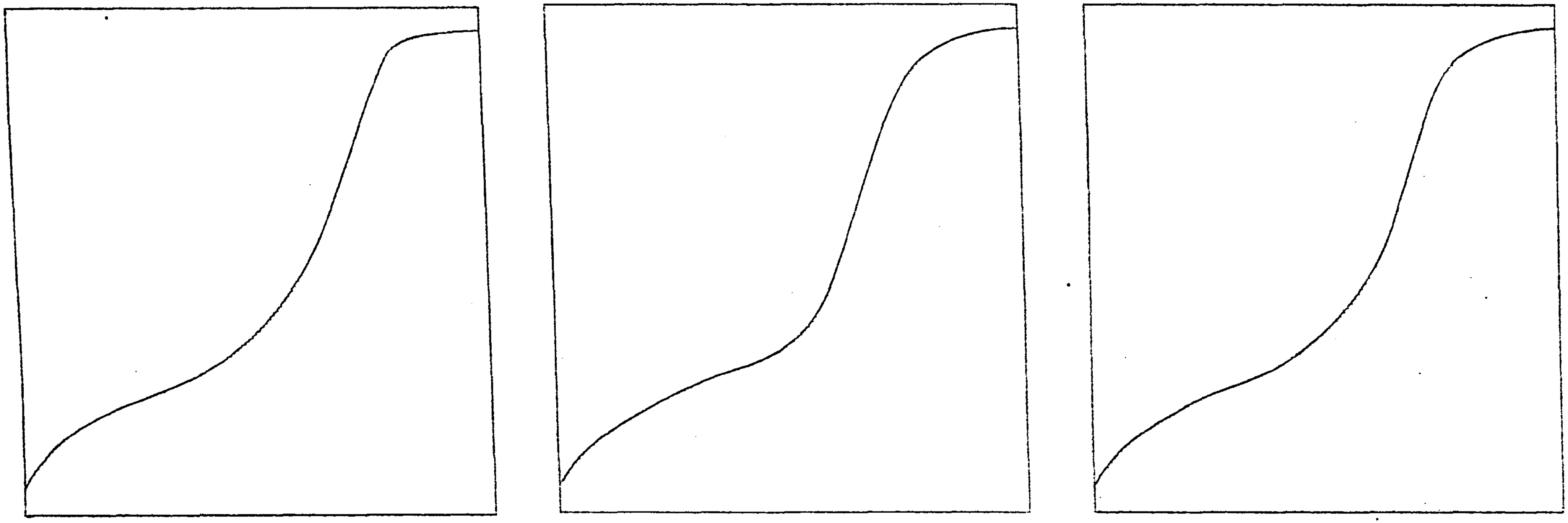
Equation (6.3.1) can be written in a numerically more appropriate form as

$$d_i = \frac{d_i^{(f)} + d_i^{(b)}(\Delta_{i-1}/\Delta_i)^p}{1 + (\Delta_{i-1}/\Delta_i)^p} \quad \text{if} \quad \Delta_{i-1}/\Delta_i \leq 1,$$

or

$$d_i = \frac{d_i^{(f)}(\Delta_i/\Delta_{i-1})^p + d_i^{(b)}}{1 + (\Delta_i/\Delta_{i-1})^p} \quad \text{if} \quad \Delta_{i-1}/\Delta_i \geq 1.$$

Graphs (i),(ii),(iii) show the effects of changing the value of p. We have taken p=1,5,10 in (6.3.1). As $p \to \infty$, $d_i \to d_i^{(f)}$ or $d_i \to d_i^{(b)}$ depending, respectively, on whether $\Delta_{i-1} < \Delta_i$ or $\Delta_{i-1} > \Delta_i$. Thus for large p, equation (6.3.1) suitably taken into account changes in the chord slopes. We have found experimentally that p=10 provided a reasonable choice.



(a) left to right recurrence

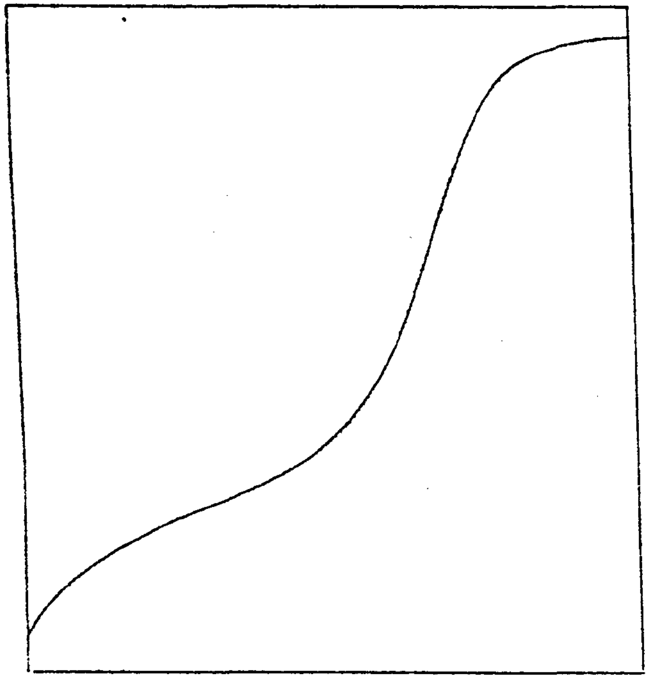
(b) right to left recurrence

(c) weighted average scheme (p = 10)

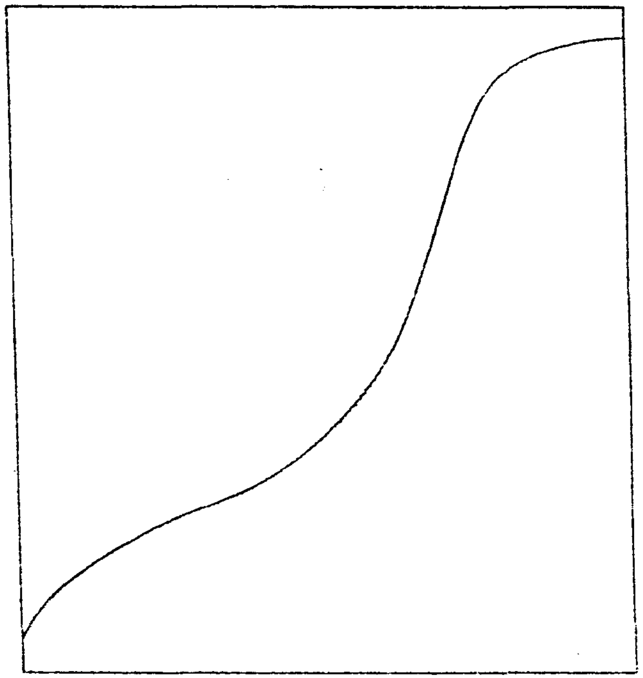
FIG. 6.3.1 RATIONAL QUADRATIC Implicit Scheme

(M2) data

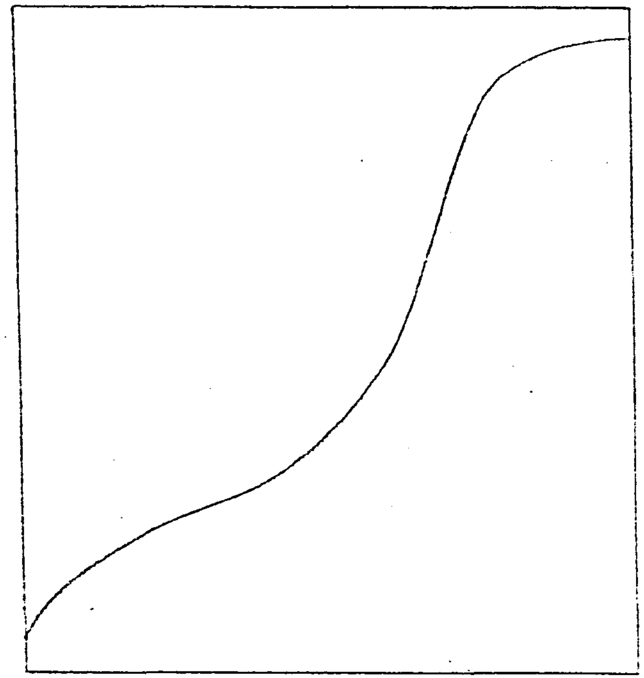
with $O(h^3)$ "G" end derivative approximations



(i) weighted average, p = 1



(ii) weighted average, p = 5

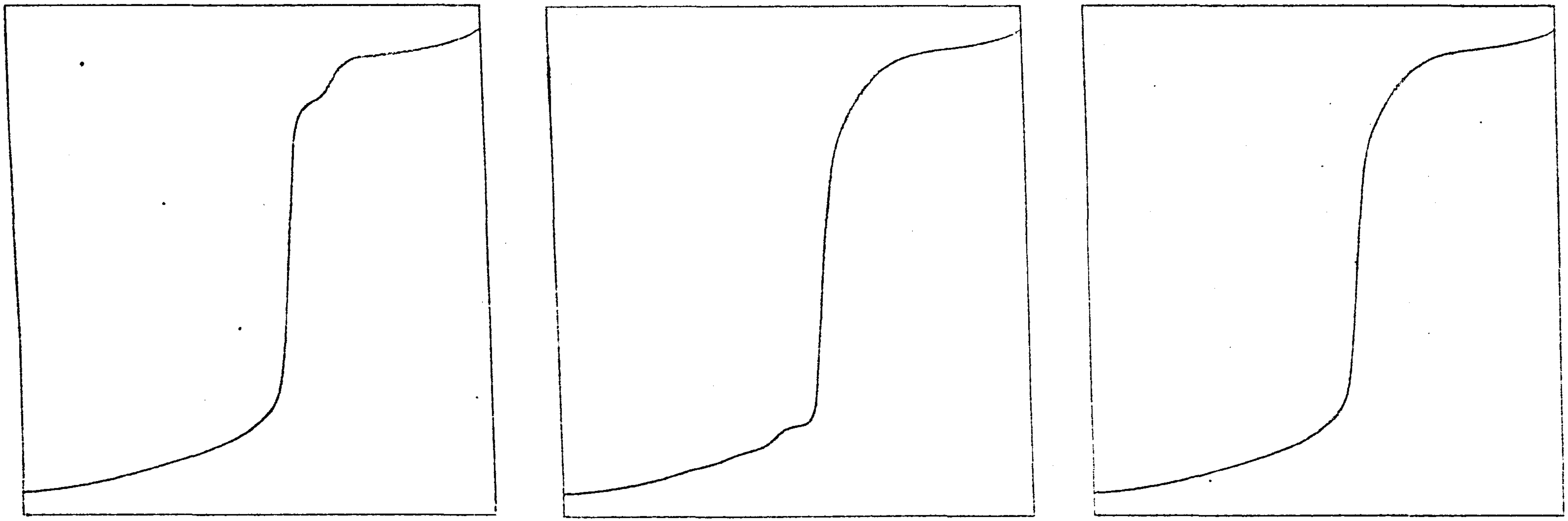


(iii) weighted average, p = 10

FIG. 6.3.2 RATIONAL QUADRATIC Implicit Scheme

(M2) data

with $O(h^3)$ "G" end derivative approximations



(a) left to right recurrence

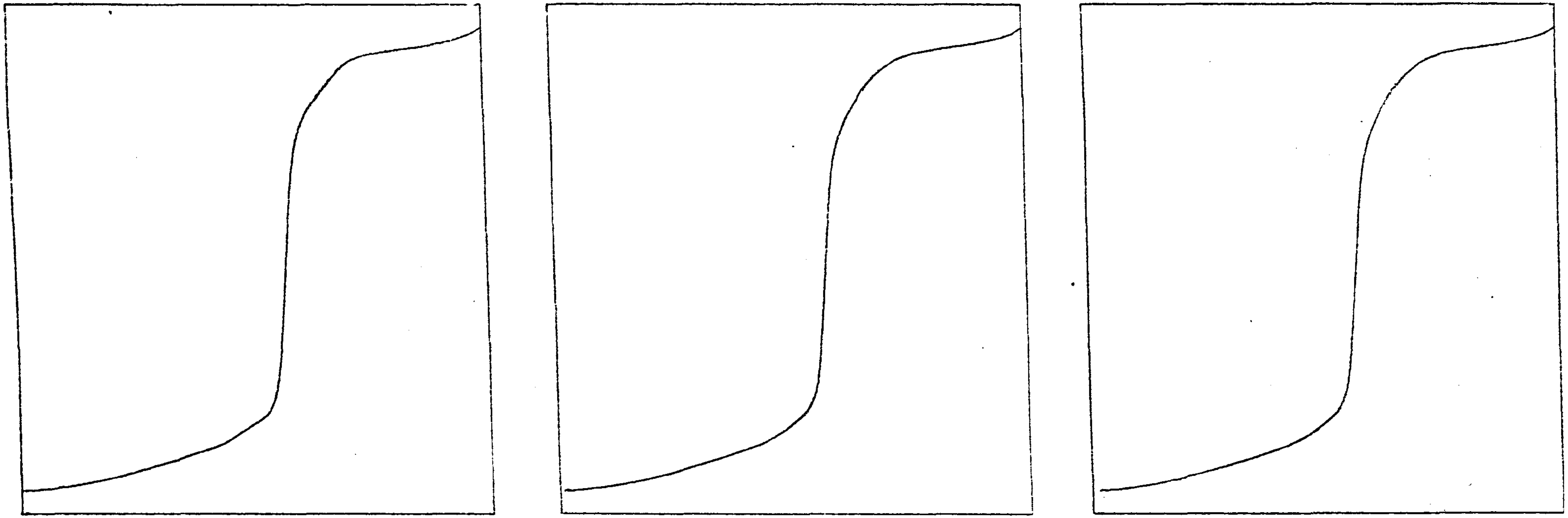
(b) right to left recurrence

(c) weighted average scheme (p = 10)

FIG. 6.3.3 RATIONAL QUADRATIC Implicit Scheme

(M3) data

with $O(h^3)$ "G" end derivative approximations

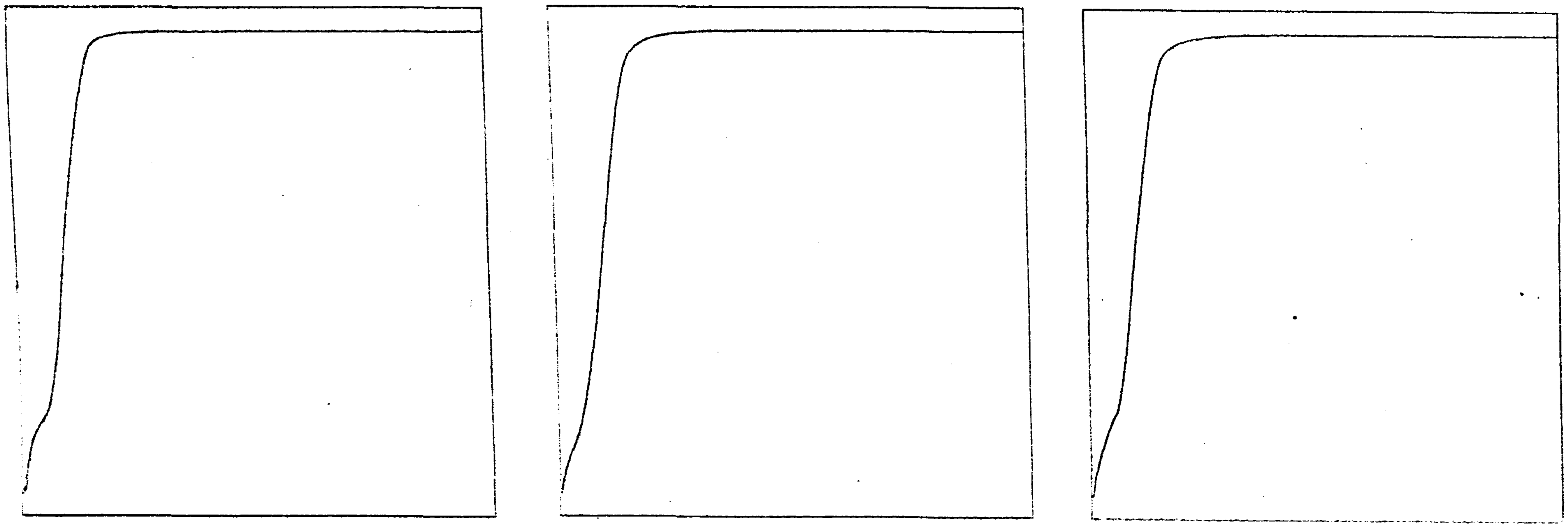


(i) weighted average, $p = 1$ (ii) weighted average, $p = 5$ (iii) weighted average, $p = 10$

FIG. 6.3.4 RATIONAL QUADRATIC Implicit Scheme

(M3) data

with $O(h^3)$ "G" end derivative approximations



(a) left to right recurrence

(b) right to left recurrence

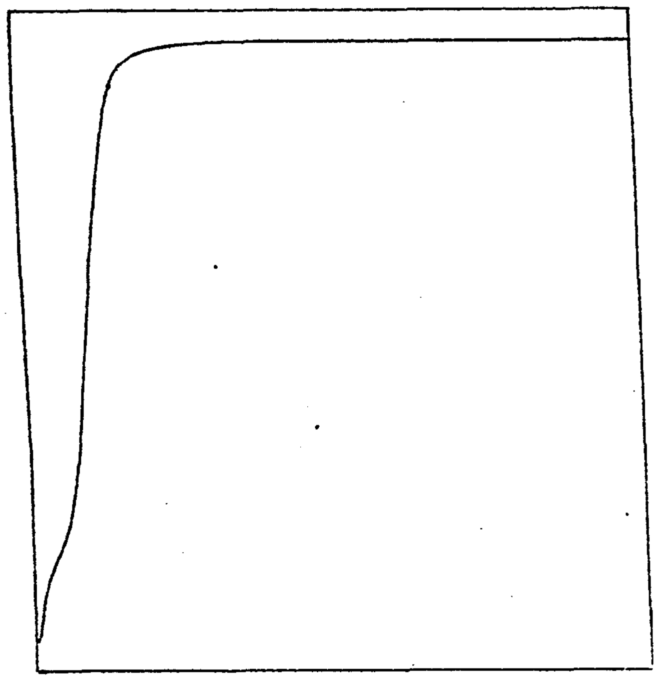
(c) weighted average scheme (p = 10)

FIG. 6.3.5

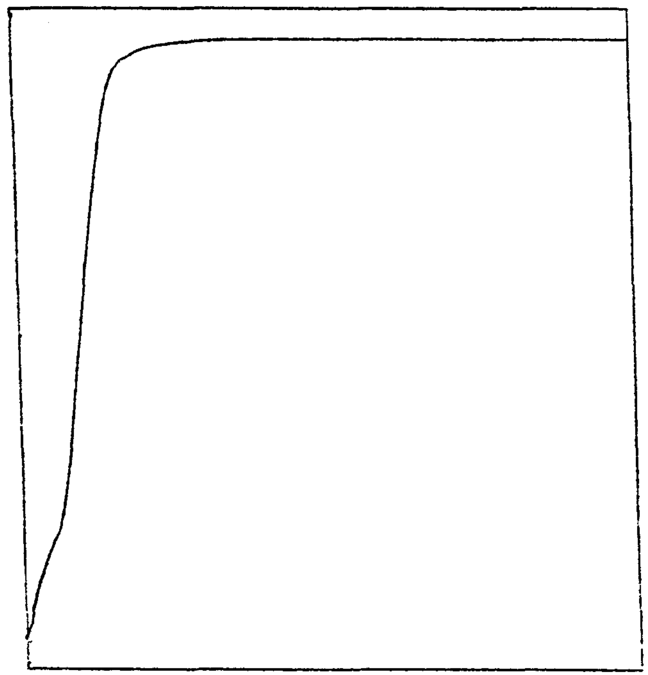
RATIONAL QUADRATIC Implicit Scheme

(M4) data

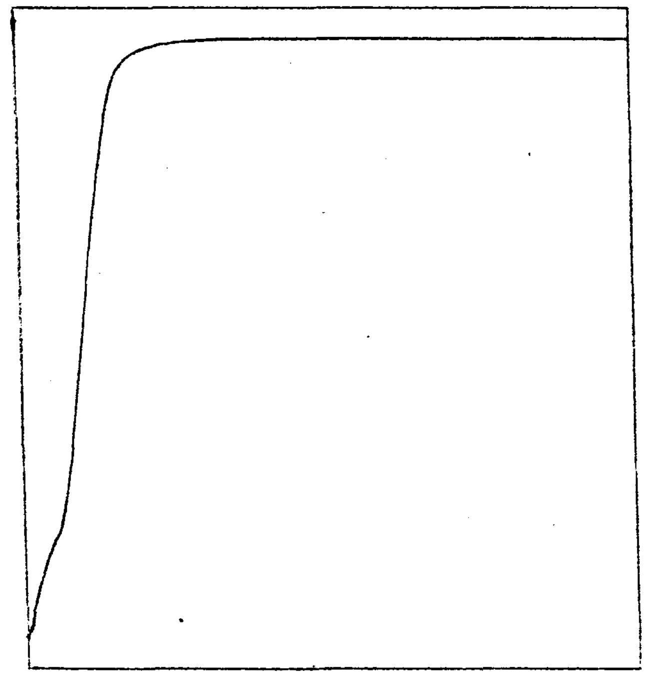
with $O(h^3)$ "G" end derivative approximations



(i) weighted average, p=1.



(ii) weighted average, p=5

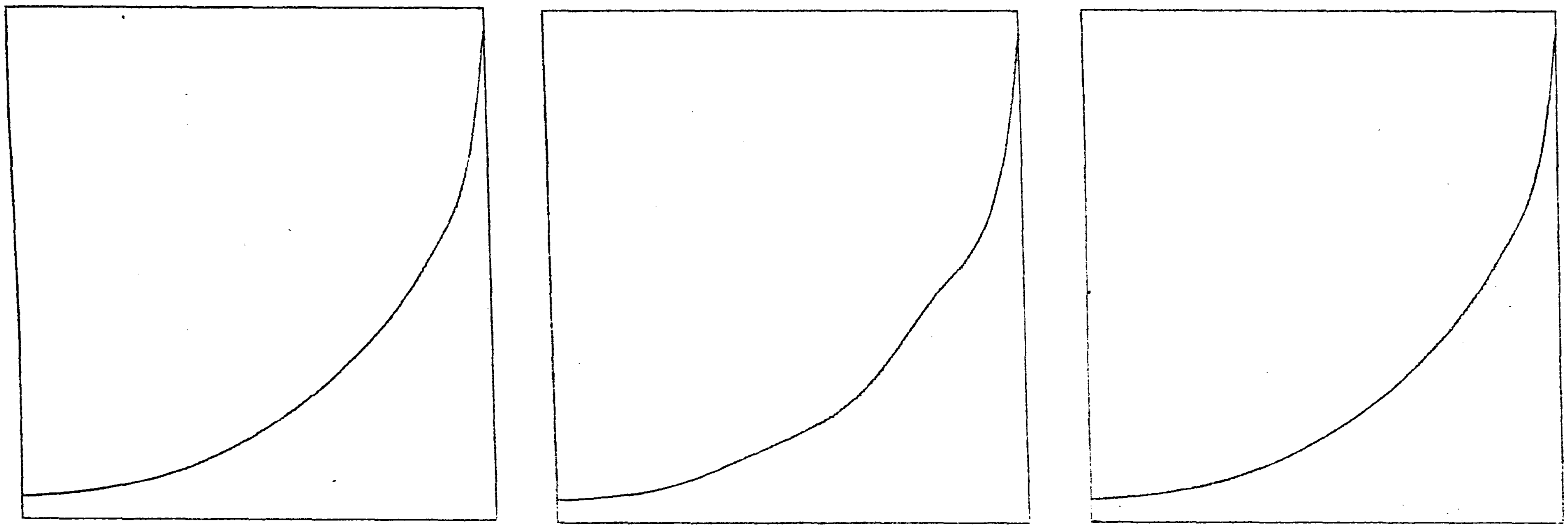


(iii) weighted average, p=10

FIG. 6.3.6 RATIONAL QUADRATIC Implicit Scheme

(M4) data

with $O(h^3)$ "G" end derivative approximations



(a) left to right recurrence

(b) right to left recurrence

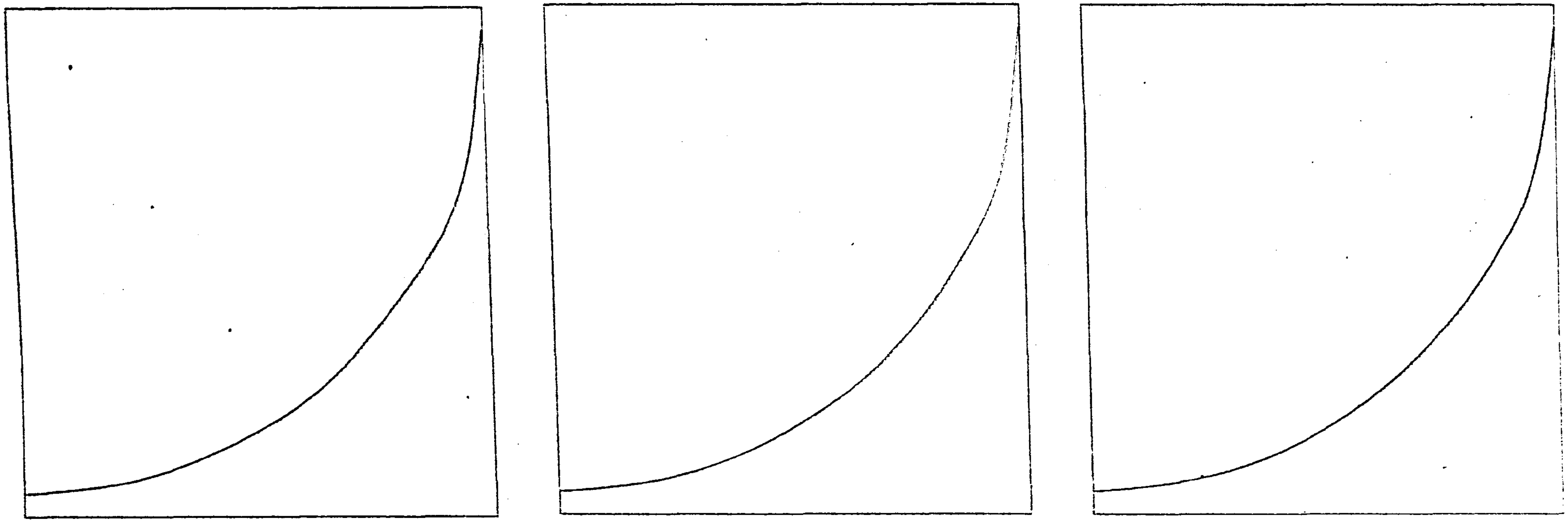
(c) weighted average scheme ($p = 10$)

FIG. 6.3.7

RATIONAL QUADRATIC Implicit Scheme

(MC3) data

with $O(h^3)$ "G" end derivative approximations



(i) weighted average, $p=1$

(ii) weighted average, $p=5$

(iii) weighted average, $p=10$

FIG. 6.3.8

RATIONAL QUADRATIC Implicit Scheme

(MC3) data

with $O(h^3)$ "G" end derivative approximations

Chapter 7

CONVEXITY CONDITIONS FOR A MONOTONIC RATIONAL QUADRATIC

7.1 Convexity conditions

The shape of a monotonic rational quadratic interpolant s between consecutive data points (x_i, f_i), (x_{i+1}, f_{i+1}) depends essentially on the two ratios d_i/Δ_i , d_{i+1}/Δ_i ; see equation (3.1.5). The interpolant may or may not have an inflexion point in (x_i, x_{i+1}). If it has not, and the interpolant is, for example, convex, certain inequalities between the above-mentioned ratios will need to be satisfied. We assume $\Delta_i > 0$. The next theorem makes the relations involved precise.

Theorem 7.1.1

A monotonic increasing rational quadratic interpolant s of the form (3.1.5) is convex in the interval $[x_i, x_{i+1}]$ if and only if

$$\begin{cases} \underline{\text{either:}} & \alpha_i + \beta_i \geq 2 \quad \text{and } 1 \geq \alpha_i(\alpha_i + \beta_i - 1) \\ \underline{\text{or:}} & \alpha_i + \beta_i \leq 2 \quad \text{and } \beta_i(\alpha_i + \beta_i - 1) \geq 1 \end{cases} \tag{7.1.1}$$

where

$$\alpha_i = d_i/\Delta_i , \qquad \beta_i = d_{i+1}/\Delta_i \qquad (\Delta_i > 0) \tag{7.1.2}$$

Proof:

With the notation of equation (7.1.2), we can write $s^{(1)}(x)$ in $[x_i, x_{i+1}]$ in the form

$$s^{(1)}(x) = \frac{\Delta_i\{\alpha_i(1-\theta)^2 + 2\theta(1-\theta) + \beta_i\theta^2\}}{\{(1-\theta)^2 + (\alpha_i + \beta_i)\theta(1-\theta) + \theta^2\}^2} \equiv \frac{\Delta_i q_i(\theta)}{\{Q_i(\theta)\}^2} \text{, say} \tag{7.1.3}$$

A differentiation gives, for $[x_i, x_{i+1}]$,

$$s^{(2)}(x) = \frac{2\Delta_i \cdot \varepsilon_i(\theta)}{h_i\{Q_i(\theta)\}^3}$$

where

$$\varepsilon_i = \tfrac{1}{2}Q_i(\theta)q_i^{(1)}(\theta) - q_i(\theta)Q_i^{(1)}(\theta)$$

$$= \{(1-\theta)^2+(\alpha_i + \beta_i)\theta(1-\theta) + \theta^2\}.\{-\alpha_i(1-\theta)+ (1-2\theta) + \beta_i\theta\}$$

$$-\{\alpha_i(1-\theta)^2 + 2\theta(1-\theta) + \beta_i\theta^2\}.\{-2(1-\theta)+(\alpha_i + \beta_i)(1-2\theta)+2\theta\} .$$

We can write this in the form

$$\varepsilon_i(\theta) = A(1-\theta)^3 + B\theta(1-\theta)^2 + C\theta^2(1-\theta) + D\theta^3 , \qquad (7.1.4)$$

where, after some algebra, we find

$$\begin{aligned} A &= 1 - \alpha_i(\alpha_i + \beta_i - 1) , \\ B &= 3(1-\alpha_i) , \\ C &= 3(\beta_i - 1) , \\ D &= -1 + \beta_i(\alpha_i + \beta_i - 1). \end{aligned} \qquad (7.1.5)$$

Now, $s^{(2)}(x) \geq 0$ throughout $[x_i, x_{i+1}]$ for a convex interpolant. Since $s^{(2)}(x)$ has the same sign as $\varepsilon_i(\theta)$, a convex interpolant exists in $[x_i, x_{i+1}]$ if and only if $\varepsilon_i(\theta) \geq 0$ for all $0 \leq \theta \leq 1$.

A differentiation of (7.1.4) gives the result

$$\varepsilon_i^{(1)}(\theta) = 3(\alpha_i + \beta_i - 2)q_i(\theta) \qquad (7.1.6)$$

where $q_i(\theta)$, given by (7.1.3), is non-negative.

We deduce that $\varepsilon_i(\theta)$ is increasing if and only if $\alpha_i + \beta_i \geq 2$ and decreasing if and only if $\alpha_i + \beta_i \leq 2$.

But $\varepsilon_i(0) = A$ and $\varepsilon_i(1) = D$. Hence

$s^{(2)}(x) \geq 0$ in $[x_i, x_{i+1}]$ if and only if

$$\begin{cases} \underline{\text{either}}: & \alpha_i + \beta_i \geq 2 \quad \text{and} \quad A \geq 0 \\ \underline{\text{or}}: & \alpha_i + \beta_i \leq 2 \quad \text{and} \quad D \geq 0 \end{cases} \qquad (7.1.7)$$

The theorem follows from this and (7.1.5).

<u>Corollary</u>: For each interval $I_i = [x_i, x_{i+1}]$, a diagram of β_i against α_i can be used to interpret the inequalities in (7.1.1). See Figure 7.1.1. This shows the hyperbolas with equations $\beta_i = 1 + 1/\alpha_i - \alpha_i$ and $\alpha_i = 1 + 1/\beta_i - \beta_i$, and the straight line $\alpha_i + \beta_i = 2$. These determine the region R, shaded in the figure.

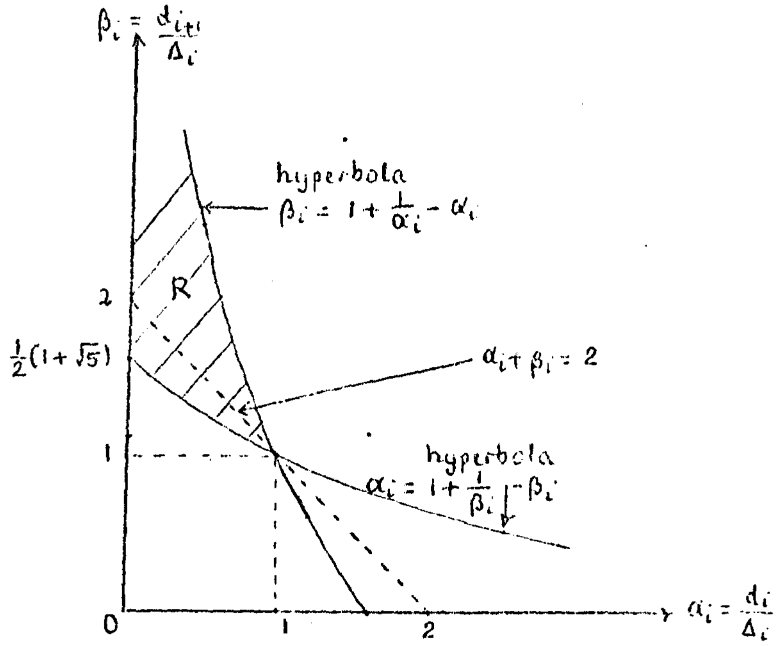


Fig.7.1.1

Provided only that the pair $(\alpha_i, \beta_i) \equiv (d_i/\Delta_i,\ d_{i+1}/\Delta_i)$ lies in the region R, the interpolant determined by them will be convex (and monotonic increasing).

Remark

The (α_i, β_i) notation employed here was used by Fritsch and Carlson in connexion with monotonic cubics.(See Chapter 2). They describe a'monotonicity region S', in the first quadrant of the $\alpha_i - \beta_i$ plane, which includes an elliptical region. Our 'convexity region R' for rational quadratics contrasts with this and involves hyperbolas. It is to be noted, once again, that the pairs (α_i, β_i), $i=2,\dots,n-1$ are not unrelated in the different intervals. For, since $\beta_{i-1} = d_i/\Delta_{i-1}$ and $\alpha_i = d_i/\Delta_i$, we have, in fact,

$$\alpha_i = \beta_{i-1}(\Delta_{i-1}/\Delta_i).$$

Given strictly increasing,convex data $(0 < \Delta_1 < \Delta_2 < \dots)$, a C^2 rational quadratic increasing interpolant, constructed as described in Chapter 4, is not necessarily itself convex. By referring to the convexity region R, one can see how this may easily occur. In terms of the α, β notation introduced, the C^2 consistency equations (4.1.1) are

$$\{(1 + 1/\alpha_i - \alpha_i) - \beta_i\} / h_i = \{\alpha_{i-1} - (1 + 1/\beta_{i-1} - \beta_{i-1})\} / h_{i-1}\ ,$$

$$i=2,\dots,n-1\ .$$

These equations have a simple geometrical interpretation. See Figure 7.1.2.

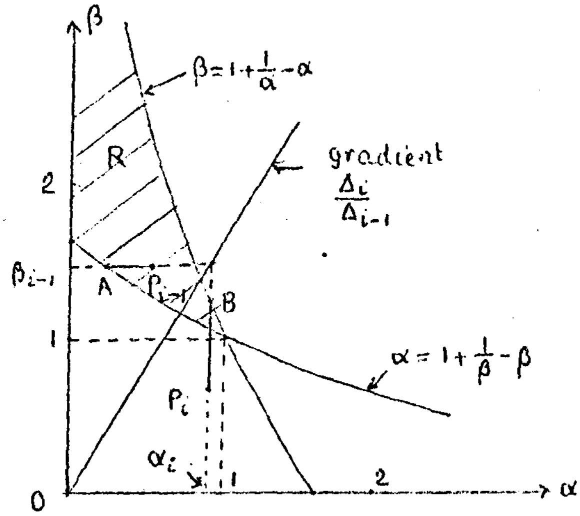


Fig.7.1.2

Consider the points $P_{i-1}(\alpha_{i-1},\beta_{i-1})$, $P_i(\alpha_i,\beta_i)$: α_i is obtained from β_{i-1} by a projection using the line of gradient Δ_i/Δ_{i-1} , and the consistency equations give the result

$$P_iB/h_i = AP_{i-1}/h_{i-1}$$

where the point A on the hyperbola $\alpha = 1 + 1/\beta - \beta$ has the same ordinate as P_{i-1} and point B on the hyperbola $\beta = 1 + 1/\alpha - \alpha$ has the same abscissa as P_i . Clearly, given P_{i-1} in the convexity region R, a sufficiently large value of the ratio h_i/h_{i-1} (for example) will force P_i to lie below the region R.

Chapter 8

SHAPE-PRESERVING INTERPOLATION USING A C^1 RATIONAL CUBIC

Earlier chapters showed how a number of rational quadratic schemes could be used for the construction of monotonic interpolants to given monotonic data. Graphical results indicated that the schemes are quite acceptable, in general. However, in some exceptional cases where data is also convex, unwanted inflexion points have occurred in the curves. As explained in Chapter 7, an inflexion in the interval $[x_i, x_{i+1}]$ is produced if the pair $(d_i/\Delta_i, d_{i+1}/\Delta_i)$ lies outside the convexity region R between two hyperbolas.

This chapter develops a C^1 piecewise rational cubic function which solves the present problem. When applied to convex and/or monotonic data, the rational cubic scheme will produce the required shape-preserving interpolant. Further, it is of interest to note that the scheme contains as a special case the monotonic rational quadratic that we have been considering thus far.

An error analysis is given which shows that the interpolant can be sufficiently accurate.

8.1 The rational cubic interpolant

Data points (x_i, f_i), $i=1,\ldots,n$ are given, where $x_1 < x_2 < \ldots < x_n$. As always, we let

$$h_i = x_{i+1} - x_i ,$$
$$\Delta_i = (f_{i+1} - f_i)/h_i \qquad (8.1.1)$$

A piecewise rational cubic function $s \in C^1[x_1, x_n]$ is defined as follows (derived in a similar manner to the rational quadratic of Section 3.1.)

If $x \in [x_i, x_{i+1}]$, let

$$\theta = (x - x_i)/h_i , \qquad (8.1.2)$$

and

$$s(x) = S(\theta) \equiv \frac{P_i(\theta)}{Q_i(\theta)}$$

$$\equiv \frac{f_{i+1}\theta^3+(r_i f_{i+1}-h_i d_{i+1})\theta^2(1-\theta)+(r_i f_i+h_i d_i)\theta(1-\theta)^2+f_i(1-\theta)^3}{1+(r_i-3)\theta(1-\theta)} \quad (8.1.3)$$

Here r_i is a parameter for the interval $[x_i,x_{i+1}]$, to be chosen so that

$$r_i > -1 \quad (8.1.4)$$

in order to ensure a strictly positive denominator in (8.1.3). This choice of rational cubic form has the following interpolatory properties:

$$s(x_i) = f_i \;, \quad s(x_{i+1}) = f_{i+1}$$
$$s^{(1)}(x_i) = d_i \;, \; s^{(1)}(x_{i+1}) = d_{i+1} \;, \quad (8.1.5)$$

where the d_i denote derivative values at the points x_i.

Remark 1.

If the choice $r_i=3$ is made, then (8.1.3) clearly reduces to the standard cubic Hermite polynomial.

Also, a little algebraic simplification shows that the choice

$$r_i = 1 + (d_i+d_{i+1})/\Delta_i \quad (8.1.6)$$

reduces the rational cubic (8.1.3) to the quadratic form of the previous chapters.

Remark 2.

The interpolant defines a non-linear operator, since the parameters r_i will depend on the data. However, the interpolant to the zero function is zero. Also, the interpolant to the data $K + f_i$, $i=1,..,n$, where K is a constant, is $K + s(x)$, provided the r_i are independent of such translations, and this will be so for all the choices of r_i that we present here.

8.2 Error bound analysis

An error bound for the rational cubic interpolant is given in the next result. Its proof follows the same lines as that of Theorem 3.2.1.

Theorem 8.2.1

Let $f \in C^4[x_1,x_n]$ and let s be the piecewise rational cubic interpolant such that $s(x_i)= f_i$ and $s^{(1)}(x_i) = d_i$, $i=1,\ldots,n$. Then for $x \in [x_i,x_{i+1}]$,

$$|f(x)-s(x)| \leq \frac{h_i}{4c_i} \max\{|\lambda_i|, |\lambda_{i+1}|\}$$

$$+ \frac{1}{384c_i}\{h_i^4 \| f^{(4)}\|_i (1+|r_i-3|/4) +4|r_i-3|(h_i^3\|f^{(3)}\|_i+3h_i^2 \| f^{(2)}\|_i)\} \tag{8.2.1}$$

where

$$\lambda_i= d_i- f_i^{(1)} \quad , \quad \lambda_{i+1}= d_{i+1} -f_{i+1}^{(1)} , \tag{8.2.2}$$

and

$$c_i = \begin{cases} (1+r_i)/4 & \text{if } -1< r_i < 3 \\ 1 & \text{if } r_i \geq 3 \end{cases} \tag{8.2.3}$$

with $\| \cdot \|_i$ denoting the uniform norm on $[x_i,x_{i+1}]$.

Proof:

On $[x_i,x_{i+1}]$, let $x(\theta)= x_i+\theta h_i$, and $F_i(\theta)= f(x(\theta))$.

Then

$$|f(x)-s(x)| = |F_i(\theta)- P_i(\theta)/Q_i(\theta)|$$

$$\leq \frac{|F_i(\theta)Q_i(\theta) -P_i^*(\theta)| + |P_i^*(\theta) -P_i(\theta)|}{|Q_i(\theta)|} \tag{8.2.4}$$

where we take

$$P_i^*(\theta) =f_i(1-\theta)^3+(r_if_i+h_if_i^{(1)})\theta(1-\theta)^2+(r_if_{i+1}-h_if_{i+1}^{(1)})\theta^2(1-\theta)+f_{i+1}\theta^3. \tag{8.2.5}$$

$P_i^*(\theta)$ is the cubic Hermite interpolant to $F_i(\theta)Q_i(\theta)$ on $0\leq \theta \leq 1$, and

$$|F_i(\theta)Q_i(\theta)-P_i^*(\theta)| \le \frac{1}{384} \max_{0\le\theta\le1} \left| \frac{d^4}{d\theta^4} F_i(\theta)Q\ (\theta)\right| ,$$

$$= \frac{1}{384} \max_{0\le\theta\le1} \left| F_i^{(4)}(\theta)Q_i(\theta) + 4F_i^{(3)}(\theta)Q_i^{(1)}(\theta) + 6F_i^{(2)}(\theta)Q_i^{(2)}(\theta)\right|$$

since $Q_i(\theta)$ is quadratic.

Now
$$|Q_i(\theta)| \le 1 + |r_i-3|/4 ,$$
$$|Q_i^{(1)}(\theta)| \le |r_i -3|$$
$$|Q_i^{(2)}(\theta)| = 2|r_i-3| ,$$

and $F_i^{(j)}(\theta) \le h_i^j \|f^{(j)}\|_i$.

Hence

$$|F_i(\theta)Q_i(\theta) -P_i^*(\theta)| \le \frac{1}{384}\{h_i^4\|f^{(4)}\|_i(1+|r_i-3|/4)+4h_i^3\|f^{(3)}\|_i\cdot|r_i-3| + 12h_i^2\|f^{(2)}\|_i|r_i-3|\} \quad (8.2.6)$$

Also, using (8.2.2)

$$|P_i^*(\theta) - P_i(\theta)| = |\theta(1-\theta)h_i[\theta\lambda_{i+1}-(1-\theta)\lambda_i]| ,$$

$$\le \frac{1}{4}h_i \max\{|\lambda_i|, |\lambda_{i+1}|\} \quad . \quad (8.2.7)$$

Finally,

$$|Q_i(\theta)| = Q_i(\theta) \ge \begin{cases} 1 & \text{if } r_i \ge 3 \\ 1-(3-r_i)/4 & \text{if } -1 < r_i < 3 \end{cases} \quad (8.2.8)$$

Using inequalities (8.2.6),(8.2.7),(8.2.8) in (8.2.4), we obtain the result in the theorem.

<u>Corollary 8.2.1</u>

Let $x \in [x_i, x_{i+1}]$,

(i) If $\lambda_i = O(h_i^2)$, $\lambda_{i+1} = O(h_i^2)$ and $r_i-3 = O(h_i)$,
then $|f(x)-s(x)| = O(h_i^3)$.

(ii) If $\lambda_i = O(h_i^3)$, $\lambda_{i+1} = O(h_i^3)$ and $r_i-3 = O(h_i^2)$,
then $|f(x)-s(x)| = O(h_i^4)$.

Ideally, therefore, r_i should be such that $r_i-3 = O(h_i^2)$, for $i=1,\dots,n-1$. In the next two sections, where we consider monotonic data and convex data, we show how r_i-3 can be chosen, with this optimal $O(h^2)$ accuracy, so that the interpolant which results conforms to the shape of the data.

8.3 Shape Preserving Interpolation: Monotonic Data

Without loss of generality, we take the monotonic data to be increasing, so that

$$\Delta_i \geq 0 \ , \ i=1,\dots,n-1 \ . \qquad (8.3.1)$$

For a monotonic increasing interpolant $s(x)$, it is necessary that

$$d_i \geq 0 \quad , \ i=1,\dots,n \ , \qquad (8.3.2)$$

and it is necessary and sufficient that

$$s^{(1)}(x) \geq 0, \text{ for all } x \in [x_1,x_n] \ . \qquad (8.3.3)$$

Now let $x \in [x_i,x_{i+1}]$. A differentiation of (8.1.3) yields the form

$$s^{(1)}(x) = \frac{d_{i+1}\theta^4 + \alpha_i\theta^3(1-\theta) + \beta_i\theta^2(1-\theta)^2 + \gamma_i\theta(1-\theta)^3 + d_i(1-\theta)^4}{[1 + (r_i-3)\theta(1-\theta)]^2} \qquad (8.3.4)$$

where

$$\begin{aligned} \alpha_i &= 2(r_i\Delta_i - d_i) \\ \beta_i &= (r_i^2+3)\Delta_i - r_i(d_i+d_{i+1}) \\ \gamma_i &= 2(r_i\Delta_i - d_{i+1}) \end{aligned} \qquad (8.3.5)$$

Thus sufficient conditions for monotonicity on $[x_i,x_{i+1}]$ are

$$\alpha_i \geq 0, \quad \beta_i \geq 0 \ , \quad \gamma_i \geq 0 \qquad (8.3.6)$$

where the necessary conditions $d_i \geq 0$, $d_{i+1} \geq 0$ are assumed.
If $\Delta_i > 0$ then a sufficient condition for (8.3.6) is

$$r_i \geq (d_i+d_{i+1})/\Delta_i \qquad (8.3.7)$$

If, in particular,

$$r_i = 1 + (d_i+d_{i+1})/\Delta_i \ ,$$

then the rational quadratic form (3.1.5) results, for which $d_i \geq 0$ and $d_{i+1} \geq 0$ are necessary and sufficient conditions for a monotonic increasing interpolant.

Clearly, if $\Delta_i=0$, then $d_i= 0 = d_{i+1}$ and $s(x)= f_i=f_{i+1}$ is a constant on $[x_i,x_{i+1}]$.

Theorem 8.2.1 and its corollary show that, if the choice $r_i= 1+(d_i+d_{i+1})/\Delta_i$ is made, then the optimal $O(h^4)$ bound on the interpolation error can be achieved if d_i and d_{i+1} are chosen with $O(h^3)$ accuracy. This follows from the fact that

$$r_i-3 =(\lambda_i+\lambda_{i+1})/\Delta_i + O(h_i^2) .$$

8.4 Shape-Preserving Interpolation: Convex data

We assume strictly convex data; that is,

$$\Delta_{i-1}< \Delta_i \quad , i=2,\ldots,n-1 \tag{8.4.1}$$

To have a convex interpolant s(x) and to avoid the possibility of s(x) having straight line segments, it is necessary that the d_i parameters should satisfy

$$d_1< \Delta_1< d_2< \ldots < \Delta_{i-1}< d_i<\Delta_i<\ldots<d_n \tag{8.4.2}$$

Now s(x) is convex if and only if

$$s^{(2)}(x)\geqslant 0 \text{ for all } x\in[x_1,x_n] \tag{8.4.3}$$

For $x\in [x_i,x_{i+1}]$, we shall have, after some algebraic manipulation, the result

$$s^{(2)}(x)= \frac{2}{h_i}.\frac{[\alpha_i\theta^3 + \beta_i\theta^2(1-\theta) + \gamma_i\theta(1-\theta)^2 + \delta_i(1-\theta)^3]}{[1 + (r_i-3)\theta(1-\theta)]^3} \tag{8.4.4}$$

where

$$\begin{aligned}
\alpha_i &= r_i(d_{i+1}-\Delta_i) -d_{i+1} +d_i \\
\beta_i &= 3(d_{i+1}-\Delta_i) \\
\gamma_i &= 3(\Delta_i - d_i) \\
\delta_i &= r_i(\Delta_i -d_i)-d_{i+1} + d_i .
\end{aligned} \tag{8.4.5}$$

Hence, from (8.4.4), necessary conditions for convexity are

$$\alpha_i\geqslant 0 \quad \text{and} \quad \delta_i\geqslant 0 . \tag{8.4.6}$$

These conditions, together with inequalities (8.4.2), are also sufficient, since we have $\beta_i > 0$ and $\gamma_i > 0$ in (8.4.4).

Thus, from (8.4.6), we have the condition that the interpolant is convex if and only if

$$r_i \geq \max\left\{\frac{d_{i+1}-d_i}{d_{i+1}-\Delta_i}, \frac{d_{i+1}-d_i}{\Delta_i - d_i}\right\}$$

$$= 1 + M_i/m_i \quad , \qquad (8.4.7)$$

where

$$M_i = \max\{d_{i+1}-\Delta_i \;,\; \Delta_i - d_i\} \;,$$
$$m_i = \min\{d_{i+1}-\Delta_i ,\; \Delta_i - d_i\} \;, \qquad (8.4.8)$$

and the necessary conditions (8.4.2) are assumed.

We have found two choices of r_i which satisfy (8.4.7) and produce pleasing graphical results. They are

$$r_i = 2 + M_i/m_i \;, \qquad (8.4.9)$$
$$r_i = 1 + M_i/m_i + m_i/M_i$$
$$= 1+(d_{i+1}-\Delta_i)/(\Delta_i-d_i) + (\Delta_i-d_i)/(d_{i+1}-\Delta_i), \qquad (8.4.10)$$

the latter being the smaller value.

We can justify the use of either (8.4.9) or (8.4.10) by Theorem 8.2.1 and its corollary. For, suppose $\lambda_i \equiv d_i - f_i^{(1)} = O(h_i^2)$ and $\lambda_{i+1} \equiv d_{i+1} - f_{i+1}^{(1)} = O(h_i^2)$. Then, for (8.4.9), $r_i - 3 = M_i/m_i - 1 = O(h_i)$ and, for (8.4.10), $r_i - 3 = (M_i/m_i - 1)^2/(M_i/m_i) = O(h_i^2)$. Therefore, in practice, the value of r_i in (8.4.10) is to be preferred, since the optimal $O(h^4)$ bound on the interpolation error can be achieved if $O(h^3)$ derivative values are given.

Remark

Strictly convex data are assumed in the above discussion. Otherwise, if $\Delta_{i-1} = \Delta_i$, then we must have $d_{i-1} = d_i = d_{i+1} = \Delta_i$. On the interval $[x_i, x_{i+1}]$, the rational cubic then reduces, as would be expected, to $s(x) = (1-\theta)f_i + \theta f_{i+1}$.

Evidently, there is a similar result on the interval $[x_{i-1}, x_i]$.

8.5 Shape Preserving Interpolation: Convex and Monotonic Data

When the given data are both monotonic increasing and strictly convex, the derivatives d_i must satisfy the inequalities

$$0 \leq d_1 < \Delta_1 < d_2 < \ldots < \Delta_{i-1} < d_i < \Delta_i < \ldots < d_n \qquad (8.5.1)$$

Any convex interpolant will then be monotonic, since we have

$$s^{(1)}(x) = \int_{x_1}^{x} s^{(2)}(x)dx + d_1 \quad ,$$

and also $d_1 \geq 0$, $s^{(2)}(x) \geq 0$.

The convex interpolation method of the previous section is therefore also suitable for the interpolation of convex and monotonic data, a result confirmed by the fact that $1+M_i/m_i \geq (d_{i+1}+d_i)/\Delta_i$ for data satisfying (8.5.1). Thus the convexity condition (8.4.7) suffices to ensure that the monotonicity condition (8.3.7) is satisfied.

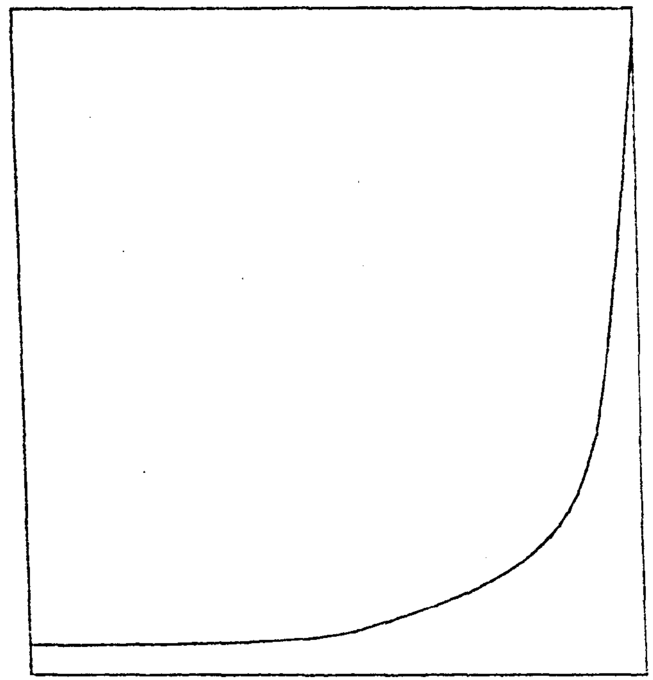
We should note that if the data is convex but not strictly convex, then the interpolant can produce straight line (and hence monotonic) segments, as previously observed.

8.6 Numerical results and discussion

In using the rational cubic scheme, the derivatives d_i usually have to be estimated from the data points, and they must satisfy inequalities (8.4.2). For the convex interpolation problem, $O(h^2)$ arithmetic mean settings for d_i will be suitable ($i=1,\ldots,n$). $O(h^2)$ geometric mean settings are suitable for the interpolation of monotonic data and also of monotonic data which is convex. Application of the rational cubic scheme was made to the three sets of data (MC1), (MC3), (C1).

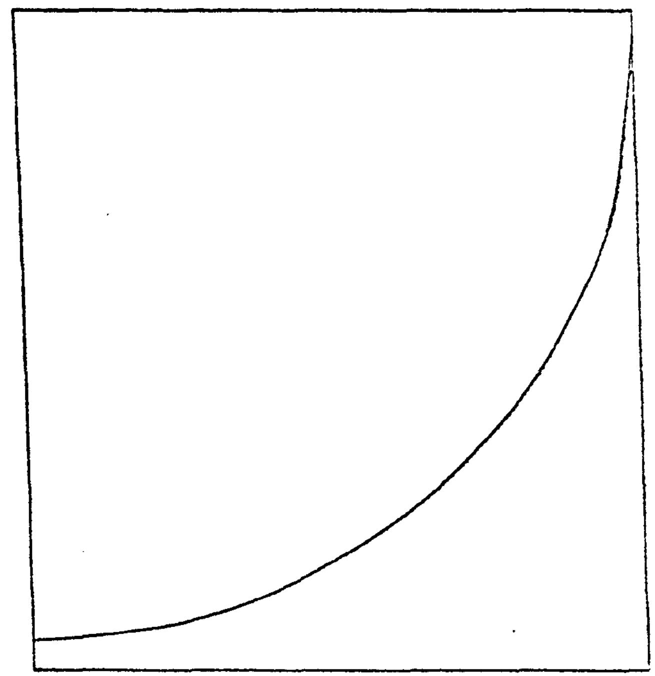
In each case the values of r_i given by the symmetrical forms of equation (8.4.10) have been used, and all the graphs are convex, the possibility of inflexion points appearing anywhere having

been eliminated by choice of the r_i.
On (MC1) and (MC3), positive $O(h^2)$ geometric settings for d_i were used, but on (C1) $O(h^2)$ arithmetic settings were more appropriate since the data here is not monotonic.
There are few data points in (MC1), so we have a fairly severe test of the rational cubic scheme here, particularly since the derivatives are estimated for the data. Even if exact derivatives had been used, the rational interpolants on the subintervals cannot be expected to reproduce the function $1/x^2$, because of the non-linear nature of the interpolation method.
It is of interest to compare the present graphs with those from the rational quadratic C^1 and C^2 schemes already drawn. The convexity constraint on a rational cubic eliminates unnecessary inflexion points on convex data, whereas a rational quadratic only maintains monotonicity and inflexion points may sometimes occur in the graphs.



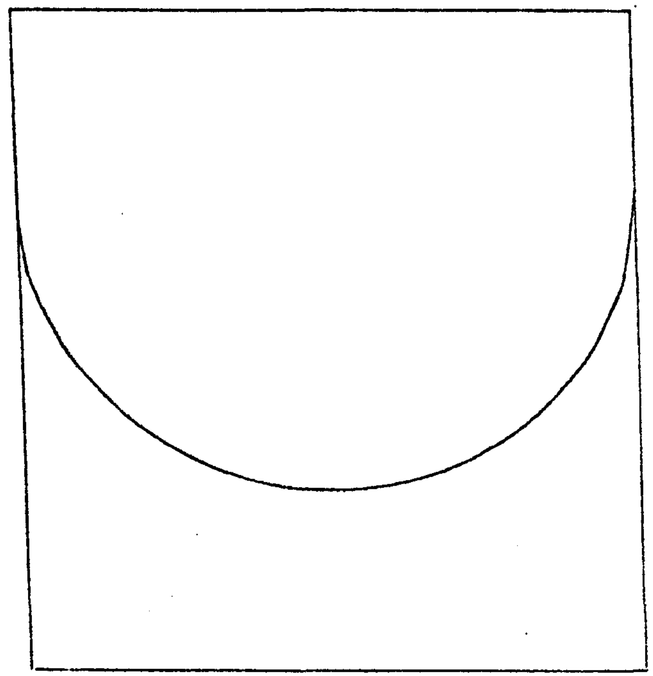
$O(h^2)$ "G" settings

FIG. 8.6.1 C^1 RATIONAL CUBIC (MCI) data



$O(h^2)$ "G" settings

FIG. 8.6.2 C^1 RATIONAL CUBIC (MC3) data



$O(h^2)$ "A" settings

FIG. 8.6.3 C^1 RATIONAL CUBIC

(C1) data

Chapter 9

A C^2 RATIONAL CUBIC CONVEX INTERPOLANT FOR CONVEX DATA

This concluding chapter extends the theory of Chapter 8 by considering the possibility of constructing a C^2 rational cubic spline interpolant for data given to be convex.

9.1 C^2 consistency equations. Choice of parameters r_i.

Our data is such that

$$\Delta_1 < \Delta_2 < \dots < \Delta_{n-1} \tag{9.1.1}$$

We recall equations (8.1.3),(8.3.4),(8.4.4) which give expressions for $s(x)$, $s^{(1)}(x)$, $s^{(2)}(x)$ in a rational cubic interpolant. These expressions contain a parameter r_i for the i th interval $[x_i,x_{i+1}]$, and r_i must be chosen to satisfy inequality (8.4.7) in order to obtain a convex interpolant.

The choice

$$r_i = 1+ (d_{i+1}-\Delta_i)/(\Delta_i-d_i)+(\Delta_i-d_i)/(d_{i+1}-\Delta_i) \tag{9.1.2}$$

was seen to yield pleasing results. However, in Chapter 8, the interpolants constructed were only of class C^1.

For a C^2 interpolant, we require, for i=2,...,n-1,

$$s^{(2)}(x_i+) = s^{(2)}(x_i-) \ .$$

Using (8.4.4) and the corresponding result for the interval $[x_{i-1},x_i]$ this equality translates to

$$\frac{2}{h_i}\zeta_i = \frac{2}{h_{i-1}}\alpha_{i-1} \ ,$$

in the notation of section 8.4.

This gives

$$\frac{1}{h_i}\{ r_i(\Delta_i-d_i)-d_{i+1}+d_i\} = \frac{1}{h_{i-1}}\{ r_{i-1}(d_i-\Delta_{i-1})-d_i+d_{i-1}\} \ ,$$

$$i=2,\dots,n-1. \tag{9.1.3}$$

We choose $d_1 < \Delta_1$ and $d_n > \Delta_{n-1}$. for the end conditions, and try to solve the equations (9.1.3) using the values of r_i given in (9.1.2). When these substitutions are made, the following system of non-linear equations in the derivatives will result:

$$\frac{(\Delta_i - d_i)^2}{h_i(d_{i+1} - \Delta_i)} = \frac{(d_i - \Delta_{i-1})^2}{h_{i-1}(\Delta_{i-1} - d_{i-1})} \quad , \; i=2,\ldots,n-1 \; . \qquad (9.1.4)$$

In the next section we shall show that these equations have a solution satisfying the necessary conditions for convexity, namely,

$$d_1 < \Delta_1 < d_2 < \Delta_2 < \ldots . < d_{n-1} < \Delta_{n-1} < d_n \; . \qquad (9.1.5)$$

Further, we show that such a solution is unique.

9.2 Solution of the consistency equations

The existence of a solution to equations (9.1.4) depends on the result of the following lemma.

Lemma 9.2.1

Let $I = [a_1,b_1] \times \ldots \times [a_n,b_n] \subset R^n$ and $\underline{G} = (G_1,\ldots,G_n)$ be a continuous mapping, $\underline{G}: I \rightarrow I$.

Then there exists $\underline{\xi} = (\xi_1,\ldots,\xi_n) \in I$ such that

$$\underline{\xi} = \underline{G}(\underline{\xi}) \; ,$$

that is, $\xi_i = G_i(\xi_1,\ldots,\xi_n)$, $i=1,\ldots,n$,

or, $\underline{\xi}$ is a fixed point of the mapping $\underline{G}$.

Remark

Lemma 9.2.1 is a particular example of Schauder's Fixed Point Theorem.

Theorem 9.2.1 (Existence)

A solution of the consistency equations (9.1.4) exists satisfying the necessary convexity conditions (9.1.5).

Proof:

From (9.1.4),

$$\frac{\Delta_i - d_i}{(d_i-\Delta_{i-1})} = \left(\frac{h_i}{h_{i-1}}\right)^{\frac{1}{2}} \left(\frac{d_{i+1}-\Delta_i}{\Delta_{i-1}-d_{i-1}}\right)^{\frac{1}{2}} , \quad i=2,\ldots,n-1 ,$$

where positive square roots are taken, consistent with the conditions (9.1.5).

Hence,

$$d_i\{h_{i-1}^{\frac{1}{2}}(\Delta_{i-1}-d_{i-1})^{\frac{1}{2}}+h_i^{\frac{1}{2}}(d_{i+1}-\Delta_i)^{\frac{1}{2}}\}=\Delta_i h_{i-1}^{\frac{1}{2}} (\Delta_{i-1}-d_{i-1})+\Delta_{i-1} h^{\frac{1}{2}}(d_{i+1}-\Delta_i)^{\frac{1}{2}},$$

i.e.,

$$d_i = \frac{\Delta_i \; a_i(d_{i-1}) + \Delta_{i-1} \; b_i(d_{i+1})}{a_i(d_{i-1}) + b_i(d_{i+1})} \; ; \; i=2,\ldots,n-1 \; , \quad (9.2.1)$$

where

$$a_i(d_{i-1}) = h_{i-1}^{\frac{1}{2}}(\Delta_{i-1} - d_{i-1})^{\frac{1}{2}}, \quad b_i(d_{i+1}) = h_i^{\frac{1}{2}}(d_{i+1} - \Delta_i)^{\frac{1}{2}} \qquad (9.2.2)$$

Define

$$\xi_1 = d_1 \text{ (constant)}$$
$$\xi_n = d_n \text{ (constant)} \qquad (9.2.3)$$

and for $\underline{\xi} = (\xi_2,\ldots,\xi_{n-1})$, define $\underline{G} = (G_2,\ldots,G_{n-1})$ by

$$G_i(\underline{\xi}) = \frac{\Delta_i \; a_i(\xi_{i-1}) + \Delta_{i-1} \; b_i(\xi_{i+1})}{a_i(\xi_{i-1}) + b_i(\xi_{i+1})} \; , \; i=2,\ldots,n-1 \quad (9.2.4)$$

where

$$a_i(x) = h_{i-1}^{\frac{1}{2}}(\Delta_{i-1} - x)^{\frac{1}{2}} \quad , \quad b_i(x) = h_i^{\frac{1}{2}}(x - \Delta_i)^{\frac{1}{2}} \; . \qquad (9.2.5)$$

The values d_1, d_n are given end conditions chosen so that $d_1 < \Delta_1$, $d_n > \Delta_{n-1}$.

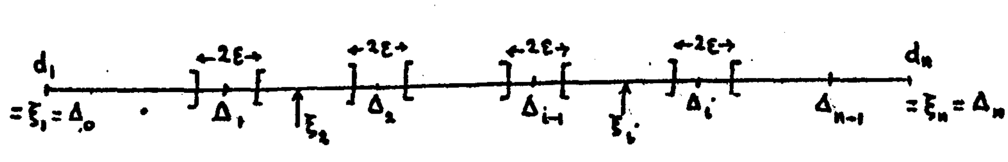
Clearly, if $\Delta_{i-1} < \xi_i < \Delta_i$,$i=2,\ldots,n-1$, then $a_i(\xi_{i-1}) > 0$ and $b_i(\xi_{i+1}) > 0$, and hence, from (9.2.4),

$$\Delta_{i-1} < G_i(\underline{\xi}) < \Delta_i \; .$$

In other words, $\underline{G}$ maps $I = (\Delta_1, \Delta_2) x \ldots x (\Delta_{n-2}, \Delta_{n-1})$ to itself.

We note, here, that I is the cartesian product of <u>open</u> intervals. To make use of the previous lemma, it is necessary for the constituent intervals to be <u>closed</u>. We now show that we can produce a map from $[\Delta_1 + \epsilon, \Delta_2 - \epsilon] x \ldots x [\Delta_{n-2} + \epsilon, \Delta_{n-1} - \epsilon]$ to itself, when $\epsilon > 0$ is chosen to be sufficiently small.

See Figure 9.2.1.



<u>Fig.9.2.1.</u>

For obvious reasons, we begin by restricting ϵ such that

$$0 < \epsilon \leq \epsilon^*$$

where
$$\epsilon^* = \min\{ \min_{2 \leq i \leq n-1} \tfrac{1}{2}(\Delta_i - \Delta_{i-1}), \; \Delta_1 - d_1, d_n - \Delta_{n-1}\} \qquad (9.2.6)$$

Let $\xi_i \in [\Delta_{i-1}+\varepsilon, \Delta_i - \varepsilon]$ for all $i=2,\ldots,n-1$.

For notational simplicity, let

$$\Delta_0 = d_1 \;, \quad \Delta_n = d_n \qquad (9.2.7)$$

Then the following lower and upper bounds for $a_i(\xi_{i-1}), b_i(\xi_{i+1})$ are obtained:

$$\left.\begin{array}{l} h_{i-1}^{\frac{1}{2}}\varepsilon^{\frac{1}{2}} \le a_i(\xi_{i-1}) \le h_{i-1}^{\frac{1}{2}}(\Delta_{i-1}-\Delta_{i-2})^{\frac{1}{2}}, \\ h_i^{\frac{1}{2}}\varepsilon^{\frac{1}{2}} \le b_i(\xi_{i+1}) \le h_i^{\frac{1}{2}}(\Delta_{i+1}-\Delta_i)^{\frac{1}{2}} \end{array}\right\} i=2,\ldots,n-1 \qquad (9.2.8)$$

We shall arrange to have

$$\Delta_{i-1}+\varepsilon \le G_i(\underline{\xi}) \le \Delta_i - \varepsilon \quad \text{for all } i=2,\ldots,n-1.$$

This is equivalent to writing, for $i=2,\ldots,n-1$,

$$\varepsilon \le (\Delta_i-\Delta_{i-1}) \cdot \frac{a_i(\xi_{i-1})}{a_i(\xi_{i-1})+b_i(\xi_{i+1})} \;, \text{ and}$$

$$\varepsilon \le (\Delta_i-\Delta_{i-1}) \cdot \frac{b_i(\xi_{i+1})}{a_i(\xi_{i-1})+b_i(\xi_{i+1})} \qquad (9.2.9)$$

From the bounds in (9.2.8) we obtain

$$\frac{a_i(\xi_{i-1})}{a_i(\xi_{i-1})+b_i(\xi_{i+1})} \ge \frac{h_{i-1}^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}}{h_{i-1}^{\frac{1}{2}}(\Delta_{i-1}-\Delta_{i-2})^{\frac{1}{2}}+h_i^{\frac{1}{2}}(\Delta_{i+1}-\Delta_i)^{\frac{1}{2}}} \;, \text{ and}$$

$$\frac{b_i(\xi_{i+1})}{a_i(\xi_{i-1})+b_i(\xi_{i+1})} \ge \frac{h_i^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}}{h_{i-1}^{\frac{1}{2}}(\Delta_{i-1}-\Delta_{i-2})^{\frac{1}{2}}+h_i^{\frac{1}{2}}(\Delta_{i+1}-\Delta_i)^{\frac{1}{2}}} \,.$$

Hence, from (9.2.9) it is sufficient to make

$$\varepsilon^{\frac{1}{2}} \le \min_{2\le i\le n-1} \left\{ \frac{(\Delta_i-\Delta_{i-1})h_{i-1}^{\frac{1}{2}}}{h_{i-1}^{\frac{1}{2}}(\Delta_{i-1}-\Delta_{i-2})^{\frac{1}{2}}+h_i^{\frac{1}{2}}(\Delta_{i+1}-\Delta_i)^{\frac{1}{2}}} \;,\; \frac{(\Delta_i-\Delta_{i-1})h_i^{\frac{1}{2}}}{h_{i-1}^{\frac{1}{2}}(\Delta_{i-1}-\Delta_{i-2})^{\frac{1}{2}}+h_i^{\frac{1}{2}}(\Delta_{i+1}-\Delta_i)^{\frac{1}{2}}} \right\}$$

where we note $\Delta_0 = d_1$, $\Delta_n = d_n$ (from (9.2.7)), and the initial restriction on ε as indicated in (9.2.6).

Thus $\varepsilon > 0$ exists such that $\underline{G}$ maps $[\Delta_1+\varepsilon, \Delta_2-\varepsilon] \times \ldots \times [\Delta_{n-2}+\varepsilon, \Delta_{n-1}-\varepsilon]$ to itself.

An application of Lemma 9.2.1 now shows that the solution to (9.1.4) exists, which satisfies convexity conditions (9.1.5).

<u>Theorem 9.2.2</u> (<u>Uniqueness</u>)

The solution of the consistency equations (9.1.4) satisfying conditions (9.1.5) is unique.

<u>Proof</u>:

Suppose that there are solutions $d_1, \ldots, d_n$, and $e_1, \ldots, e_n$ to the consistency equations, where

$$d_1 < \Delta_1 < d_2 < \Delta_2 < \ldots < d_{n-1} < \Delta_{n-1} < d_n \quad ,$$

$$e_1 < \Delta_1 < e_2 < \Delta_2 \ \ldots < e_{n-1} < \Delta_{n-1} < e_n \qquad (9.2.10)$$

and $e_1 = d_1$, $e_n = d_n$ are fixed end conditions.

We show that it is impossible to have $e_2 \neq d_2$.

Without loss of generality, assume that

$$d_2 > e_2 \ .$$

We make the following substitutions:

$$\Delta_1 - d_1 = \Delta_1 - e_1 = c \quad ,$$

$$d_n - \Delta_{n-1} = e_n - \Delta_{n-1} = c' \quad ,$$

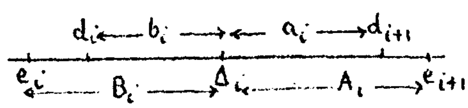
and

$$\Delta_i - d_i = b_i \quad , \quad d_{i+1} - \Delta_i = a_i \quad ,$$

$$\Delta_i - e_i = B_i \quad , \quad e_{i+1} - \Delta_i = A_i \ . \qquad (9.2.11)$$

We note a_i, A_i, b_i, $B_i > 0$ (strictly).

See Figure 9.2.2.



<u>Fig.9.2.2</u>

The two systems of consistency equations are, in this notation,

$$\frac{b_2^2}{h_2 a_2} = \frac{a_1^2}{h_1 c} \ , \ \frac{b_3^2}{h_3 a_3} = \frac{a_2^2}{h_2 b_2} \ , \ \frac{b_4^2}{h_4 a_4} = \frac{a_3^2}{h_3 b_3} \ , \ldots \ , \frac{b_{n-1}^2}{h_{n-1} c'} = \frac{a_{n-2}^2}{h_{n-2} b_{n-2}}$$

$$\frac{B_2^2}{h_2 A_2} = \frac{A_1^2}{h_1 c} \ , \ \frac{B_3^2}{h_3 A_3} = \frac{A_2^2}{h_2 B_2} \ , \ \frac{B_4^2}{h_4 A_4} = \frac{A_3^2}{h_3 B_3} \ , \ldots \ , \frac{B_{n-1}^2}{h_{n-1} c'} = \frac{A_{n-2}^2}{h_{n-2} B_{n-2}} \qquad (9.2.12)$$

The assumption $d_2 > e_2$ is equivalent to the pair of inequalities $a_1 > A_1$, $B_2 > b_2$. We will show this will imply

$$d_3 < e_3,\ d_4 > e_4,\ d_5 < e_5\ ,\ldots$$

or, equivalently,

$$(A_2 > a_2 \text{ and } b_3 > B_3), (a_3 > A_3 \text{ and } B_4 > b_4),\ (A_4 > a_4 \text{ and } b_5 > B_5), \ldots \quad (9.2.13)$$

Hence the equality $d_n = e_n$ is eventually contradicted.

To this end, let P(m) denote, for $m \in N^+$, the statement:

$$\left(\frac{A_{2m}}{a_{2m}}\right)^{2^{2m-2}} = \left(\frac{B_{2m}}{b_{2m}}\right)^{2^{2m-1}} \cdot \left(\frac{a_{2m-1}}{A_{2m-1}}\right)^{3.2^{2m-3}} \cdot \left(\frac{A_{2m-2}}{a_{2m-2}}\right)^{3.2^{2m-4}} \cdots \left(\frac{A_2}{a_2}\right)^3 \left(\frac{a_1}{A_1}\right)^2 ,$$

$$\left(\frac{a_{2m+1}}{A_{2m+1}}\right)^{2^{2m-1}} = \left(\frac{b_{2m+1}}{B_{2m+1}}\right)^{2^{2m}} \cdot \left(\frac{A_{2m}}{a_{2m}}\right)^{3.2^{2m-2}} \cdot \left(\frac{a_{2m-1}}{A_{2m-1}}\right)^{3.2^{2m-3}} \cdots \left(\frac{A_2}{a_2}\right)^3 \left(\frac{a_1}{A_1}\right)^2 . \quad (9.2.14)$$

We use induction to show P(m) is true.

From (9.2.12), by division, $\dfrac{A_2}{a_2} = \left(\dfrac{B_2}{b_2}\right)^2 \left(\dfrac{a_1}{A_1}\right)^2$, and

$$\left(\frac{a_3}{A_3}\right)^2 = \left(\frac{b_2}{B_2}\right)^2 \left(\frac{b_3}{B_3}\right)^4 \left(\frac{A_2}{a_2}\right)^4 ,\ = \left(\frac{b_3}{B_3}\right)^4 \left(\frac{A_2}{a_2}\right)^3 \left(\frac{a_1}{A_1}\right)^2 ,$$

on substituting for B_2/b_2 from the previous step. Hence P(1) holds.

We prove $P(m) \Rightarrow P(m+1)$. From (9.2.12) by division

$$\left(\frac{A_{2m+2}}{a_{2m+2}}\right)^{2^{2m}} = \left(\frac{B_{2m+1}}{b_{2m+1}}\right)^{2^{2m}} \cdot \left(\frac{B_{2m+2}}{b_{2m+2}}\right)^{2^{2m+1}} \cdot \left(\frac{a_{2m+1}}{A_{2m+1}}\right)^{2^{2m+1}} ,$$

$$= \left(\frac{B_{2m+2}}{b_{2m+2}}\right)^{2^{2m+1}} \cdot \left(\frac{a_{2m+1}}{A_{2m+1}}\right)^{3.2^{2m-1}} \cdot \left(\frac{A_{2m}}{a_{2m}}\right)^{3.2^{2m-2}} \cdots \left(\frac{A_2}{a_2}\right)^3 \left(\frac{a_1}{A_1}\right)^2 ,$$

on substituting for B_{2m+1}/b_{2m+1} using the inductive hypothesis.

Also, from (9.2.12) by division,

$$\left(\frac{a_{2m+3}}{A_{2m+3}}\right)^{2^{2m+1}} = \left(\frac{b_{2m+2}}{B_{2m+2}}\right)^{2^{2m+1}} \cdot \left(\frac{b_{2m+3}}{B_{2m+3}}\right)^{2^{2m+2}} \cdot \left(\frac{A_{2m+2}}{a_{2m+2}}\right)^{2^{2m+2}} ,$$

$$= \left(\frac{b_{2m+3}}{B_{2m+3}}\right)^{2^{2m+2}} \cdot \left(\frac{A_{2m+2}}{a_{2m+2}}\right)^{3.2^{2m}} \cdot \left(\frac{a_{2m+1}}{A_{2m+1}}\right)^{3.2^{2m-1}} \cdots \left(\frac{A_2}{a_2}\right)^3 \left(\frac{a_1}{A_1}\right)^2 ,$$

on substituting for b_{2m+2}/B_{2m+2} from the previous step of this calculation. Thus P(m) implies P(m+1) and hence by induction equations (9.2.14) hold. The assumption $a_1 > A_1$, $B_2 > b_2$ (equivalent to $d_2 > e_2$) now allows us to conclude that (9.2.13) is true, leading to the contradiction $d_n \neq e_n$. Thus $e_2=d_2$ and by recursion $e_i=d_i$,$i=3,\dots,n$. Hence there is at most one solution $d_2,\dots d_{n-1}$ satisfying the convexity conditions (9.1.5).

9.3 Numerical solution. The theoretical and practical difficulties

In the previous sections of this chapter we formulated a set of consistency equations for a rational cubic convex spline, having a unique solution for the derivatives d_i in the allowed intervals: $\Delta_{i-1} < d_i < \Delta_i$, $i=2,\dots,n-1$. We assume $d_1 < \Delta_1$ and $d_n > \Delta_{n-1}$ are given as end conditions.

The consistency equations, (9.1.4),, arose through the use of the parameters $r_i = 1+(d_{i+1}-\Delta_i)/(\Delta_i-d_i)+(\Delta_i-d_i)/(d_{i+1}-\Delta_i)$ in the rational cubic. However, these equations present us with a number of difficulties, both theoretical and practical.

Theoretical problems

A satisfactory error bound analysis has not, so far, been obtained. The analysis is more complicated than that of the rational quadratic and an upper bound for $\max_{2\leq i\leq n-1} |d_i-f_i^{(1)}|$ cannot, at present, be given. Also, no contraction map could be discovered which provides a basis for an iteration method for solving the consistency equations. Thus, it has not been found possible to prove, for example, that either a Jacobi or Gauss-Seidel type of iteration, based on (9.2.1), actually converges (though in practice they often do).

Practical problems

From the existence and uniqueness proof it is almost evident that there will be problems of finding d_i numerically when the derivatives have values close to the ends of their allowed intervals: $\Delta_{i-1} < d_i < \Delta_i$. Also, if used as a recursive system, the equations (9.1.4) show a marked instability when n is large. This is evident from the induction argument given above. Further, numerical experimentation shows that there exist many solutions of the consistency equations which give rise to non-convex curves, even though there is only one unique solution giving a convex curve.

<u>A method of solution (n not large)</u>

The details given in the uniqueness proof suggest a possible procedure for attempting to obtain the required solution and thus the unique convex interpolant. This procedure relies on the relationship discovered between the set d_i and the associated set e_i of the proof of Theorem 9.2.2.; see Figure 9.3.1.

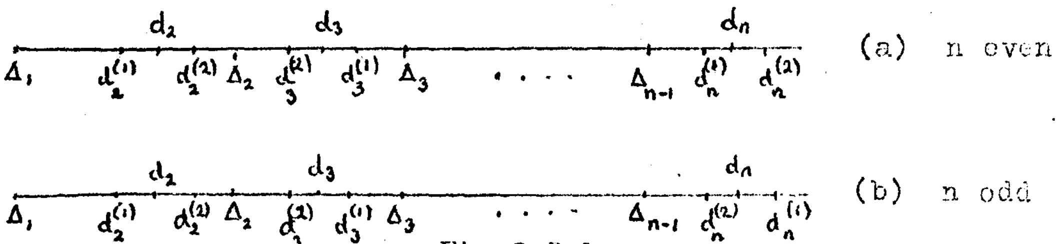


Fig.9.3.1.

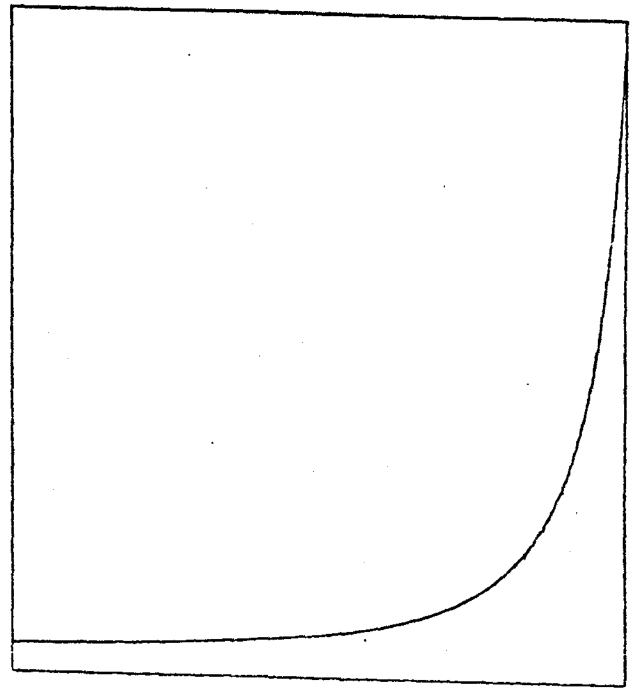
We use the consistency equations as a recursive system, and consider d_2 as a variable which requires determination. We seek two values of d_2 , namely $d_2^{(1)}$, $d_2^{(2)}$, such that $\Delta_1 < d_2^{(1)} < d_2 < d_2^{(2)} < \Delta_2$ and which lead, via the system, to values $d_n^{(1)}$, $d_n^{(2)}$ on either side of the true value d_n. If n is even, we require $\Delta_{n-1} < d_n^{(1)} < d_n < d_n^{(2)}$; but if n is odd, we require $\Delta_{n-1} < d_n^{(2)} < d_n < d_n^{(1)}$. Further, we must ensure that the values $d_i^{(1)}$, $d_i^{(2)}$ for each of $i=3,\dots,n-1$ remain in the allowed intervals: $\Delta_{i-1} < d_i^{(1)}, d_i^{(2)} < \Delta_i$.

A search for suitable values $d_2^{(1)}$, $d_2^{(2)}$ is made. Unfortunately, due to the presence of unwanted solutions, the values of $d_2^{(1)}$, $d_2^{(2)}$ are likely to be very close. A binary search method in the interval $[\Delta_1, \Delta_2]$ can be used to locate a value $d_2^{(1)}$ in a small interval with right hand point d_2. Following that, a value $d_2^{(2)}$ can be produced, by converging from the right. Using the values $d_2^{(1)}$, $d_2^{(2)}$ and the corresponding values $d_n^{(1)}$, $d_n^{(2)}$, a secant approach may be used, which after a few iterations gives d_2 corresponding to the correct value d_n, to the demanded accuracy.

On the data sets of our experiments, the method works fairly well. We have taken data sets (MC1), (MC3), and have found that the binary search is not slow, the number of iterations is small (9,12 respectively) and the graphs are quite acceptable. The number of data points in each case is, of course, relatively small.

An experiment is made on the set (MC2) to ascertain the order of convergence (as done in the other chapters). With a uniform h=0.2, we have found $\max_{2\le i\le n-1} |f_i^{(1)}-d_i| = 0.281355\times10^{-5}$, and with h=0.1, we have found $\max_{2\le i\le n-1} |f_i^{(1)}-d_i| = 0.16583\times10^{-6}$. The ratio of these errors gives 16.97 indicating $O(h^4)$ convergence. However, with h=0.05 (when the number of points is 21) the proposed method has failed in a search for $d_2^{(1)}$, $d_2^{(2)}$ due to instability in the method and no results are available.

These difficulties indicate that further study of the consistency equations is needed, with the object of finding a stable iterative method for which convergence is guaranteed.



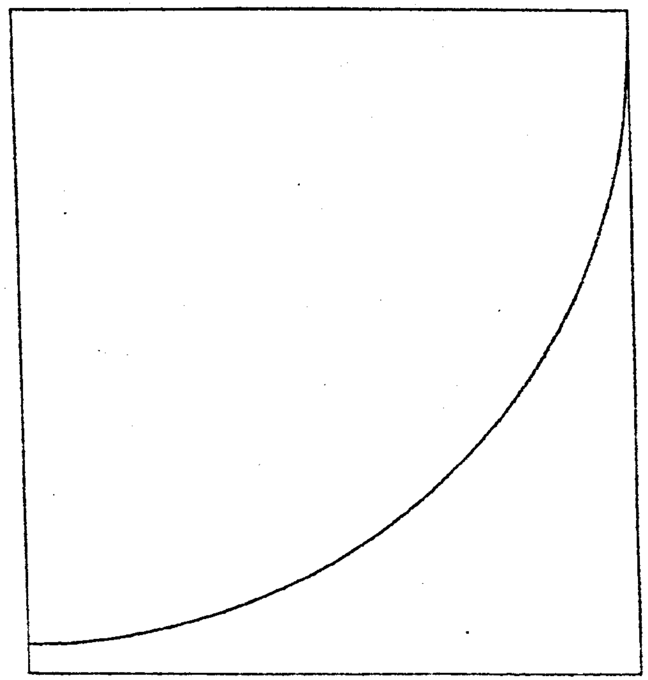
exact end derivatives ($d_1 = 0{\cdot}25$, $d_n = 250$)
[9 iterations]

FIG. 9.3.1

RATIONAL CUBIC

CONVEX SPLINE

(MC1) data



end derivatives $d_1 = 0$, $d_n = 25$
[12 iterations]

FIG. 9.3.2

RATIONAL CUBIC

CONVEX SPLINE

(MC3) data

CONCLUSIONS

We have been concerned in this thesis with a variety of methods for solving to a large extent the problem of shape-preserving interpolation. We have been led to consider the special cases of monotonic and convex data, and the possibility of effecting the interpolation by means of piecewise defined rational quadratic and cubic functions. An important factor influencing the acceptance or otherwise of a particular rational scheme is the order of accuracy of the resulting interpolant, which is determined in part by the accuracy of the derivatives d_i at the knots. Their determination has been achieved by both explicit and implicit means. There exists a choice between rational quadratic and rational cubic schemes, and one between providing C^1 or C^2 interpolants. In the following remarks we attempt to guide the reader into making a specific decision regarding this choice.

Thus, we suggest that rational quadratic schemes should be used whenever the data set is simply monotonic. For a C^1 interpolant, any of three $O(h^2)$ explicitly evaluated derivative settings for d_i may be used to provide reasonably accurate results quickly. In particular, experiments indicate that geometric or harmonic mean approximations are to be preferred. If greater accuracy is required, we suggest the two-term recurrence relations between derivatives and the averaging procedure, using $O(h^3)$ end conditions. These are easy to apply, do not need iteration and give results which are probably marginally better than those where use is made of explicit $O(h^3)$ derivatives. When strictly monotonic data is given, we would propose a C^2 rational quadratic spline interpolant in most examples, and, when end conditions have to be estimated from the data, the geometric or harmonic $O(h^2)$ end conditions may prove sufficient.

By contrast, the rational cubic interpolation scheme is more appropriate for use on convex data. It is dependent on the choice of a parameter set r_i. When the data is not also monotonic, the $O(h^2)$ arithmetic explicit settings d_i together with suitably calculated values of r_i determined from them should give good C^1 interpolants. To give satisfactory C^2 spline interpolants we have to seek a solution to a set of non-linear equations in the derivatives which is not easy to solve numerically. A numerically efficient procedure for solving them has yet to be found, using an iteration which can be proved theoretically and practically to be convergent. Further work will show if this is achievable, and it might then be possible to investigate the problem of providing a spline solution when the data is composed of a union of convex and concave pieces. This remains an ultimate objective.

APPENDIX A.1

The BERNSTEIN polynomials and WEIERSTRASS's Approximation Theorem

Let $f:[0,1]\to R$ be continuous on $[0,1]$.

Define the sequence of BERNSTEIN polynomials $(B_n(f;\theta))$, depending on f, by

$$B_n(f;\theta)=\sum_{k=0}^{n}\binom{n}{k}f(\tfrac{k}{n})\ \theta^k(1-\theta)^{n-k} \qquad (n=1,2,\ldots)$$

The WEIERSTRASS Approximation Theorem states that for all $\varepsilon>0$, there exists $n_0(\varepsilon)$ such that

$$|f(\theta)-B_n(f;\theta)|<\varepsilon \ , \text{ for all } n\geq n_0 \ , \text{ all } 0\leq\theta\leq 1.$$

For a proof of this theorem and properties of the polynomials B_n in the approximation of functions, reference may be made to Chapter VI of [6] .

Example: $f:[0,1]\to R$, $f(\theta)=|\theta-\frac{1}{2}|$. We compute $B_1,\ldots,B_5$, associated with this function.

$$B_1(f;\theta) = f(0)(1-\theta)+f(1)\theta=\tfrac{1}{2} ,$$

$$B_2(f;\theta) = f(0)(1-\theta)^2+2f(\tfrac{1}{2})\theta(1-\theta)+f(1)\theta^2=\tfrac{1}{2}\{(1-\theta)^2+\theta^2\}=\tfrac{1}{2}-\theta(1-\theta) ,$$

$B_3(f;\theta) = B_2(f;\theta)$ may be checked

$$B_4(f;\theta) = f(0)(1-\theta)^4+4f(\tfrac{1}{4})\theta(1-\theta)^3+6f(\tfrac{1}{2})\theta^2(1-\theta)^2+4f(\tfrac{3}{4})\theta^3(1-\theta)+f(1)\theta^4,$$

$$= \tfrac{1}{2}\{(1-\theta)^4+\theta^4\}+\theta(1-\theta)\{(1-\theta)^2+\theta^2\}$$

$$= \tfrac{1}{2}-\theta(1-\theta)\{1+\theta(1-\theta)\}$$

and $B_5(f;\theta) = B_4(f;\theta)$ may be checked.

Figure A.1.1 shows the manner in which $B_1,\ldots,B_5$ progressively approximate f.

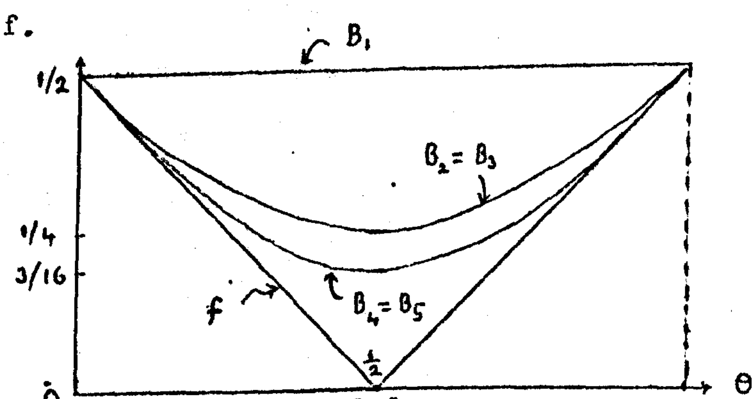


Fig.A.1.1

APPENDIX A.2

PEANO's Kernel Theorem

Let L: $C^{n+1}[a,b] \rightarrow R$ be defined as

$$L(f) = \int_a^b \{a_o(x)f(x)+a_1(x)f^{(1)}(x)+\ldots+a_n(x)f^{(n)}(x)\}\,dx$$

$$+\sum_{k=o}^{m_o} a_{ko}f(x_{ko})+\sum_{k=o}^{m_1} a_{k1}f^{(1)}(x_{k1})+\ldots+\sum_{k=o}^{m_n} a_{kn}f^{(n)}(x_{kn}).$$

where $a_i(x)$ are functions assumed piecewise continuous on $[a,b]$, a_{ij} are constants, and the values x_{ij} belong to $[a,b]$. Then L is a linear functional (since $L(\alpha f+\beta g)=\alpha L(f)+\beta L(g)$, for all $\alpha,\beta \in R$ and $f,\ g \in C^{n+1}[a,b]$).

PEANO's Kernel Theorem states that:

If, for all polynomials p of degree at most n, $L(p)=0$, then for any $f \in C^{n+1}[a,b]$,

$$L(f)=\int_a^b f^{n+1}(t)K(t)\,dt \ , \text{ where } K(t) = \frac{1}{n!} L_x[(x-t)_+^n] \ ,$$

in which

$$(x-t)_+^n = \begin{cases} (x-t)^n & \text{if } x \geq t \\ 0 & \text{if } x < t \end{cases}$$

and L_x means L is applied to its argument considered as a function of x.

Corollary:

If the Kernel, K(t), maintains a constant sign in $a \leq t \leq b$, then

$$L(f) = \frac{f^{n+1}(\xi)}{(n+1)!}L(x^{n+1}) \ , \quad \text{for some } \xi \in (a,b).$$

For a proof of the theorem and its consequences, consult, for instance, Chapter III in [6] .

Example 1

If $f \in C^3[0,1]$, then $L(f) \equiv f(1)-f(0)- f^{(1)}(\frac{1}{2}) = f^{(3)}(\xi)/24$, where $\xi \in (0,1)$.

Proof:

If $f= p_2$, an arbitrary quadratic, $L(p_2)= 0$, but if $f=p_3$, an arbitrary cubic, $L(p_3) \neq 0$.

By the theorem, $K(t) = \frac{1}{2!} L_x[(x-t)_+^2] = \frac{1}{2}\{(1-t)_+^2-(-t)_+^2 -2(\frac{1}{2}-t)_+\}$

$$= \tfrac{1}{2}\{ (1-t)^2-(1-2t)_+\}$$

$$= \begin{cases} \frac{1}{2}t^2 & \text{if } t \leq \frac{1}{2} \\ \frac{1}{2}(1-t)^2 & \text{if } t > \frac{1}{2} . \end{cases}$$

Now $K(t) \geq 0$ for all $0 \leq t \leq 1$. Hence by the corollary, for some $\xi \in (0,1)$,

$$L(f) = \frac{f^{(3)}(\xi)}{3!}L(x^3) = \frac{f^{(3)}(\xi)}{6}\{1-3(\tfrac{1}{2})^2\} = f^{(3)}(\xi)/24 .$$

Example 2

If $f \in C^4[0,1]$, then $L(f) \equiv \int_0^1 f(x)dx - \frac{1}{2}\{f(0)+f(1)\}- \frac{1}{12}\{f^{(1)}(0)- f^{(1)}(1$

$$= f^{(4)}(\xi)/720, \text{ for some } \xi \in (0,1) .$$

Proof:

If $f= p_3$, an arbitrary cubic, $L(p_3)= 0$, but if $f=p_4$, an arbitrary quartic, $L(p_4)\neq 0$.

The theorem gives

$$K(t) = \frac{1}{3!}L_x[(x-t)_+^3] = \frac{1}{6}\{\int_0^1(x-t)_+^3dx -\tfrac{1}{2}((-t)_+^3+(1-t)_+^3)-\tfrac{3}{12}((-t)_+^2-(1-t)_+^2)\}$$

$$= \frac{1}{6}\{ \frac{1}{4}(1-t)^4- \tfrac{1}{2}(1-t)^3+ \tfrac{1}{4}(1-t)^2\}$$

$$= \frac{1}{24}(1-t)^2t^2 .$$

Thus $K(t) \geq 0$ throughout $0 \leq t \leq 1$, so by the corollary,

$$L(f)= \frac{f^{(4)}(\xi)}{4!}L(x^4) = \frac{f^{(4)}(\xi)}{24}\{ \frac{1}{5} - \tfrac{1}{2}.1- \frac{1}{12}.(-4)\} = \frac{f^{(4)}(\xi)}{720},\ \xi \in (0,1).$$

APPENDIX A.3

Useful Expansions

Let $f \in C^{\infty}[x_1, x_n]$ interpolate the data $(x_i, f_i), i=1,2,\ldots,n$ and let Δ_i, $\Delta_{i,j}$ be as defined under the section headed Notation.

The following Taylor expansions may be written down:

For $i=1,2,\ldots,n-1$,

$$\Delta_i = \Delta_{i,i+1} = f_i^{(1)} + h_i.(f_i^{(2)}/2) + h_i^2.(f_i^{(3)}/6) + h_i^3.(f_i^{(4)}/24) + \ldots \quad (A.3.1)$$

For $i=2,\ldots,n$,

$$\Delta_{i-1} = \Delta_{i-1,i} = f_i^{(1)} - h_{i-1}(f_i^{(2)}/2) + h_{i-1}^2(f_i^{(3)}/6) - h_{i-1}^3(f_i^{(4)}/24) + \ldots \quad (A.3.2)$$

For $i=1,\ldots,n-2$,

$$\Delta_{i,i+2} = f_i^{(1)} + (h_i + h_{i-1})(f_i^{(2)}/2) + (h_i + h_{i+1})^2(f_i^{(3)}/6) + \ldots \quad (A.3.3)$$

For $i=3,\ldots,n$

$$\Delta_{i-2,i} = f_i^{(1)} - (h_{i-2} + h_{i-1})(f_i^{(2)}/2) + (h_{i-2} + h_{i-1})^2(f_i^{(3)}/6) + \ldots \quad (A.3.4)$$

When $f \in C^r[x_1, x_n]$, finite expansions are taken. If $r=4$, for example, equation (A.3.1) would be used as

$$\Delta_i = f_i^{(1)} + h_i.(f_i^{(2)}/2) + h_i^2.(f_i^{(3)}/6) + h_i^3.(f_{i+\alpha}^{(4)}/24) \,,$$

where $f_{i+\alpha}^{(4)}$ is an abbreviation for $f^{(4)}(x_i + \alpha h_i)$, and $0 < \alpha < 1$.

A proof of (A.3.2) would proceed thus:

$$\Delta_{i-1} = (f_i - f_{i-1})/h_{i-1} = \{f_i - (f_i - h_{i-1}.f_i^{(1)} + h_{i-1}^2.\frac{f_i^{(2)}}{2} - \ldots)\}/h_{i-1}$$

$$= f_i^{(1)} - h_{i-1}.(f_i^{(2)}/2) + \ldots$$

An example

If $f_1^{(1)} > 0$, $\Delta_1 > 0$, $\Delta_{1,3} > 0$, then $d_1 = \Delta_1^{(1+h_1/h_2)}.\Delta_{1,3}^{(-h_1/h_2)}$ approximates $f_1^{(1)}$ to $O(h^2)$ accuracy.

Proof:

$$d_1 = \{f_1^{(1)} + h_1\frac{f_1^{(2)}}{2} + O(h_1^2)\}^{(1+h_1/h_2)}.\{f_1^{(1)} + (h_1+h_2)\frac{f_1^{(2)}}{2} + O((h_1+h_2)^2)\}^{(-h_1/h_2)}$$

therefore

$$d_1/f_1^{(1)} = \{1 + h_1\frac{f_1^{(2)}}{2f_1^{(1)}} + O(h_1^2)\}^{(1+h_1/h_2)}.\{1 + (h_1+h_2)\frac{f_1^{(2)}}{2f_1^{(1)}} + O((h_1+h_2)^2)\}^{(-h_1/h_2)}$$

$$= \{1 + h_1(1+h_1/h_2).\tfrac{1}{2}f_1^{(2)}/f_1^{(1)} + O(h^2)\}\{1 - (h_1/h_2)(h_1+h_2).\tfrac{1}{2}f_1^{(2)}/f_1^{(1)} + O(h^2)\}$$

$$= 1 + O(h^2).$$

REFERENCES

[1] H.AKIMA, A new method of interpolation and smooth curve fitting based on local procedures, J.Assoc.Comput.Mach., 17(1970),pp.589-602.

[2] G.H.BEHFOROOZ and N.PAPAMICHAEL , End conditions for cubic spline interpolation,J.Inst.Maths.Applics.23(1979) pp.355-366.

[3] G.BIRKHOFF and A.PRIVER, Hermite interpolation errors for derivatives, J.Math.Phys.46 (1967),pp.440-447.

[4] R.L.BURDEN, J.D.FAIRES and A.C.REYNOLDS. Numerical Analysis, Chapter 3,pp.73-123. PWS Publishers (1981).

[5] J.BUTLAND, A method of interpolating reasonably-shaped curves through any data. Proceedings of Computer Graphics (1980), On line Publications Ltd.,pp.409-421.

[6] P.J.DAVIS, Interpolation and Approximation. Chapters 2, 3,6 (1975). Dover Publications, Inc.

[7] R.DELBOURGO and J.A.GREGORY, C^2 Rational Quadratic Spline Interpolation to Monotonic Data, IMA J.Num.Analysis, (1983) 3,pp.141-152.

[8] R.DELBOURGO and J.A.GREGORY, Shape-Preserving Piecewise Rational Interpolation, Brunel University TR/10/83. (To appear in SIAM Journal of Scientific and Statistical Computing).

[9] R.DELBOURGO and J.A.GREGORY, The determination of derivative parameters for a monotonic rational quadratic interpolant, Brunel University TR/07/84.

[10] R.DELBOURGO and J.A.GREGORY, An implicit method for determining the derivative parameters for a monotonic rational quadratic interpolant. In preparation.

[11] T.M.R.ELLIS and D.H.McLAIN, Algorithm 514. A new method of cubic curve fitting using local data. ACM Trans. Math.Software 3(1977),pp.175-178.

[12] F.N.FRITSCH and J.BUTLAND, A method for constructing local monotone piecewise cubic interpolants, SIAM Journal of Scientific and Statistical Computing (to appear June 1984).

[13] F.N.FRITSCH and R.E.CARLSON, Monotone piecewise cubic interpolation,SIAM J.Num.Analysis 17(1980),pp.238-246.

[14] J.A.GREGORY and R.DELBOURGO, Piecewise Rational Quadratic Interpolation to Monotonic Data, IMA J.Num.Analysis (1982)2,pp.123-130.

[15] D.F.McALLISTER,E.PASSOW and J.A.ROULIER, Algorithms for computing shape preserving spline interpolants to data. Math.Comp.31 (1977),pp.717-725.

[16] D.F.McALLISTER and J.A.ROULIER, An algorithm for computing a shape preserving osculatory quadratic spline. ACM Trans.Math.Software 7(1981),pp.331-347.

[17] H.METTKE, Convex cubic Hermite-spline interpolation, Journal of Comp.& Appl.Maths.9(1983),pp.205-211.

[18] E.PASSOW, Piecewise monotone spline interpolation. J.Approximation Theory 12(1974),pp.240-241.

[19] E.PASSOW and J.A.ROULIER, Monotone and convex spline interpolation,SIAM J. Num.Anal.14(1977),pp.904-909.

[20] S.PRUESS, Alternatives to the exponential spline in tension, Maths.Comp.33(1979)pp.1273-1281.

[21] L.L.SCHUMAKER, On shape preserving Quadratic Spline Interpolation. SIAM J. Numer.Anal.20(1983),pp.854-864.