M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique

R. S. Anderssen
F. R. de Hoog
L. B. Wahlbin

On pointwise stability of cubic smoothing splines with nonuniform sampling points

M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique, tome 25, nº 6 (1991), p. 671-692
<http://www.numdam.org/item?id=M2AN_1991__25_6_671_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(Vol. 25, n° 6, 1991, p. 671 à 692)

ON POINTWISE STABILITY OF CUBIC SMOOTHING SPLINES WITH NONUNIFORM SAMPLING POINTS (*)

R. S. ANDERSSEN (1), F. R. DE HOOG (1), and L. B. WAHLBIN (2)

Communicated by J. H. BRAMBLE

Abstract. — *A general technique is developed for analyzing pointwise stability of smoothing splines when the design points are unevenly spaced. In particular, it is not assumed that the distributions of design points approach a limit. The technique can also be applied to the analysis of pointwise convergence.*

Résumé. — *Nous développons une technique générale pour l'analyse de la stabilité ponctuelle dans le lissage des fonctions splines dans le cas où les points sont inégalement distribués. En particulier, nous ne supposons pas que les distributions de points approchent une limite. La technique s'applique aussi à l'analyse de la convergence ponctuelle.*

1. INTRODUCTION

Given a grid of n design, or sampling, points $0 = x_0 < x_1 < \cdots < x_{n-1} = 1$ on the normalized interval [0, 1], data d_i, $i = 0, 1, \ldots, n-1$, at these design points, and a penalty, or smoothing, parameter $\lambda_n > 0$, the piecewise cubic twice continuously differentiable smoothing spline $s_n(x)$ is the solution of the following minimization problem,

$$\min_{s_n \in H^2} \frac{1}{n} \sum_{i=0}^{n-1} (s_n(x_i) - d_i)^2 + \lambda_n \int_0^1 (s_n''(x))^2 \, dx \,, \tag{1.1}$$

cf. Holladay [2], Reinsch [4] and Schoenberg [5]. Here $H^2 = H^2[0, 1]$ is the Sobolev space of functions with square integrable second derivatives in the distribution sense.

For the underlying idea of striking a happy medium between approximat-

(*) Received May 1990.
(1) CSIRO Division of Mathematics and Statistics, GPO Box 1965, Canberra, ACT 2601, Australia.
(2) Department of Mathematics, Cornell University, Ithaca, NY 14853, USA.

ing data, typically contaminated by errors, and pleasing the eye, see Whittaker [7]. For a critique of smoothing splines, see de Boor [1, p. 248-9].

The major question addressed in this paper is that of pointwise stability of the smoothing spline. To phrase this question precisely we adopt an asymptotic point of view. Let

$$G_n = \{x_0, x_1, \dots, x_{n-1}\}\ , \quad x_i = x_i(n)\ ,$$

be a sequence of grids of design points, parametrized by $n = 2, 3, \dots,$ and let these grids together with the smoothing parameters λ_n form a design family

$$\mathscr{D} = \{\mathscr{D}_n = \{G_n, \lambda_n\}\ ; n = 1, 2, \dots\ \}\ . \tag{1.2}$$

Further, for a continuous function $v(x)$ on [0, 1], let

$$\|v\|_\infty = \max_{0 \leq x \leq 1} |v(x)|$$

and, for an n-vector $v = \{v_i\ ; i = 0, \dots, n-1\}$,

$$|v|_\infty = |v|_{\infty, n} = \max_{0 \leq i \leq n-1} |v_i|\ . \tag{1.3}$$

Then, what conditions can be placed on the design family $\mathscr{D}$ so that there exists a constant $C = C(\mathscr{D})$, independent of n, such that

$$\|s_n\|_\infty \leq C\, |d|_\infty \tag{1.4}$$

for all data d? (For the reader who likes to think in terms of a fixed set of design points, and a fixed penalty parameter, the equivalent question is to ask exactly what features of this set influences C in (1.4).)

Our investigation gives a technique for treating the stability question (1.4). We are aiming at generality, and it is, then, perhaps not surprising that our technique is somewhat implicit in nature. We will give examples of design families in which it becomes explicit.

Most published research into smoothing splines involves assumptions that the grid sequence G_n finally settles down to a distinguishable pattern, in the sense that the function

$$F_n(y) = (\text{number of design points} \leq y\,)/n$$

approaches a limit $F(y)$. Our technique does not assume that such a limit exists. (In two of our applications, Sections 4.2 and 4.3, the limiting distributions do exist. In contrast to other investigations, we can allow F' to become infinite or to vanish, corresponding to systematic meshrefinement around a point, or meshthinning, respectively, or even to be a Dirac delta function.)

The motivation for this paper is the earlier work of Silverman [6] and de Hoog and Anderssen [3]. In these papers the properties of the Green's (kernel) functions $g_n(x\,;y)$, which recovers the unique minimizer $s_n(y)$ of (1.1) as

$$s_n(y) = \frac{1}{n}\sum_{i=1}^{n-1} g_n(x_i\,;y)\, d_i\,,$$

are the key to any analysis of stability and convergence. Following this pattern, the analysis in the present paper focuses on first obtaining suitable estimates for the Green's functions; thereafter the desired results follow using standard techniques. The present analysis will pinpoint the local and global influences of data and the design family on the behaviour of the smoothing spline.

We mention two generalizations of (1.1) which, in the interest of brevity, we do not pursue in this paper. The first is to replace the penalty on second derivatives by a penalty on some other derivative, thus generating other than cubic smoothing splines. The second is to replace the sum occurring by a weighted sum,

$$\sum_{i=0}^{n-1} W_{i,n}(s_n(x_i) - d_i)^2\,,$$

where $W_{i,n}$ (and λ_n) could be chosen according to statistical rules. (The statistical aspects of choosing λ_n in (1.1) are not considered in this investigation.)

We now collect together conventions and notation which will be used throughout the paper. In the interest of clean notation, dependence on n will usually be suppressed. Thus $s(x) = s_n(x)$, $\lambda = \lambda_n$, *cf.* also (1.3). For an interval $I \subseteq [0, 1]$,

$$\text{(1.5)}\qquad \begin{aligned} \langle v, w\rangle_I &= \langle v, w\rangle_{I,n} = \frac{1}{n}\sum_{x_i\in I} v(x_i)\, w(x_i)\,,\\ |v|_I &= |v|_{I,n} = \langle v, v\rangle_{I,n}^{1/2} \end{aligned}$$

and

$$\text{(1.6)}\qquad (v, w)_I = \int_I v(x)\, w(x)\, dx\,, \quad \|v\|_I = (v, v)_I^{1/2}\,.$$

When $I = [0, 1]$ we drop the indicator I so that $\langle v, w\rangle = \langle v, w\rangle_{[0,1]} = \langle v, w\rangle_{[0,1],n}$ etc.

An outline of the paper is as follows. In Section 2 we give the heart of our technique, an investigation into the decay properties of the associated Green's function. As a follow-up, we give, in Section 3, a technique for estimating the total energy in the Green's function. Applications to the stability problem (1.4) occupy Section 4, and applications to convergence when data come from sampling a function having two derivatives is the subject of Section 5.

2. DECAY ESTIMATES FOR THE GREEN'S FUNCTION

The main result of this section is stated in precise form in Theorem 2.1 at the end.

We start by motivating why decay estimates for the Green's (kernel) function are connected with stability bounds of the form (1.4) (see (2.5) below). For $y \in [0, 1]$, we define the Green's function $g = g_n(x ; y)$ centered at y by the relation

$$\langle g, v \rangle + \lambda (g'', v'') = v(y), \quad \text{for all} \quad v \in H^2. \tag{2.1}$$

It exists and is unique ; e.g. by the Lax-Milgram lemma. We remark that, unless y is a design point, the Green's function, as given in (2.1), is not a cubic spline on the design interval containing y.

The Euler-Lagrange equation corresponding to the minimization problem (1.1) reduces to finding $s \in H^2$ such that

$$\langle s, v \rangle + \lambda (s'', v'') = \langle d, v \rangle , \quad \text{for all} \quad v \in H^2. \tag{2.2}$$

It is then immediate from (2.1) that

$$s(y) = \langle d, g \rangle . \tag{2.3}$$

We shall assume that $[0, 1]$ has been subdivided into disjoint (except for endpoints) closed intervals $I_j = [y_j, y_{j+1}]$, $j = 0, 1, \ldots, J = J(\mathcal{D}_n)$,

$$[0, 1] = I_0 \cup I_1 \cup \ldots \cup I_J . \tag{2.4}$$

The subdivision points y_j are, in general, distinct from the design points x_i. The subdivision (2.4) will, as indicated, depend on n ; it could also depend on y, but we have found no use for this extra generality. Thus the subdivisions (2.4) will be independent of y in this investigation. A major point in this section is to introduce reasonably practical conditions on the subdivisions relating them to properties of the design family (1.2). In applications, the central problem will be to construct such a suitable partition.

From (2.3) and (2.4) it follows by the triangle inequality and by Cauchy-Schwarz' inequality that

$$
\begin{aligned}
|s(y)| &= |\langle d, g \rangle| \leq \sum_{j=0}^{J} |\langle d, g \rangle_{I_j}| \\
&\leq \sum_{j=0}^{J} |d|_{I_j} |g|_{I_j} \leq \sum_{j=0}^{J} (f_j)^{1/2} |d|_{\infty,j} E_j
\end{aligned}
\tag{2.5}
$$

where

$$|d|_{\infty,j} = |d|_{\infty,I_j} = \max_{x_i \in I_j} |d_i| , \tag{2.6}$$

$$f_j = f(I_j) = (\text{number of design points in } I_j)/n , \tag{2.7}$$

and

$$E_j^2 = E^2(I_j) = |g|_{I_j}^2 + \lambda \|g''\|_{I_j}^2 . \tag{2.8}$$

Here f_j denotes the fraction of design points in I_j (with design points x_i counted twice if they coincide with an endpoint of I_j other than 0 or 1), and E_j will be referred to as the *energy* of g contained in I_j.

At the heart of our argument is a description of properties of the intervals $I_j = I_j(\mathscr{D}_n)$ such that the energy E_j exhibits a constant fraction decay as we move left or right from a central interval I_{j_0} containing y. To be precise, we shall seek the subdivisions (2.4) so that, with $p < 1$ independent of n,

$$E_j \leq p^{|j - j_0| - 1} \mathscr{E} , \quad y \in I_{j_0} , \tag{2.9}$$

where $\mathscr{E}$ is the total energy of g in $[0, 1]$,

$$\mathscr{E}^2 = \mathscr{E}^2(y) = |g|^2 + \lambda \|g''\|^2 . \tag{2.10}$$

When this has been accomplished, the result will be reported back into (2.5). Techniques for estimating $\mathscr{E}(y)$ will be given in Section 3.

We proceed to give an analysis of conditions that may be placed on the intervals I_j so that the constant fraction energy decay estimate (2.9) holds. We give the details for going to the right of $I_{j_0} \ni y$; the analysis to the left is completely analogous. For $j \geq j_0 + 1$ let

$$L_j = |I_j| = y_{j+1} - y_j , \tag{2.11}$$

$$R_{j+1} = I_{j+1} \cup I_{j+2} \cup \dots \cup I_J = [y_{j+1}, 1] , \tag{2.12}$$

$$Q_{j+1}^2 = E_{j+1}^2 + E_{j+2}^2 + \cdots + E_J^2 . \tag{2.13}$$

In particular, Q_{j+1} is the energy of g contained in R_{j+1}. Let $0 \leq k_j(x) \leq 1$ be a cut-off function such that

$$k_j(x) = \begin{cases} 0, & x \leq y_j, \\ 1, & x \geq y_{j+1}. \end{cases} \tag{2.14}$$

By a scaling argument from the case $L_j = 1$, it is clear that we may assume

$$\| k_j^{(\gamma)} \|_\infty \leq C L_j^{-\gamma}, \quad k^{(\gamma)} = d^\gamma k / dx^\gamma, \quad \gamma = 1, 2, 3, 4, \tag{2.15}$$

with C a universal constant.

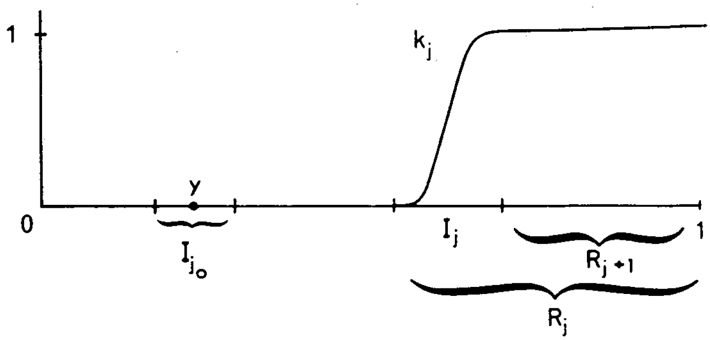

Figure 1.1.

Writing $k = k_j$ for simplicity in notation, we have, since k is nonnegative, and by (2.14),

$$\begin{aligned} Q_{j+1}^2 &= \langle g, g \rangle_{R_{j+1}} + \lambda (g'', g'')_{R_{j+1}} \leq \langle g, kg \rangle + \lambda (g'', kg'') \\ &= [\langle g, kg \rangle + \lambda (g'', (kg)'')] - \lambda (g'', k'' g) - 2 \lambda (g'', k' g'). \end{aligned} \tag{2.16}$$

From (2.1) the terms in square brackets equal $(kg)(y)$, which vanishes since $k(y)$ does.

We proceed with some elementary transformations of the two remaining terms on the right of (2.16). (In the interest of readability we give the trivial details.)

Integrating by parts,

$$(g'', k' g') = - \int g' (k' g')' = - \int g' k'' g' - \int g' k' g''$$

so that, integrating by parts once more,

$$2(g'', k' g') = -\int g' k'' g' = \int g k'' g'' + \int g k''' g' . \tag{2.17}$$

Further, again integrating by parts,

$$\int g k''' g' = -\int g k^{iv} g - \int g' k''' g$$

or,

$$\int g k''' g' = \frac{1}{2} \int g k^{iv} g . \tag{2.18}$$

Combining (2.17) and (2.18),

$$2(g'', k' g') = \int g k'' g'' - \frac{1}{2} \int k^{iv} g^2$$

so that from (2.16), noting that by (2.14) the integrals involved actually extend only over I_j,

$$Q_{j+1}^2 \leqslant -2\lambda \int_{I_j} g'' k'' g + \frac{\lambda}{2} \int_{I_j} k^{iv} g^2 .$$

Using (2.15), the Cauchy-Schwarz inequality, the arithmetic-geometric mean inequality, and changing the universal constant C slightly,

$$\begin{aligned} Q_{j+1}^2 &\leqslant \frac{2C\lambda}{L_j^2} \int_{I_j} |g''|\,|g| + \frac{C\lambda}{2L_j^4} \int_{I_j} g^2 \\ &\leqslant C\left[\frac{\lambda}{L_j^2} \|g''\|_{I_j} \|g\|_{I_j} + \frac{\lambda}{L_j^4} \|g\|_{I_j}^2 \right] \leqslant C\left[\lambda \|g''\|_{I_j}^2 + \frac{\lambda}{L_j^4} \|g\|_{I_j}^2 \right]. \end{aligned} \tag{2.19}$$

The idea is now to convert the right hand side of (2.19) into an expression involving the energy E_j^2. Noting that the first term is a part of this energy, while the second is not, we attack the latter. For this a mild condition on the distribution of design points x_i within I_j will be introduced ; namely, (2.20) below. We isolate the present step as a separate lemma.

LEMMA 2.1 : *Let I be an interval of length L, and let $f(I)$ be the fraction of design points contained in I* (cf. (2.7)). *Assume that there is a constant $c_0 > 0$ such that*

$$c_0 L^2 f(I) \leqslant \frac{1}{n} \sum_{x_i \in I} (x_i - \xi)^2 \equiv |x - \xi|_I^2 , \text{ for all } \xi \text{ in } I. \tag{2.20}$$

Then there exists a constant $C = C(c_0^{-1})$ *such that for any* $u \in H^2(I)$,

$$\|u\|_I^2 \leqslant C\left[\frac{L}{f(I)}\,|u|_I^2 + L^4\|u''\|_I^2\right]. \tag{2.21}$$ □

Remark 2.1 : The condition (2.20) ensures that the design points x_i in I are *not* clustered only around *one* point. As an example, if $I = [x_i, x_{i+1}]$ is a design interval, then $f(I) = 2/n$, and (2.20) is fulfilled with $c_0 = 1/4$. □

The proof of Lemma 2.1 will be postponed until later in this section. Admitting it for the moment, and assuming that (2.20) holds for $I = I_j$, uniformly in j and n, we obtain from (2.19) and (2.21),

$$Q_{j+1}^2 \leqslant C(c_0^{-1})\left(\lambda\|g''\|_{I_j}^2 + \frac{\lambda}{L_j^3 f_j}\,|g|_{I_j}^2\right). \tag{2.22}$$

To proceed we now make the further assumption that

$$\frac{\lambda}{L_j^3 f_j} \leqslant C_1\,, \quad \text{independent of } j \text{ and } n\,. \tag{2.23}$$

Remark 2.2 : When (2.23) holds one has the following bound for the number $J+1$ of intervals in the partition (2.4),

$$(J+1) \leqslant (2\,C_1)^{1/4}\,\lambda^{-1/4}. \tag{2.24}$$

For, by Hölder's inequality with conjugate indices 4/3 and 4,

$$(J+1) = \sum_0^J 1 = \sum_0^J L_j^{3/4}\,.\,L_j^{-3/4} \leqslant \left[\sum_0^J L_j\right]^{3/4}\left[\sum_0^J L_j^{-3}\right]^{1/4}.$$

Since $\sum_0^J L_j = 1$, and by (2.23),

$$(J+1) \leqslant C_1^{1/4}\lambda^{-1/4}\left[\sum_0^J f_j\right]^{1/4}.$$

Since the intervals I_j are disjoint except for endpoints, $\sum_0^J f_j \leqslant 2$, and (2.24) follows. □

Assuming (2.23) we now have from (2.22) that, with $K = K(c_0^{-1}, C_1)$, cf. (2.13),

$$Q_{j+1}^2 \leqslant KE_j^2 = K(Q_j^2 - Q_{j+1}^2), \quad j \geqslant j_0 + 1\,, \tag{2.25}$$

so that with $p = p(c_0^{-1}, C_1) = (K/(K+1))^{1/2} < 1$,

$$Q_{j+1}^2 \leq p^2 Q_j^2, \quad j \geq j_0 + 1 . \tag{2.26}$$

Performing the analogous construction to the left, applying induction to (2.26), and noting that $E_j \leq Q_j \leq \mathscr{E}$, we have the decay estimate (2.9), subject to the conditions (2.20) and (2.23) relating the partitioning $I_j = I_j(n)$ of (2.4) to the design family (1.2).

We next prove Lemma 2.1.

Proof of Lemma 2.1 : By the mean-value theorem for averages,

$$\frac{1}{f(I)} |u|_I^2 = \frac{1}{nf(I)} \sum_{x_i \in I} u^2(x_i) = u^2(\eta), \text{ for some } \eta \in I . \tag{2.27}$$

Now,

$$\|u\|_I^2 = \int_I u^2(x)\, dx = Lu^2(\eta) + \int_I (u^2(x) - u^2(\eta))\, dx . \tag{2.28}$$

By Cauchy-Schwarz' inequality and the arithmetic-geometric mean inequality,

$$\int_I (u^2(x) - u^2(\eta))\, dx = \int_I \left[\int_\eta^x 2\, u(\tau)\, u'(\tau)\, d\tau \right] dx \leq \tag{2.29}$$

$$\leq 2\, L \|u\|_I \|u'\|_I \leq \frac{\|u\|_I^2}{2} + 2\, L^2 \|u'\|_I^2 .$$

Hence from (2.27)-(2.29),

$$\|u\|_I^2 \leq 2\, Lu^2(\eta) + 4\, L^2 \|u'\|_I^2 = \frac{2\, L}{f(I)} |u|_I^2 + 4\, L^2 \|u'\|_I^2 . \tag{2.30}$$

The first term on the right of (2.30) already occurs in the energy-related expression on the right of our desired result (2.21). To treat the second term on the right of (2.30), we first derive an auxiliary result, (2.31) below. For this, let ξ_0 be the point where $|u(\xi)|^2$ attains its minimum on I, and let ξ_1 be the minimum point for $|u'(\xi)|^2$. By the mean-value theorem,

$$(x - \xi_0)\, u'(\bar{\xi}) = u(x) - u(\xi_0), \text{ for some } \bar{\xi} \in \text{Int}\,(x, \xi_0) .$$

It follows that

$$(x - \xi_0)^2\, (u'(\xi_1))^2 \leq (u(x) - u(\xi_0))^2 \leq u^2(x), \text{ for all } x \in I ,$$

since either $u(\xi_0) = 0$ or u does not change sign on I. In particular,

$$|x - \xi_0|_I^2 (u'(\xi_1))^2 \leqslant |u|_I^2$$

so that the assumption (2.20) gives

$$(u'(\xi_1))^2 \leqslant \frac{1}{c_0 L^2 f(I)} |u|_I^2 . \tag{2.31}$$

To apply this estimate to the last term in (2.30) note first that by the mean-value theorem we have for some ξ in I,

$$\| u' \|_I^2 = L(u'(\xi))^2 = L(u'(\xi_1))^2 + L[(u'(\xi))^2 - (u'(\xi_1))^2] . \tag{2.32}$$

Using again Cauchy-Schwarz' inequality and the arithmetic-geometric mean inequality,

$$(u'(\xi))^2 - (u'(\xi_1))^2 = 2 \int_{\xi_1}^{\xi} u'(\tau) u''(t) \, d\tau \leqslant \frac{1}{2L} \| u' \|_I^2 + 2L \| u'' \|_I^2 .$$

From this and (2.31) applied to (2.32),

$$\| u' \|_I^2 \leqslant \frac{2}{c_0 L f(I)} |u|_I^2 + 4 L^2 \| u'' \|_I^2 . \tag{2.33}$$

Reporting back into (2.30) completes the proof of Lemma 2.1. □

We conclude this section by writing up the result of our investigation into the energy decay of the Green's function in precise form, and by giving three additional remarks. For emphasis we relabel the assumptions (2.20) and (2.23) made on the partitioning (2.4).

THEOREM 2.1 : *Let $\mathcal{D} = \{\mathcal{D}_n\}$ be a design family (cf. (1.2)). Assume that for each n a partition $\{I_j\}_0^J$ of* [0,1] *exists (cf. (2.4)) such that with constants $c_0 > 0$ and C_1 independent of n and j,*

$$c_0 L_j^2 f_j \leqslant |x - \xi|_{I_j}^2 , \quad \textit{for all } \xi \textit{ in } I_j \tag{A.1}$$

(cf. (1.5), (2.7) and (2.11) for notation, and cf. also Remark 2.1), and

$$\frac{\lambda}{L_j^3 f_j} \leqslant C_1 . \tag{A.2}$$

(cf. Remark 2.2). Then there exists $p = p(c_0^{-1}, C_1) < 1$ such that for $y \in I_{j_0}$,

$$E_j \leqslant p^{|j - j_0| - 1} \mathscr{E}(y) , \tag{2.34}$$

(cf. (2.8) and (2.10) for notation). □

Remark 2.3 : If $f_j \simeq \omega_j L_j$, where $\omega_j \simeq \omega(y_j)$, for some "frequency function" ω, and if approximate equality holds in (A.2), then $L_j \simeq (\omega_j/\lambda)^{1/4}$. This fits in with the decay properties of the kernel of the equivalent smoothing kernel method given, under rather stringent conditions, in Silverman [6]. □

Remark 2.4 : The energy decay result of Theorem 2.1 gives a pointwise decay estimate for the Green's function. With the break-up (2.4), if $x \in I_k$ and $y \in I_\ell$,

$$|g(x;y)| \leqslant Cp^{|k-\ell|}\, \mathscr{E}(x)\, \mathscr{E}(y)\,.$$

For, by (2.1),

$$\begin{aligned} |g(x;y)| &= |\langle g(\,.\,;x), g(\,.\,;y)\rangle + \lambda(g''(\,.\,;x), g''(\,.\,;y))| \\ &\leqslant \sum_{j=0}^{J} E_j(x)\, E_j(y) \leqslant \left[\sum_{j=0}^{J} p^{|k-j|-1} p^{|\ell-j|-1}\right] \mathscr{E}(x)\, \mathscr{E}(y) \\ &\leqslant Cp^{|k-\ell|}\, \mathscr{E}(x)\, \mathscr{E}(y)\,. \end{aligned}$$

□

Remark 2.5 : The techniques introduced in this section (and the next) can easily be adapted to investigating the pointwise stability of solutions to the fourth order boundary value problem

$$\begin{cases} \lambda s^{iv} + \omega s = \omega d, & \text{for} \quad 0 \leqslant x \leqslant 1\,, \\ s'' = s''' = 0 & \text{for} \quad x = 0,1\,. \end{cases}$$

Here the nonnegative weight function ω is subject to $\int_0^1 \omega = 1$. Since λ may be small, we could be in a singularly perturbed case, with rough ω. The adaptation referred to merely consists of redefining $\langle v, w\rangle_I = \int_I \omega v w\, dx$ and $f(I) = \int_I \omega\, dx$. The obvious analogue of Theorem 2.1 then holds, with the same proof. □

3. ESTIMATING THE TOTAL ENERGY IN THE GREEN'S FUNCTION

In this brief section we shall give methods for estimating the total energy in the Green's function $g = g_n(x;y)$ of (2.1),

$$\mathscr{E}^2(y) = |g|^2 + \lambda \|g''\|^2\,.$$

By (2.1) we have

$$\mathscr{E}^2(y) = g(y)\,. \tag{3.1}$$

Let I be an interval containing y, and assume that (A.1) holds on I,

$$c_0 L^2 f(I) \leq |x - \xi|_I^2, \quad \text{for all} \quad \xi \in I, \tag{A.1}$$

where $L = |I|$, *cf.* (1.5) and (2.7) for notation, and also Remark 2.1. Let further x_i denote any design point in I. Then

$$g(y) = g(x_i) + (g(y) - g(x_i)) = g(x_i) + \int_{x_i}^{y} g'(\tau)\, d\tau$$

so that by Cauchy-Schwarz' inequality,

$$g^2(y) \leq 2\, g^2(x_i) + 2\, L \, \|g'\|_I^2 .$$

Assuming (A.1) we have then from (2.33) in the proof of Lemma 2.1,

$$g^2(y) \leq 2\, g^2(x_i) + \frac{4}{c_0 f(I)} |g|_I^2 + 8\, L^3 \|g''\|_I^2 .$$

Summing this over the design points x_i in I, and dividing by n, we obtain

$$f(I)\, g^2(y) \leq C\, |g|_I^2 + 8\, L^3 f(I) \|g''\|_I^2$$

which, on using (3.1), becomes

$$\mathscr{E}^4(y) = g^2(y) \leq C \left[\frac{1}{f(I)} + \frac{L^3}{\lambda} \right] \mathscr{E}^2(y)$$

so that, with (A.1),

$$\mathscr{E}(y) \leq C \left[\frac{1}{f^{1/2}(I)} + \frac{L^{3/2}}{\lambda^{1/2}} \right] . \tag{3.2}$$

Under the further assumption that

$$0 < c_2 \leq \frac{\lambda}{L^3 f(I)} \tag{A.3}$$

(the complement to (A.2)), we have

$$\mathscr{E}(y) \leq C / f^{1/2}(I) .$$

In our applications we shall sometimes be able to use this with $I = I_{j_0}$, the interval from the subdivision (2.4) which contains the central point y. Then, assuming that (A.1) and (A.3) hold for $I = I_{j_0}$,

$$\mathscr{E}(y) \leq C / f_{j_0}^{1/2}, \quad \text{for} \quad y \in I_{j_0} . \tag{3.3}$$

Finally, we make the following trivial observation : if $y = x_i$ is a design point, then

$$\mathscr{E}^2(x_i) = g(x_i) \leqslant \sqrt{n}\, \mathscr{E}(x_i)$$

so that

$$\mathscr{E}(x_i) \leqslant \sqrt{n}\,. \tag{3.4}$$

4. APPLICATIONS TO POINTWISE STABILITY

From (2.5), which was the motivation for the investigations in Section 2, we have by Theorem 2.1, assuming (A.1) and (A.2) there, that

$$|s(y)| \leqslant \left[\sum_{j=0}^{J} |d|_{\infty,j} (f_j)^{1/2} p^{|j-j_0|-1} \right] \mathscr{E}(y), p < 1\,, \tag{4.1}$$

where $y \in I_{j_0}$ (*cf.* (2.4), (2.6), (2.7) and (2.10) for notation). In this section we examine how to actually construct the fundamental partition (2.4), and estimate the total energy, so that the pointwise stability estimate (1.4) is obtained from (4.1).

We shall treat three examples of meshfamilies. Unfortunately, we have not been able to identify a useful wider class of meshfamilies, and some details of the analysis will therefore be specific to each example.

We start by considering design families where local irregularities in the placement of design points are smoothed out when averaged on the scale of $\lambda^{1/4}$.

4.1. $\lambda^{1/4}$ almost-regular meshes

Assume that there are constants $0 < k_0 < k_1$, independent of n, such that for any interval I of length comparable to $\lambda^{1/4} = \lambda_n^{1/4}$,

$$k_0\, \lambda^{1/4} \leqslant f(I) \leqslant k_1\, \lambda^{1/4}\,, \tag{4.2}$$

where $f(I)$ is the fraction of design points contained in I (*cf.* (2.7)). This, then, is a condition only on the average distribution of design points on the scale $\lambda^{1/4}$. For it to make sense we must have, essentially,

$$\lambda \geqslant n^{-4}\,. \tag{4.3}$$

We further make the weak assumption (A.1) of Theorem 2.1 on the internal distribution of design points within each interval of length $|I| \simeq \lambda^{1/4}$, namely that they do not cluster too much around a single point in

I_j (*cf.* Remark 2.1). We refer to design families satisfying (4.2) and (A.1) when $|I| \simeq \lambda^{1/4}$ as being $\lambda^{1/4}$ *almost-regular meshes*. We may then take for the I_j of (2.4) simply intervals of length $\lambda^{1/4}$ (with some adjustment, perhaps, at one end of [0,1]). Clearly (A.2) of Theorem 2.1 and (A.3) of Section 3 are satisfied and we obtain immediately from (3.3), (4.1) and (4.2) that

$$|s(y)| \leq C \sum_{j=0}^{J} |d|_{\infty,j}\, p^{|j-j_0|-1}, \quad p < 1, \tag{4.4}$$

where $C = \max\limits_j (f_j)^{1/2}/(f_{j_0})^{1/2} \leq (k_1/k_0)^{1/2}$. Clearly, then, the stability estimate (1.4) holds for such meshes.

Remark 4.1 : The local dependence of smoothing splines is clearly identified in the above stability estimate (4.4). For a point y, the terms in the summation on the right hand side of (4.4) bound the contributions of the individual data points to the value of the smoothing spline at y. The essential nature of the local dependence is encapsulated in the exponential decay associated with each of these terms. It shows that the influence of a data point is dominated by an exponential decay the exponent of which depends linearly on the "distance" of the data point from y. □

As remarked above, in (4.3), if $\lambda \ll n^{-4}$ the condition (4.2) for $\lambda^{1/4}$ almost-regular meshes does not make sense. In the interest of testing the power of our general technique, one may wonder what stability results can be obtained for $\lambda < n^{-4}$. We shall give a simple one : assume that the grid family is *quasi-uniform* ; i.e., with $h_i = x_{i+1} - x_i$, the following condition holds

$$\tilde{k}_0\, n^{-1} \leq \min_i\, h_i \leq \max_i\, h_i \leq \tilde{k}_1\, n^{-1}, \tag{4.5}$$

where $\tilde{k}_0$ and $\tilde{k}_1$ are independent of n. We take the intervals I_j in (2.4) as design point intervals, $I_j = [x_j, x_{j+1}]$. Then $f_j = 2/n$ and Theorem 2.1 clearly applies so that by (4.1),

$$|s(y)| < 2^{1/2}\, n^{-1/2} \left[\sum_{j=0}^{n-2} |d|_{\infty,j}\, p^{|j-j_0|-1} \right] \mathscr{E}(y).$$

From (3.4) then, for a design point x_i,

$$|s(x_i)| \leq 2^{1/2} \left[\sum_{j=0}^{n-2} |d|_{\infty,j}\, p^{|j-j_0|-1} \right], \quad p < 1.$$

This establishes the stability of the smoothing splines at the knot points.

Their global stability then follows from the well known stability of the natural spline interpolants on quasi-uniform meshes.

4.2. Meshes which are systematically refined or thinned

We start by describing, in some generality, a rational procedure for finding the subdivisions $\{I_j\}$ in (2.4). Let $F : [0,1] \to [0,1]$, one-to-one and onto, be a strictly monotone continuous function, and let, apart from local irregularities, the gridpoints be "regularly distributed with density dF/dx".

Before turning to discussion of local irregularities, we give our rational procedure for selecting the subdivision. Assume that F' exists, that

$$f_j \simeq \int_{I_j} F' \simeq L_j\, F'(y_j)\,,$$

and, if $y_j = y(j)$ for some function $y(\,.\,)$, that

$$L_j = y_{j+1} - y_j \simeq y'(j) \left(= \frac{dy}{dj}(j) \right).$$

Considering (A.2) of Theorem 2.1, and its complement (A.3) of Section 3, it is natural to try to find $y(\,.\,)$ as a solution of

$$\frac{\lambda}{(y')^4\, F'(y)} = 1, y(0) = 0\,,$$

i.e.,

$$y' = \left[\frac{\lambda}{F'(y)} \right]^{1/4}, \quad y(0) = 0\,. \tag{4.6}$$

Assuming that we have a solution $y(\,.\,)$ of this separable differential equation, we then attempt

$$y_j = y(j)\,, \quad j = 0, \dots, J\,, \tag{4.7}$$

with possible adjustment, if needed, at the right end. It remains, of course, to verify in each application (i.e., for the design family) that (A.1), (A.2), and (A.3) if desired, hold.

In order to show that the effect of local irregularities can often be ignored, we consider the following situation. Assume that the design points are $2\,n + 1$ in number, and let every other point be given in terms of the inverse function to F as

$$x_i = F^{-1}\left[\frac{i}{2\,n} \right], \quad i = 0, 2, 4, \dots, 2\,n\,. \tag{4.8}$$

Each of the points $x_0, x_2, \ldots, x_{2n-2}$ now has a companion $x_1, x_3, \ldots, x_{2n-1}$, possibly chosen quite close to it.

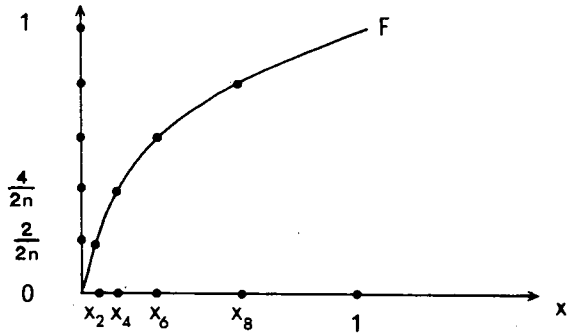

Figure 4.1.

In fact, we examine this situation when, in (4.8), $F(y) = y^{\alpha}$, $0 < \alpha$. The case $\alpha < 1$ corresponds to a systematic meshrefinement towards the left endpoint, whereas $\alpha > 1$ corresponds to a mesh that is thinned towards the left (*cf. fig.* 4.1). The procedure in (4.6), (4.7) yields

(4.9) $$y_j^{\frac{3+\alpha}{4}} = \frac{\alpha+3}{4}\left[\frac{\lambda}{\alpha}\right]^{1/4} j\,.$$

For simplicity in writing, we take instead

(4.10) $$y_j^{\frac{3+\alpha}{4}} = \lambda^{1/4} j\,, \qquad j = 0, \ldots, J \simeq \lambda^{-1/4}\,,$$

assuming proper adjustment on the right.

We shall check that, under certain conditions on λ_n to be given below, (A.1), (A.2) and (A.3) hold for y_j as given in (4.10). Firstly,

(4.11) $$L_j = y_{j+1} - y_j = y_{j+1}\left[1 - \left[\frac{y_j}{y_{j+1}}\right]\right]$$

$$= y_{j+1}\left[1 - \left[1 - \frac{1}{j+1}\right]^{\frac{4}{3+\alpha}}\right] \simeq y_{j+1}(j+1)^{-1}\,.$$

Here and below, $\simeq$ indicates equivalence of quantities, independently of j. The constants in such equivalences will depend on α. For notational simplicity we do not trace this dependence, but if desired, this is easily done. Secondly (*cf. fig.* 4.1)

$$(4.12)\quad f_j \simeq \left(\# i : y_j \leqslant \left[\frac{i}{2n} \right]^{\frac{1}{\alpha}} \leqslant y_{j+1} \right) / n$$

$$\simeq \left(\# i : 2\, n y_{j+1}^{\alpha} \left[1 - \frac{1}{(j+1)} \right]^{\frac{4\alpha}{3+\alpha}} \leqslant i \leqslant 2\, n y_{j+1}^{\alpha} \right) / n$$

$$\simeq y_{j+1}^{\alpha} (j+1)^{-1}$$

provided each I_j contains at least a few gridpoints. For $\alpha < 1$, meshrefinement, an elementary calculation shows that this is the case if

$$(4.13)\qquad \lambda_n \geqslant c_\alpha n^{-4}, \qquad \alpha < 1,$$

while for the thinning mesh,

$$(4.14)\qquad \lambda_n \geqslant c_\alpha n^{-\left(\frac{3+\alpha}{\alpha}\right)}, \qquad \alpha > 1.$$

The positive constants c_α can easily be made explicit, if desired. It is easy to check that (A.1) is satisfied. For (A.2) and (A.3), by (4.10), (4.11) and (4.12),

$$\frac{\lambda}{L_j^3 f_j} \simeq \frac{\lambda (j+1)^4}{y_{j+1}^{3+\alpha}} = 1 .$$

Thus, under the condition (4.13) or (4.14), we may apply (4.1) and (3.3) so that for $y \in I_{j_0}$,

$$|s(y)| \leqslant C \left[\sum_{j=0}^{J} \left[\frac{f_j}{f_{j_0}} \right]^{1/2} p^{|j - j_0| - 1} \right] |d|_\infty$$

which by (4.12) and (4.10) gives

$$|s(y)| \leqslant C \left[\sum_{j=0}^{J} \left[\frac{(j+1)}{(j_0+1)} \right]^{\frac{3(\alpha-1)}{2(3+\alpha)}} p^{|j - j_0| - 1} \right] |d|_\infty .$$

It is easily seen that

$$\left[\frac{j+1}{j_0+1} \right]^{\alpha - 1} \leqslant C (1 + |j - j_0|)^{|\alpha - 1|}$$

so that, finally,

$$\begin{aligned}|s(y)| &\leqslant C\left[\sum_{j=0}^{J}(1+|j-j_0|)^{\frac{3|\alpha-1|}{2(3+\alpha)}}\,p^{|j-j_0|-1}\right]|d|_\infty\\ &\leqslant C\ |d|_\infty,\quad \text{independent of } y,\end{aligned}$$

since the algebraic growth of the first factor in the sum is held in check by the exponential decay of the second. Stability therefore holds.

Remark 4.2 : Although we have only examined systematically refined and thinned grids in this section, the ideas generalize. In fact, if a mesh has a subgrid obtained from a systematic refinement such that the remaining points satisfy the conditions (A.1) and (A.3), then stability will be guaranteed. □

4.3. A "Universal" meshrefinement

Let $x_0 = 0$ and $x_1 = x_1(n)$ be given. let further

$$F_n^{-1}(z) = e^{C_n(z-1)}, \qquad \text{for} \qquad z \geqslant 1/n,$$

and

$$x_i = F_n^{-1}\left[\frac{i}{n-1}\right], \qquad i = 1, \dots, n-1. \tag{4.15}$$

In particular, this determines $C_n = C_n(x_1(n)) \simeq \ln\left[\frac{1}{x_1(n)}\right]$. A distinguishing feature of these grids is that, with $k = k_n$ a factor independent of i,

$$x_{i+1} - x_i = kx_i, \qquad i \geqslant 1.$$

Since $F_n'(y) = 1/(C_n y)$ they are also, in a sense, a natural limit of our previous meshrefinements as α approaches zero.

The rational guess of Section 4.2, see (4.6) and (4.7), easily adapts to the present case in which $F_n$depends on n. Instead of (4.10) we are now led to

$$y_j^{3/4} = \frac{3}{4}(\lambda C_n)^{1/4} j.$$

Provided

$$\lambda_n \geqslant C \max\left[\frac{x_1^3(n)}{C_n}, \frac{1}{n^4 C_n^5}\right]$$

one obtains pointwise stability.

We remark that in this case the distribution function $\{ \# i : x_i \leqslant s\}/n$ approaches unity for all $0 < s \leqslant 1$, if $x_1(n)$ approaches zero. Thus, in the limit, everything is concentrated at the origin, and hence the name "universal" meshrefinement. The limiting frequency function is a Dirac delta at the origin. (The name "universal" is also motivated by the following folklore: meshrefinement according to (4.15), with suitably chosen second meshpoint $x_1 = x_1(n)$, is sufficient to resolve any localized singularity found in actual practice.)

5. APPLICATIONS TO POINTWISE CONVERGENCE

Let data d_i equal $D(x_i)$ for some underlying function $D(x)$ in H^2. Then from (2.1) and (2.3) (with exact measurements),

$$s(y) - D(y) = -\lambda(D'', g'') . \tag{5.1}$$

We first give a simple result, (5.2) below, which nevertheless is of interest since it shows that pointwise convergence occurs for an extremely wide range of design families. By Cauchy-Schwarz' inequality,

$$|(s-D)(y)| \leqslant \lambda \; \|g''\| \; \|D''\| \leqslant \lambda^{1/2} \mathscr{E}(y) \; \|D''\| \; .$$

By (3.2), taking $y \in [x_{i_0}, x_{i_0+1}]$ a meshinterval, since then $f_{i_0} = 2/n$, and (A.1) holds with $c_0 = 1/4$,

$$\mathscr{E}(y) \leqslant C \left(\sqrt{n} + \left[\frac{h_{i_0}^3}{\lambda} \right]^{1/2} \right)$$

where $h_{i_0} = x_{i_0+1} - x_{i_0}$. Thus

$$|(s-D)\,(y)| \leqslant C\,((\lambda n)^{1/2} + h_{i_0}^{3/2})\;\|D''\|\;, \quad y \in [x_{i_0}, x_{i_0+1}]\,, \tag{5.2}$$

with C completely independent of the design family and of y. (In fact, $C = 1$ works.) Hence convergence occurs for $D \in H^2[0, 1]$ if $\lambda_n n$ and $h_{i_0} = h_{i_0}(y, n)$ both approach zero as n tends to infinity.

We next consider the examples of design families given in Section 4, and we shall assume that $D(x)$ is twice continuously differentiable on [0, 1]. The object is to derive estimates which are sharper than (5.2). For the breakup (2.4) we now have from (5.1) and Theorem 2.1, using the Cauchy-Schwarz inequality,

$$\begin{aligned} |(S-D)(y)| &\leqslant \lambda \left[\sum_{j=0}^{J} \int_{I_j} |g''| \right] \|D''\|_\infty \\ &\leqslant \lambda^{1/2} \left[\sum_{j=0}^{J} L_j^{1/2} E_j \right] \|D''\|_\infty \leqslant \lambda^{1/2} \left[\sum_{j=0}^{J} L_j^{1/2} p^{\,|j-j_0|-1} \right] \mathscr{E}(y)\; \|D''\|_\infty \,. \end{aligned} \tag{5.3}$$

5.1. $\lambda^{1/4}$ almost-regular meshes

Using (3.3), which applies in this situation in which $L_j \simeq \lambda^{1/4} \simeq f_j$, (5.3) immediately gives the following convergence result :

$$\|s - D\|_\infty \leq C\lambda^{1/2} \|D''\|_\infty . \tag{5.4}$$

Again, localization is easily performed.

Recall that $\lambda \geq n^{-4}$ is required. It is well known that the natural spline interpolant to D has convergence bounded by $C\bar{h}^2 \|D''\|_\infty$, $\bar{h} = \max (x_{i+1} - x_i)$, completely independent of the mesh-distribution. One may wonder what happens for the smoothing spline when λ_n is very small. We shall give a partial answer.

Let $\underline{h} = \min (x_{i+1} - x_i)$ and assume that

$$\lambda_n \leq \underline{h}^3/n \ (\leq n^{-4}) . \tag{5.5}$$

We may then take $I_j = [x_j, x_{j+1}]$ as the partition intervals in (2.4). Theorem 2.1 applies under the condition (5.5), and we use (3.2) to estimate $\mathscr{E}(y)$. Thus, from (5.3),

$$\|s - D\|_\infty \leq C\lambda^{1/2} \bar{h}^{1/2}\left[\sqrt{n} + \frac{\bar{h}^{3/2}}{\lambda^{1/2}}\right] \|D''\|_\infty \leq C\bar{h}^2 \|D''\|_\infty ,$$

independent of the distribution of design points, but subject to (5.5). In the formal limit $\lambda = 0$ we recover the result for the spline interpolant mentioned above.

5.2. Meshes which are, disregarding local irregularities, systematically refined or thinned

Using the breakup of Section 4.2 (recall the conditions (4.13) and (4.14)) we have from (5.3), and from (4.11), (4.12) and (3.3),

$$\begin{aligned} |(s - D)(y)| &\leq \\ &\leq C\lambda^{1/2}\left[\sum_{j=0}^{J} \left[\frac{y_{j+1}(j_0 + 1)}{y_{j_0+1}^{\alpha}(j+1)}\right]^{1/2} p^{|j - j_0| - 1}\right] \|D''\|_\infty \\ &= \lambda^{\frac{2}{3+\alpha}}\left[\sum_{j=0}^{J} ((j+1)(j_0+1)^3)^{\frac{1-\alpha}{2(3+\alpha)}} p^{|j - j_0| - 1}\right] \|D''\|_\infty , \end{aligned} \tag{5.6}$$

after some elementary algebra.

For the case $\alpha < 1$ of meshrefinement toward the left,

$$|(j+1)| \leqslant 1 + j_0 + |j - j_0|$$

so that from (5.6),

$$|(s-D)(y)| \leqslant C\lambda^{\frac{2}{3+\alpha}} (j_0+1)^{\frac{4(1-\alpha)}{2(3+\alpha)}} \|D''\|_\infty$$
$$= C\lambda^{1/2}(y_{j_0+1})^{\frac{1-\alpha}{2}} \|D''\|_\infty, \quad \lambda_n \geqslant c_\alpha n^{-4}.$$

Thus we have uniform convergence of order $\lambda^{1/2}$. The convergence estimate gets better as y moves towards the left, and in the first interval, $y \in I_0 = \left[0, \lambda^{\frac{1}{3+\alpha}}\right]$, it is actually $C\lambda^{\frac{2}{3+\alpha}} \|D''\|_\infty$.

A similar calculation in the case of meshthinning, $\alpha > 1$, under the constraint (4.14), $\lambda_n \geqslant c_\alpha n^{-(3+\alpha)/\alpha}$, establishes uniform convergence of order $\lambda^{\frac{2}{3+\alpha}}$. The convergence estimate now gets better towards the right, where it approaches $\lambda^{1/2}$.

We remark that if $D(x)$ has four derivatives and $D''(x) = D'''(x) = 0$ for $x = 0, 1$, then from integrating by parts in (5.1),

$$(s-D)(y) = \lambda(g, D^{iv}).$$

Analyses similar to those above are readily made, but we leave the details to the interested reader (*cf.* (2.21)).

We also leave it to the reader to investigate pointwise convergence in the « universal » meshrefinement of Section 4.3.

ACKNOWLEDGEMENT

The third author thanks the Center for Mathematical Analysis, Canberra, and the Commonwealth Scientific and Industrial Research Organization, Canberra, for their hospitality during the Fall of 1987. He also thanks the National Science Foundation, USA, and the Army Research Office through the Mathematical Sciences Institute at Cornell for their support. The first two authors wish to acknowledge the continuing support of the Centre for Mathematical Analysis, ANU, Canberra.

REFERENCES

[1] C. DE BOOR, A Practical Guide to Splines, Springer, New York, 1978.

[2] J. C. HOLLADAY, Smoothest curve approximation, Math. Comp. 11, 1957, 233-243.

[3] F. R. DE HOOG and R. S. ANDERSSEN, Convergence of kernel functions for cubic smoothing splines on non-equispaced grids, Austral. J. Statist. 30A, 1988, 90-99.

[4] C. H. REINSCH, Smoothing by spline functions, Numer. Math. 10, 1967, 177-183.

[5] I. J. SCHOENBERG, Spline functions and the problem of graduation, Proc. Nat. Acad. Sci. 52, 1964, 947-950.

[6] B. W. SILVERMAN, Spline smoothing : the equivalent variable kernel method, Ann. Statist. 12, 1984, 898-916.

[7] E. WHITTAKER, On a new method of graduation, Proc. Edinburgh Math. Soc. 41, 1923, 63-75.