GEORGIA INSTITUTE OF TECHNOLOGY OFFICE OF CONTRACT ADMINISTRATION

PROJECT ADMINISTRATION DATA SHEET

[x] ORIGINAL [ ] REVISION NO. ______

Project No. G-37-607 GTRI/~~GIT~~ DATE 7 / 2 / 84

Project Director: Dr. Marcus C. Spruill School/~~Lab~~ Math

Sponsor: National Science Foundation

Type Agreement: Grant No. DMS-8401759

Award Period: From 7/1/84 To ~~12/31/85~~* 12/31/86 (Performance) ~~3/31/86~~ 3/31/87 (Reports)

Sponsor Amount: This Change / Total to Date

Estimated: $ ______ $ 15,800

Funded: $ ______ $ 15,800

Cost Sharing Amount: $ 4,393 Cost Sharing No: G-37-315

Title: "Mathematical Sciences: Optimal Experimental Designs"

ADMINISTRATIVE DATA OCA Contact Lynn Boyd x4820

1) Sponsor Technical Contact:

Jerome Sacks
National Science Foundation
Data Support Services Section
Div. of Math & Computer Science
1800 G. St., N.W. Washington, DC 20550
(202) 357-9764

2) Sponsor Admin/Contractual Matters:

Myra B. Galinn
Grants Official
National Science Foundation
Washington, DC 20550
(202) 357-9630

Defense Priority Rating: n/a Military Security Classification: n/a

(or) Company/Industrial Proprietary: n/a

RESTRICTIONS

See Attached NSF Supplemental Information Sheet for Additional Requirements.

Travel: Foreign travel must have prior approval – Contact OCA in each case. Domestic travel requires sponsor approval where total will exceed greater of $500 or 125% of approved proposal budget category.

Equipment: Title vests with GIT

COMMENTS:

*includes usual 6-month unfunded flexibility period.

This will be a continuing grant and $16,500 is anticipated Level of funding for FY 85

RECEIVED JUL 1984 Research Reports Office

COPIES TO: Sponsor I.D. #02.107.000.84.031

Project Director
Research Administrative Network
Research Property Management
Accounting

Procurement/EES Supply Services
Research Security Services
Reports Coordinator (OCA)
Research Communications (2)

GTRI
Library
Project File
Other Newton

FORM OCA 4:383

ORGIA INSTITUTE OF TECHNOLOGY

OFFICE OF CONTRACT ADMINISTRATION

SPONSORED PROJECT TERMINATION/CLOSEOUT SHEET

Date 2/27/87

ject No. G-37-607

School/XXX Mathematics

ludes Subproject No.(s) N/A

ject Director(s) M. C. Spruill

GTRC / XXXX

onsor National Science Foundation

le "Mathematical Sciences: Optimal Experimental Designs"

fective Completion Date: 12/31/86 (Performance) 3/31/87 (Reports)

ant/Contract Closeout Actions Remaining:

- [ ] None
- [ ] Final Invoice or Final Fiscal Report
- [ ] Closing Documents
- [X] Final Report of Inventions - Questionnaire to P.I.
- [ ] Govt. Property Inventory & Related Certificate
- [ ] Classified Material Certificate
- [ ] Other

ontinues Project No.

Continued by Project No.

OPIES TO:

oject Director
esearch Administrative Network
esearch Property Management
ccounting
ocurement/GTRI Supply Services
esearch Security Services
eports Coordinator (OCA)
XXXXXXX

Library
GTRC
~~Research Communications (2)~~
Project File
Other Ina Lashley
Angela Jones
~~Russ Embry~~

ORM OCA 69.285

G-37-601

FINAL TECHNICAL REPORT

OPTIMAL EXPERIMENTAL DESIGNS

M.C. Spruill

Prepared for

National Science Foundation
Washington, DC 20550

Under

Grant No. DMS 8401759

September 1986

GEORGIA INSTITUTE OF TECHNOLOGY
A UNIT OF THE UNIVERSITY SYSTEM OF GEORGIA
SCHOOL OF MATHEMATICS
ATLANTA, GEORGIA 30332

1986

APPENDIX VI

NATIONAL SCIENCE FOUNDATION
Washington, D.C. 20550

FINAL PROJECT REPORT
NSF FORM 98A

PLEASE READ INSTRUCTIONS ON REVERSE BEFORE COMPLETING

PART I–PROJECT IDENTIFICATION INFORMATION

1. Institution and Address	2. NSF Program	3. NSF Award Number
Georgia Institute of Technology Atlanta, GA 30332	Probability & Statistics	DMS 8401759
	4. Award Period From 1984 To 1986	5. Cumulative Award Amount $32,800

6. Project Title

Optimal Experimental Designs

PART II–SUMMARY OF COMPLETED PROJECT *(FOR PUBLIC USE)*

In scientific experiments in which the influence of certain independent variables on a response variable is to be assessed there are good and bad choices of the values of the independent variables. Each run of the experiment has a cost and if this cost is high it is especially important to select the values of the independent variable correctly. Our research has centered on these problems when the information concerns evaluation of the relation of the response to a single variable at an inaccessible value of the independent variable. This includes extrapolation and interpolation problems. We quite successfully solved these problems, including in the former case the extrapolation of rates of change.

PART III–TECHNICAL INFORMATION *(FOR PROGRAM MANAGEMENT USES)*

1. ITEM *(Check appropriate blocks)*	NONE	ATTACHED	PREVIOUSLY FURNISHED	TO BE FURNISHED SEPARATELY TO PROGRAM Check (✓)	Approx. Date
a. Abstracts of Theses					
b. Publication Citations		X		X	1987
c. Data on Scientific Collaborators					
d. Information on Inventions					
e. Technical Description of Project and Results					
f. Other *(specify)*					

2. Principal Investigator/Project Director Name *(Typed)*	3. Principal Investigator/Project Director Signature	4. Date
M. C. Spruill		10-1-86

NSF Form 98A (5-78) Supersedes All Previous Editions

Form Approved OMB No. 99R0013

FINAL TECHNICAL REPORT

Optimal Experimental Designs

Prepared by

M. C. Spruill
School of Mathematics
Georgia Institute of Technology
Atlanta, Georgia 30332

Prepared for

National Science Foundation
Washington, DC 20550

Grant No. DMS 8401759

September 1986

Optimal Experimental Designs

by

M. C. Spruill
Georgia Institute of Technology

Table of Contents

Introduction

Over the course of the last few years, the author of this report has developed a technique of searching for optimal designs. The method is general enough to apply to discrete or continuous data from time series, whether stationary or not, vector data, or scalar data and can be used in finding good designs for estimating any linear functional of interest. A basic characteristic of this technique is that a companion variational problem is identified which, although not necessarily easily solved, is at least devoid of any considerations involving measures. The interplay between the design problem and the variational problem has led to the solution of several sample problems. Each new statistical design problem results in a new variational problem. Reported below are the results of this program applied to the extrapolation of derivatives and to the interpolation of functions. As can be seen below these results are almost complete; only the case of non-parametric interpolation deserving further study.

The first chapter contains the results pertaining to extrapolation of derivatives and at the end contains the key theorem validating the technique of search in the case of scalar observations.

Chapter two deals with interpolation and employs the polynomials developed in chapter three. For more information on these see also De Boor and Rice (1982), Lebedev (1968), Achiezer (1956) and Spruill (1987).

Optimal Extrapolation of Derivatives

1.1. Introduction

For estimating the k^{th} derivative of a mean function outside the interval on which observations can be taken the following generalized Hoel-Levine procedure is investigated. Let the experimenter have control over the values x within the interval [-1,1] at which the observations can be taken and let $c \geq 1$ if $k > 0$ or $c > 1$ if $k = 0$. To estimate $\theta^{(k)}(c)$ take observations only at the $m \geq k+1$ distinct points $x_j = -\cos(\frac{(j-1)\pi}{m-1})$, allocating a proportion

$$\xi_j = \frac{|\phi_{x_j}^{(k)}(c)|}{\sum_{i=1}^{m} |\phi_{x_i}^{(k)}(c)|}$$

at x_j, $j = 1,\ldots,m$. Here ϕ_{x_j} are the Lagrange interpolation polynomials. Estimate $\theta^{(k)}(c)$ by

$$\sum_{i=1}^{m} \bar{y}(x_i)\phi_{x_i}^{(k)}(c),$$

the k^{th} derivative of the polynomial of degree m-1 passing through $\{(x_i,\bar{y}(x_i))\}_{i=1}^{m}$ evaluated at c, where $\bar{y}(x_i)$ is the arithmetic mean of the observations at x_i.

The given procedure is shown to be optimal for the unbiased estimation of $\theta^{(k)}(c)$ when the observations are uncorrelated, $E[Y(x)] = \theta(x)$, $Var(Y(x))$ is constant on [-1,c], and θ is a polynomial of degree m-1. Furthermore, if θ is not a polynomial, but close, then the given procedure performs nearly as well as an optimal estimator in terms of maximum mean square error.

For example, suppose the observations $\{Y(x_1),\ldots,Y(x_N)\}$ are uncorrelated, where $Y(x) = \theta(x) + \gamma$, $E(\gamma) = 0$, $E(\gamma^2) = 1$, and θ is essentially a polynomial of degree three in the sense that $\int_{-1}^{1.5} (\theta^{(4)}(t))^2 dt \le \varepsilon^2 = 1$. The optimal estimator of $\theta(1.5)$ has maximum mean square error over $\|\theta^{(4)}\|_2 \le 1$ of $(81.157)/20$ when $N = 20$. The procedure suggested is not optimal here, but to this degree of accuracy, it has the same maximum mean square error (mmse). The optimal estimator of $\theta^{(1)}(1.5)$ has mmse $(578.250)/20$ which compares with $(578.251)/20$ for the suggested procedure. In estimating $\theta^{(2)}(1.5)$ the optimal mmse is $(1308.811)/20$ while that of the suggested is $(1308.818)/20$. The suggested procedure has mmse $(601.951)/20$ for estimating $\theta^{(3)}(1.5)$ while the optimal is $(601.945)/20$.

These maximum mean square errors indicate the potential folly of extrapolation to a point as far as one fourth of the length of the observation interval. Only a slight improvement occurs when the model is known to be correct and a polynomial of degree 3, for the optimal variances are then $\frac{81}{20}$, $\frac{576}{20}$, $\frac{1296}{20}$, and $\frac{576}{20}$ respectively in estimating $\theta^{(0)}(1.5)$ through $\theta^{(3)}(1.5)$. Even for a sample size of 1000 the estimator of the second derivative has an optimal mmse of 1.921, the suggested an mmse of 1.936 and, if the third degree model is correct, an optimal variance of 1.296.

In the example, and generally, if c were closer to 1, or ε were smaller, or if σ^2 were larger then the suggested procedure would compare even more favorably with the optimal, both performing better on an absolute scale as $c \downarrow 1$.

Spruill [1984, 1985] has previously investigated in some depth the case $k=0$ and we proceed in our analysis here by the same techniques. Our conclusions are similar; when estimating the k^{th} derivative of θ the generalized Hoel-Levine procedure yields estimates which are locally model robust.

1.2. Optimal Minimax Extrapolation

Let $Y(x_1),\ldots,Y(x_N)$ be uncorrelated, $Y(x_i) = \theta(x_i) + \gamma_i$, $E(\gamma_i) = 0$, and $E(\gamma_i^2) = 1$ for $i = 1,\ldots,N$. If the function θ is in the Sobolev space $W_m^2[-1,c]$ and $\varepsilon > 0$ is arbitrary then there is a linear estimator $\ell_1'Y$ of $\theta^{(k)}(c)$, $k < m$ satisfying

$$\sup_{\|\theta^{(m)}\|_2 \le \varepsilon} E_\theta(\ell_1'Y-\theta^{(k)}(c))^2 \le \sup_{\|\theta^{(m)}\|_2 \le \varepsilon} E_\theta(\ell'Y-\theta^{(k)}(c))^2$$

for all constant N-vectors ℓ. Speckman [1979] developed these estimators and showed among other things that $\ell_1'Y$ has the same value as $\tilde{\theta}^{(k)}(c)$, where $\tilde{\theta} \in W_m^2[-1,c]$ minimizes the form

$$\sum_{i=1}^{N} (Y(x_i)-\theta(x_i))^2 + \frac{1}{\varepsilon^2}\int_{-1}^{c} (\theta^{(m)}(t))^2 dt \tag{1.1}$$

It follows from [Wahba, 1978] that Speckman's estimator is a limit of Bayes rules if γ_i are i.i.d. $N(0,1)$.

The mmse of Speckman's estimator is $N^{-1}d_{\eta,k}(\xi)$ employing the design ξ, where

$$d_{\eta,k}(\xi) = \sup\left(\frac{(\theta^{(k)}(c))^2}{\int_{-1}^{1} \theta^2(x)d\xi(x) + \eta \int_{-1}^{c} (\theta^{(m)}(x))^2 dx} \right),$$

$\eta = (N\varepsilon^2)^{-1}$, and the supremum is over those elements θ of $W_m^2[-1,c]$ for which the denominator does not vanish. The design ξ_0 is called optimal (in the approximate theory) if ξ_0 is a Borel measure and for all other such measures ξ, $d_{\eta,k}(\xi_0) \leq d_{\eta,k}(\xi)$.

Let $k \in \{0,\ldots,m-1\}$ be fixed, where $m \geq 1$ is also fixed. The Lagrange interpolation polynomials to the points $-1 \leq x_1 < x_2 < \ldots < x_m \leq 1$ are denoted by $\phi_{x_j}(x)$, $j = 1,\ldots,m$. Define for x and $t \in [-1,c]$

$$h_x(t) = \frac{(x-t)_+^{m-1}}{(m-1)!} - \sum_{i=1}^{m} \phi_{x_i}(x) \frac{(x_i-t)_+^{m-1}}{(m-1)!},$$

let

$$h_x^{(k)}(t) = \left.\frac{\partial^k h_y(t)}{\partial y^k}\right|_{y=x}, \quad R^2 = \|h_c^{(k)}\|_2^2 = \int_{-1}^{c} (h_c^{(k)}(t))^2 dt,$$

and

$$Q = \sum_{i=1}^{m} |\phi_{x_i}^{(k)}(c)|.$$

Let $\eta > 0$ be fixed and

$$\delta(x) = (\eta Q^2+R^2)^{-1}[\eta Q \sum_{j=1}^{m} (-1)^{j-m}\phi_{x_j}(x) + \int_{-1}^{c} h_c^{(k)}(t)h_x(t)dt].$$

Also introduce the extremal problems

$$P_{\eta,k}\text{: minimize } \|\theta\|_\infty^2 + \eta\|\theta^{(m)}\|_2^2 \text{ over all } \theta \in W_m^2[-1,c]$$

$$\text{such that } \theta^{(k)}(c) = 1.$$

Here and below

$$\|\theta\|_\infty^2 = \sup_{-1\le x\le 1} |\theta(x)|^2 \quad\text{and}\quad \|\theta^{(m)}\|_2^2 = \int_{-1}^{c} (\theta^{(m)}(t))^2 dt.$$

<u>Theorem 1.2.1.</u> There is an optimal design ξ_0 for estimating $\theta^{(k)}(c)$ whose support is $-1 \le x_1 < x_2 < \dots < x_m \le 1$ if and only if the corresponding δ equioscillates at the points $\{x_1,\dots,x_m\}$ in the sense that for $i = 1,\dots,m$ $\delta(x_i) = (-1)^{i-m}\|\delta\|_\infty$. If there is such an m point optimal design ξ_0 then

$$\xi_0(x_j) = \frac{|\phi_{x_j}^{(k)}(c)|}{\sum_{i=1}^{m} |\phi_{x_i}^{(k)}(c)|}, \quad j = 1,\dots,m,$$

ξ_0 is unique among m point optimal designs, and δ solves $P_{\eta,k}$.

In Figure 1.1 can be found plots of the functions δ associated with the optimal designs for $m = 3$, $c = 1.5$, and $\eta = .005$. Typically the solutions are found, as were the δ's there, numerically. When $m = 2$ and $k = 1$ the solutions can be found analytically and are as follows for estimating $\theta^{(1)}(c)$ from the interval $[a,b]$, $b \le c$. If $\eta \ge \frac{(b-a)^3}{24}$ then $x_1 = a$ and $x_2 = b$. Otherwise, $x_1 = b-2(3\eta)^{1/3}$ and $x_2 = b$. We have previously discussed in [Spruill, 1984] the estimation of $\theta(c)$. When $m = 1$ it is easy to show that all observations should be taken at b.

The proof of Theorem 1.2.1 can be carried out utilizing the program established in [Spruill, 1984] paralleling that of Theorem 4.1 therein and employing instead Theorem 1.2.2 below, whose proof is given in the appendix. Denote the support of ξ by $S(\xi)$.

<u>Theorem 1.2.2.</u> Suppose there is a function δ_0 in the set $\Delta = \{\theta \in W_m^2[-1,c]: \theta^{(k)}(c) = 1\}$ and a probability measure ξ_0 on the Borel subsets of $[-1,1]$ such that

i) $S(\xi_0) \subset \{x: |\delta_0(x)| = \max_{[-1,1]} |\delta_0(x)|\}$,

ii) there is an $\alpha > 0$ such that for all $\theta \in W_m^2[-1,c]$

$$\int_{-1}^{1} \theta(x)\delta_0(x)d\xi_0(x) + \eta \int_{-1}^{c} \theta^{(m)}(x)\delta_0^{(m)}(x)dx = \alpha\theta^{(k)}(c),$$

and

iii) $$\int_{-1}^{1} \theta^2(x)d\xi_0(x) + \eta \int_{-1}^{c} (\theta^{(m)}(x))^2dx = 0$$

entails $\theta^{(k)}(c) = 0$. Then ξ_0 is optimal for estimating $\theta^{(k)}(c)$ and

iv) δ_0 solves $P_{\eta,k}$.

Furthermore, if there is a Borel probability measure ξ_0 such that $d_{\eta,k}(\xi_0) = \inf d_{\eta,k}(\xi) < \infty$ then there is a solution δ_0 to $P_{\eta,k}$ satisfying i) - iv).

The course from here to a proof of Theorem 1.2.1 following [Spruill, 1984] is clear except for the following fact which corresponds to Lemma 3.1 there and leads to the proof of Theorem 1.2.3 below. Let $f_j(x) = x^j$, $j = 0,\dots,m-1$ and $a < b \leq c$.

<u>Lemma 1.2.1</u>. If $0 < k \leq m-1$ and

$$g_j(x) = f_j(x) - \frac{f_j^{(k)}(c)}{f_k^{(k)}(c)} f_k(x)$$

$j \neq k$ then $\{g_j\}_{j=0,j\neq k}^{m-1}$, except possibly for the sign of one of them, is a T system on $[a,b]$.

<u>Proof</u>: Let $a \leq \tau_0 < \tau_1 < \ldots < \tau_{k-1} < \tau_{k+1} < \ldots < \tau_{m-1} \leq b$ be given and form the determinant

$$D = \begin{vmatrix} g_0(\tau_0) & g_0(\tau_1) & \cdots & g_0(\tau_{m-1}) \\ \vdots & \vdots & \vdots & \vdots \\ g_{m-1}(\tau_0) & \cdots & \cdots & g_{m-1}(\tau_{m-1}) \end{vmatrix}_{m-1\times m-1} = \frac{1}{f_k^{(k)}(c)} \psi^{(k)}(c),$$

where

$$\psi(x) = \begin{vmatrix} f_k(x) & f_k(\tau_0) & \cdots & f_k(\tau_{m-1}) \\ f_0(x) & f_0(\tau_0) & \cdots & f_0(\tau_{m-1}) \\ \vdots & \vdots & \vdots & \vdots \\ f_{m-1}(x) & f_{m-1}(\tau_0) & \cdots & f_{m-1}(\tau_{m-1}) \end{vmatrix}_{m\times m}.$$

Suppose we have chosen $\tau_0,\ldots,\tau_{m-1}$ to make D vanish. Then $\psi^{(k)}(c) = 0$. Since ψ has m-1 zeros in $[a,b]$ it follows that $\psi^{(k)}$ has m-1-k in (a,b) and hence at least m-k in $[a,c]$. Since ψ is of degree m-k-1, $\psi \equiv 0$ on $[a,c]$. This contradicts the fact that $\{x^0,x^1,\ldots,x^{m-1}\}$ is a T system on $[a,c]$ and proves the assertion. □

Now one can prove the following.

Theorem 1.2.3. The problem $P_{\eta,k}$ has a solution $\theta_0 \in W_m^2[-1,c]$. If $\eta > 0$ it is unique and there are points $-1 \leq x_1 < \ldots < x_{m+r} \leq 1$ and a $q = 0$ or 1 such that

$$\theta_0(x_i) = (-1)^{i+q} ||\theta_0||_\infty$$

$i = 1,\ldots,m+r$.

For c fixed, as $\eta \to \infty$ the optimal designs are always m point designs. Furthermore, the design points and masses approach those of the generalized Hoel-Levine designs. The proof of this statement can be obtained as in [Spruill, 1985] utilizing the fact, which follows for example from [Rivlin, 1969] Theorem 1.10, that the unique polynomial p of degree m-1 minimizing $||p||_\infty$ subject to $p^{(k)}(c) = 1$ is $T_{m-1}/T_{m-1}^{(k)}(c)$, where T_{m-1} is the m-1st Chebyshev polynomial of the first kind.

For η fixed, as $c \downarrow 1$ there are asymptotic optimal designs. Let us first consider the case $k = 0$. We require the following lemmas.

Lemma 1.2.2. Let f and g be continuous functions on an interval I whose right endpoint is $b \in I$. If for points $x_1 < x_2 < \ldots < x_m$ in I, $f(x_j) = (-1)^{j-m}$, and for points $y_1 < y_2 < \ldots < y_m$ in I, $g(y_j) = (-1)^{j-m}$, and if $x_m = y_m = b$, then $\sup_{x\in I} |f(x)| = \sup_{x\in I} |g(x)| = 1$ entails at least m zeros of f-g in I.

The proof is left to the reader. From Lemma 1 of [Spruill, 1985] we know that if

$$k(c) = \inf_{\theta \in W_m^2[-1,c]} \frac{(\|\theta\|_\infty^2 + \eta\|\theta^{(m)}\|_2^2)}{\|\theta\|_{[-1,c]}^2}$$

then $k(c) > 0$ for every $c > 1$ where

$$\|\theta\|_{[-1,c]}^2 = \sum_{i=0}^{m} \int_{-1}^{c} (\theta^{(i)}(t))^2 dt.$$

<u>Lemma 1.2.3.</u> $\lim\limits_{c \downarrow 1} k(c) > 0$.

<u>Proof:</u> First transform to the interval [-1,1] linearly so that 1 maps to $B = (3-c)/c+1$. Consider

$$K(B) = \inf_{f \in W_m^2[-1,1]} \frac{\left(\sup\limits_{-1 \le x \le B} |f(x)|^2 + \eta \int_{-1}^{1} (f^{(m)}(x))^2 dx \right)}{\|f\|_{[-1,1]}^2}$$

The linear transformation is $z = ux + v$ from $[-1,c]$ to $[-1,1]$, $u = 2(c+1)^{-1}$ and then

$$u^{2m-1} k(c) \le K(B) \le \frac{1}{u^{2m}} k(c).$$

Now if $B_2 > B_1$ then $K(B_2) \ge K(B_1)$ so if $u < 1$

$$0 < \frac{2^{2m-1} k(c_1)}{(c_1+1)^{2m-1}} \le K(B_1) \le \lim_{B \to 1} K(B) \le \lim_{c \downarrow 1} k(c) \frac{(c_1+1)^{2m}}{2^{2m}}$$

and $\lim\limits_{c \downarrow 1} k(c) > 0$. □

Fix η, let $c_n \downarrow 1$ as $n \to \infty$, ξ_n be the optimal design for Speckman's estimator for extrapolation from [-1,1] to c_n, θ_n minimize $\rho_n(\theta) = \|\theta\|_{\infty[-1,1]}^2 + \eta \|\theta^{(m)}\|_{2[-1,c_n]}^2$

over $\theta \in W_m^2[-1,c_n]$ such that $\theta(c_n) = 1$, and $\{\bar{x}_j\}_{j=1}^m$ be the Chebyshev points in $[-1,1]$.

Theorem 1.2.4. For n sufficiently large

i) $S(\xi_n) = \{x_{nj}\}_{j=1}^m$, where $-1 \le x_{n1} < \ldots < x_{nm} \le 1$ and

ii) $\lim_{n\to\infty} x_{nj} = \bar{x}_j$, $j = 1,\ldots,m$.

Proof: We first claim that $\theta_n(1) \to 1$ as $n \to \infty$. Since

$$\|\theta_n\|_\infty^2 \le \rho_n(\theta_n) \le \rho_n\left(\frac{T_{m-1}}{T_{m-1}(c_n)}\right) = \frac{\|T_{m-1}\|_\infty^2}{T_{m-1}^2(c_n)}$$

$\overline{\lim_{n\to\infty}}\ \theta_n(1) \le 1$. If $\underline{\lim_{n\to\infty}}\ \theta_n(1) < 1$ then for some $\varepsilon > 0$ and subsequence n_ν we have

$$\varepsilon \le \theta_{n_\nu}(c_{n_\nu}) - \theta_{n_\nu}(1) = \int_1^{c_{n_\nu}} \theta'_{n_\nu}(t)dt \le \sqrt{c_{n_\nu}-1}\ \left(\int_1^{c_{n_\nu}} (\theta'_{n_\nu}(t))^2 dt\right)^{1/2}$$

which implies $\|\theta_{n_\nu}\|_{[-1,c_n]} \to \infty$. By Lemma this implies $\rho_{n_\nu}(\theta_{n_\nu}) \to \infty$. But

$$\rho_n(\theta_n) \le \frac{T_{m-1}^2(1)}{T_{m-1}^2(c_n)} \to 1 \quad \text{so} \quad \underline{\lim_{n\to\infty}}\ \theta_n(1) \ge 1.$$

It follows that

$$\int_{-1}^{1} (\theta_n^{(m)}(t))^2 dt \to 0$$

as $n \to \infty$ and therefore that $\|\theta_{n'} - p\|_{\infty[-1,1]} \to 0$ for some polynomial p of degree m-1 and some subsequence n'. Now, using the equioscillation of $\theta_{n'}$ and arguing as in Theorem 1 of [Spruill, 1985] we conclude that p satisfies $p(y_i) = (-1)^{i-m}$ for m points $-1 \leq y_1 < y_2 < \ldots < y_m \leq 1$ and $\|p\|_{\infty[-1,1]} = 1$. It now follows from Lemma 1.2.2 that $p = T_{m-1}$ and the assertions are an immediate consequence. □

We know that as $c \downarrow 1$ the optimal designs for estimating $\theta(c)$ converge to the Hoel-Levine design. This is no longer true when estimating $\theta^{(k)}(c)$ for $k > 0$; there are asymptotic designs but they depend upon η. The function solving

$$\min_{W_m^2[-1,1]} \left(\|\theta\|^2_{\infty[-1,1]} + \eta \|\theta^{(m)}\|^2_{2[-1,1]} \right)$$

subject to $\theta^{(k)}(1) = 1$ equioscillates, for η sufficiently large, at precisely m points and these are the points of support of the asymptotic design with weights determined as usual. However, this equioscillation does not take place at the Chebyshev points, as one can verify by considering the function δ. One case, estimating $\theta^{(m-1)}(c)$, has the property that the designs do not depend upon c, so that these asymptotic designs are actually appropriate for every $c \geq 1$.

1.3. Comparison with the Polynomial Case

Theorem 1.2.2 remains valid for $\eta = +\infty$ interpreting $(+\infty)(0) = 0$ so if $\eta = +\infty$ then δ_0 is a polynomial of degree m-1 which minimizes $\|\theta\|_\infty$ subject to $\theta^{(k)}(c) = 1$. As is well known the unique solution is $T_{m-1}(x)/T^{(k)}_{m-1}(c)$. Consequently, the optimal design points are the Chebyshev points when Speckman's estimator is used. From (1.1) we see that this is just the usual least squares estimator and hence we have the following theorem. Let $\{Y(x_1),...,Y(x_N)\}$ satisfy $Y(x_i) = \theta(x_i) + \varepsilon_i$, $E(\varepsilon_i) = 0$, $E(\varepsilon_i\varepsilon_j) = \sigma^2\delta_{ij}$, and $\theta \in \mathcal{P}_{m-1}$.

<u>Theorem 1.3.1</u>. For the minimum variance linear unbiased estimation of $\theta^{(k)}(c)$ the optimal design is supported on

$$x_i = -\cos\left(\frac{(i-1)\pi}{m-1}\right) \quad i = 1,...,m$$

and assigns masses

$$\xi_i = |\phi_{x_i}^{(k)}(c)| / \sum_{j=1}^{m} |\phi_{x_j}^{(k)}(c)|$$

at these points.

These are the designs and estimation procedures described in the opening paragraph which we have called generalized Hoel-Levine designs. Since for a design ξ on m points for which $n_i \propto |\phi_{x_i}^{(k)}(c)|$, we have

$$\text{mmse} = N^{-1} d_{\eta,k}(\xi) = N^{-1}\left[\left(\sum_{i=1}^{m} |\phi_{x_i}^{(k)}(c)|\right)^2 + \frac{1}{\eta}\,\|h_c^{(k)}\|_2^2\right]$$

it follows that the mmse is continuous in the design points and hence that in the asymptotics above $(\eta \to \infty)$ we also have, besides convergence of the design points, convergence of the maximum mean square errors. Therefore, because the estimators are both least squares for η large, the generalized Hoel-Levine procedures are locally, as measured by $\eta = \sigma^2/N\varepsilon^2$, robust against departures from the model.

The tables exhibit the supports of optimal designs. Entries labeled ommse are N times the maximal mean square error of Speckman's estimator for the optimal choice of points. The entries gmmse are N times the maximal mean square errors of the generalized Hoel-Levine procedure described in the introduction. Since for estimating $\theta^{(m-1)}(c)$ the designs are independent of c no value of c is indicated for those entries.

The tabled values support the observation that as $\eta \to \infty$ the optimal designs for Speckman's estimator converge to the generalized Hoel-Levine designs and that their maximum mean square errors converge to the variance of the associated least squares estimator for the polynomial model. They also suggest that the convergence is rapid, for though the theorems are stated to hold as $\eta \to \infty$, for apparently small values of η the asymptotic behavior is seen to be very nearly realized. Actually, these values which appear small are not surprising. Let σ^2 and N be given and suppose the observations are used to test H_0: $\theta \in P_{m-1}$ against H_A: $\theta \in P_m$ instead of in the estimation of $\theta(c)$, where P_r is the collection of polynomials on [-1,1] of degree r. Then for the test which rejects H_0 when $|\hat{\beta}_m|$ is large, $p(x) = \sum_{j=0}^{m} \beta_j x^j$, and is of size α the optimal assignment of observations is at the m+1 Chebyshev points $x_i = -\cos(\frac{i\pi}{m})$, $i = 0,1,\ldots,m$, with the usual weights. The power against $p(x) = \beta x^m$, $x \in [-1,1]$ in terms of η is

$$P(\eta) = 1 - [\Phi(z_{\alpha/2} - \frac{1}{\sqrt{2\eta}\ m!E_m}) - \Phi(-z_{\alpha/2} - \frac{1}{\sqrt{2\eta}\ m!E_m})]$$

where

$$E_m = \sum_{j=0}^{m} (\prod_{i\neq j} |x_j - x_i|)^{-1}.$$

The values of E_m are 4, 24, and 192 respectively for m = 2, 3, and 4. Consulting [Spruill, 1985] we find, for example, that when m = 3, c = 1.5 and k = 0 the Hoel-Levine and optimal procedures perform much the same for $\eta \geq .01$. Since $P(.01) \doteq .0503$ when $\alpha = .05$ it should not be surprising that this is so, for whenever θ is a polynomial of degree m satisfying $\int_{-1}^{1} (\theta^{(m)}(t))^2 dt \leq \varepsilon^2$, $\eta = \sigma^2/N\varepsilon^2$, and $\eta \geq .01$ the power is no more than .0503. Thus the data cannot distinguish such a θ from a polynomial of degree m-1 and we see that in fact .01 is a "large" value of η. When m = 4, c = 1.5, and k = 0 the η value is 5×10^{-4} and $P(5 \times 10^{-4}) \doteq .05$ so again this is a "large" value. In all cases the η values at which asymptotic behavior is essentially obtained for derivatives are larger than those for k = 0 so similar statements apply. The same sort of arguments do not apply to explain the transition from m to m+1 points as η decreases.

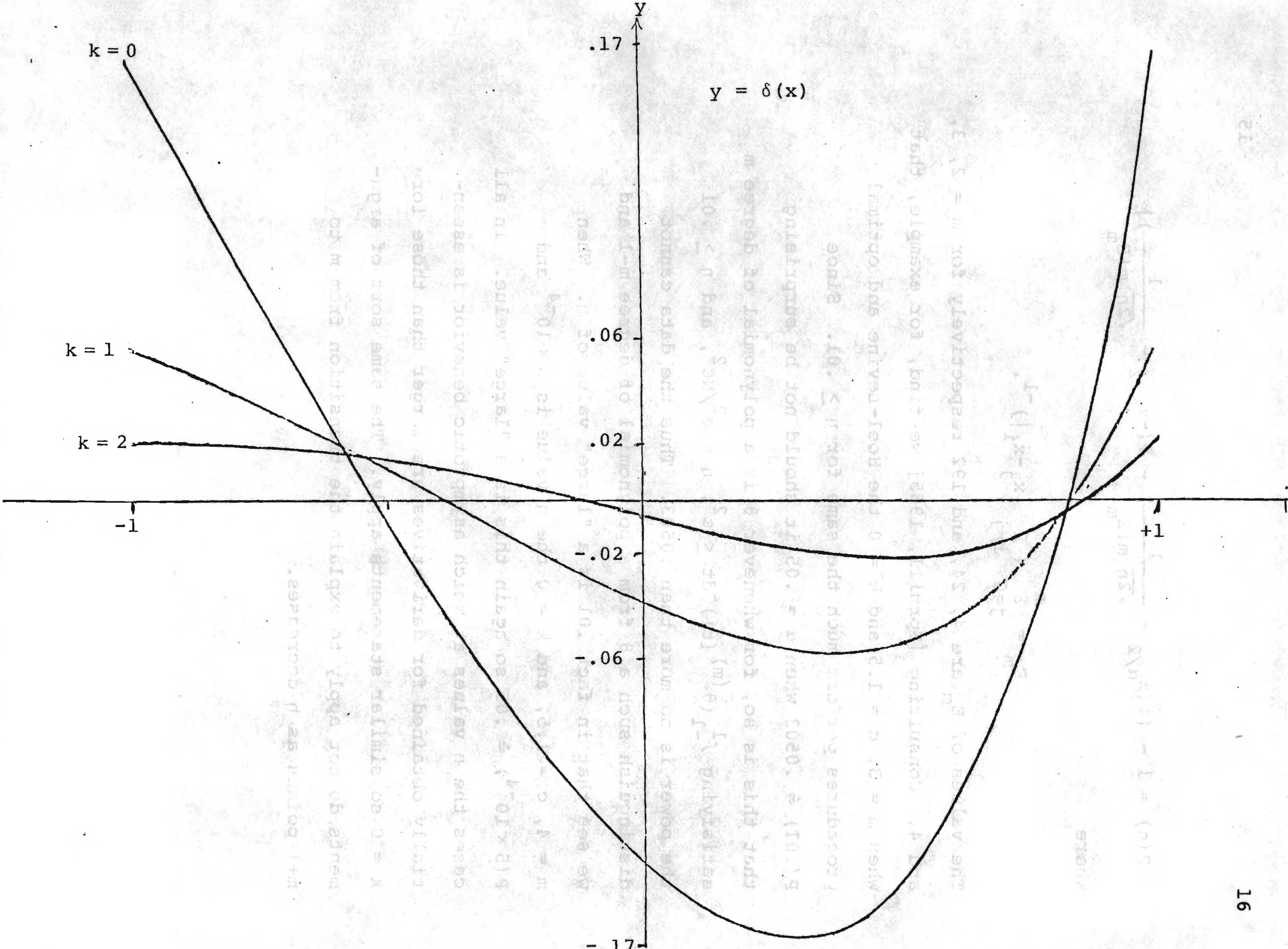

Figure 1.1. Solutions to $P_{\eta,k}$ for $\eta = .005$, $m = 3$, $c = 1.5$, $k = 0$, 1 and 2.

TABLE I.

Optimal Locations for m = 3, k = 1.

η	c	x_1	x_2	x_3	ommse	gmmse
.5	1.25	-1	.005821	+1	25.496	25.497
.1	-	-	.028793	-	27.447	27.489
.05	-	-	.056575	-	29.813	29.979
.01	-	-	.226283	-	46.359	49.897
.005	-	-	.342289	-	63.291	74.795
.001	-	-.669655	.555829	-	155.194	273.979
.5	1.5	-1	.006588	+1	37.080	37.083
.1	-	-	.032653	-	41.339	41.417
.05	-	-	.064243	-	46.527	46.834
.01	-	-	.253256	-	83.606	90.171
.005	-	-	.375663	-	123.132	144.342
.001	-	-.593443	.577075	-	357.122	577.714

TABLE II.

Optimal Locations for m = 3, k = 2.

	x_1	x_2	x_3	ommse	gmmse
.5	-1	.010436	+1	18.530	18.533
.1	-	.052204	-	28.580	28.667
.05	-	.102979	-	40.987	41.335
.01	-	.364450	-	135.364	142.679
.005	-.948596	.484811	-	246.705	269.358
.001	-.412302	.626601	-	1031.659	1282.792

TABLE III.

Optimal Locations for m = 4, k = 1.

η	c	x_1	x_2	x_3	x_4	ommse	gmmse
.5	1.25	-1	-.499962	.500048	+1	248.119	248.119
.1	-	-	-.499811	.500244	-	248.347	248.347
.05	-	-	-.499623	.500489	-	248.632	248.632
.01	-	-	-.498116	.502429	-	250.902	250.912
.005	-	-	-.496229	.504820	-	253.723	253.761
.001	-	-	-.481073	.522627	-	275.645	276.558
5×10^{-4}	-	-	-.462065	.541891	-	301.584	305.055
10^{-4}	-	-	-.320988	.627155	-	470.874	533.026
5×10^{-5}	-	-1	-.193678	.669966	-	631.742	817.990
10^{-5}	-	-.730681	.086434	.746349	-	1410.121	3097.702
.5	1.5	-1	-.499954	.500060	+1	576.225	576.225
.1	-	-	-.499770	.500303	-	577.125	577.125
.05	-	-	-.499540	.500606	-	578.250	578.251
.01	-	-	-.497698	.503012	-	587.225	587.259
.005	-	-	-.495387	.505973	-	598.383	598.519
.001	-	-	-.476652	.527844	-	685.339	688.596
5×10^{-4}	-	-	-.452823	.551033	-	788.817	801.193
10^{-4}	-	-	-.277606	.645202	-	1481.064	1701.965
5×10^{-5}	-	-	-.135323	.687775	-	2168.709	2827.931
10^{-5}	-	-.658454	.126388	.758829	-	5843.933	11835.658

TABLE IV.

Optimal Locations for m = 4, k = 2.

η	c	x_1	x_2	x_3	x_4	ommse	gmmse
.5	1.25	-1	-.499941	.500078	+1	900.632	900.632
.1	-	-	-.499704	.500394	-	903.163	903.164
.05	-	-	-.499409	.500787	-	906.324	906.328
.01	-	-	-.497039	.503908	-	931.552	931.642
.005	-	-	-.494057	.507738	-	962.926	963.285
.001	-	-	-.469545	.535624	-	1207.791	1216.428
5×10^{-4}	-	-	-.437834	.564126	-	1499.973	1532.856
10^{-4}	-	-	-.219296	.664968	-	3477.730	4064.281
5×10^{-5}	-	-	-.070944	.705025	-	5492.423	7228.563
.5	1.5	-1	-.499935	.500086	+1	1297.281	1297.281
.1	-	-	-.499679	.500434	-	1302.407	1302.409
.05	-	-	-.499358	.500867	-	1308.811	1308.818
.01	-	-	-.496778	.504304	-	1359.934	1360.090
.005	-	-	-.493528	.508520	-	1423.561	1424.180
.001	-	-	-.466637	.539069	-	1921.992	1936.904
5×10^{-4}	-	-	-.431550	.569877	-	2520.982	2577.808
10^{-4}	-	-	-.194599	.673714	-	6692.122	7705.043
5×10^{-5}	-	-.987601	-.045008	.712990	-	11124.502	14114.086

TABLE V

Optimal Locations for m = 4, k = 3.

η	x_1	x_2	x_3	x_4	ommse	gmmse
.5	-1	-.499910	.500127	+1	578.595	578.595
.1	-	-.499550	.500634	-	588.974	588.975
.05	-	-.499100	.501268	-	601.945	601.951
.01	-	-.495467	.506282	-	705.616	705.755
.005	-	-.490854	.512411	-	834.956	835.511
.001	-	-.451468	.565649	-	1860.073	1873.557
5×10^{-4}	-	-.398388	.595989	-	3119.381	3171.115
10^{-4}	-	-.092491	.704666	-	12653.569	13551.576

1.4. Key Theorem

Theorem 1.2.2 follows, as a special case, from Theorem 1.4.2 below. We assume that for each distinct set of points $\{x_1,\ldots,x_k\}$, $k < \infty$, in some factor space X and for each set of positive integers $\{n_1,\ldots,n_k\}$ an experiment can be run, resulting in the uncorrelated observations $\{Y_1(x_1), Y_2(x_1),\ldots, Y_{n_k}(x_k)\}$. Here $Y(x) = (m_x,\theta) + \varepsilon$, where $E(\varepsilon) = 0$, $E(\varepsilon^2) = \sigma^2$, θ is an unknown element of the Hilbert space Θ, $m_x \in \Theta$ for each $x \in X$, and $(\cdot,\cdot)$ is the inner product on Θ. There is an auxiliary separable Hilbert space H and a bounded linear operator T from Θ into H. We seek the linear estimator $\ell'Y$ of (τ,θ), where τ is some fixed known element of Θ, for which $\sup_{\|T\theta\|_H \leq \varepsilon} E_\theta(\ell'Y-(\tau,\theta))^2 = V(\ell,\tau)$ is minimized. Here $\varepsilon > 0$ is also fixed. The extrapolation problem above can be put into this framework with $\Theta = W_m^2[-1,c]$, $H = L_2[-1,c]$, $(\tau,\theta) = \theta^{(k)}(c)$, and $X = [-1,1]$. In the general case Speckman has found the estimator $\ell_0'Y$ and its maximum mean square error $V(\ell_0,\tau) = d(\tau)$. Of course this quantity depends upon the points x_i and numbers n_i and we emphasize this in the notation by writing $d(\tau,\xi)$. One can show that $d(\tau,\xi) = N^{-1}(\tau, M^{\#}(\xi)\tau)$ where $M^{\#}(\xi)$ is the Moore-Penrose inverse of the operator $M(\xi)$ which we now describe.

Let $V = R\times H$ and introduce the inner product $(,)_V$ on V by $(v,w)_V = v_1w_1 + (v_2,w_2)_H$. For each Borel probability measure ξ on $[-1,1]$ let $L_2(\xi)$ denote the class of V-valued measurable functions on $S(\xi)$ with inner product

$$(f,g)_\xi = \int [f_1(x)g_1(x) + (f_2(x),g_2(x))_H]d\xi(x).$$

Let $L_\xi: \Theta \to L^2(\xi)$ be defined by $L_\xi\theta(x) = [(m_x,\theta),T\theta]$ for $x \in S(\xi)$. Then $M(\xi) = L_\xi^* L_\xi$.

Introduce the set

$$R = \{\int L_x^* \phi(x)d\xi(x) : \xi \in \Xi, \ \phi \in F\}$$

where Ξ is the Borel probability measures on $[-1,1]$, and F is the collection of measurable functions from X into V with $\|\phi(x)\|_V \leq 1$. One can use the arguments of [Spruill, 1980], replacing m_x there by L_x, and prove the following theorem. Certain assumptions are needed.

(A1) For every $\theta \in \Theta$ and $\xi \in \Xi$ the functions (m_x,θ) are measurable on $S(\xi)$.

(A2) For all $\xi \in \Xi$, $\int_{S(\xi)} \|m_x\|^2 d\xi(x) < \infty$.

(A3) For each $\xi \in \Xi$, L_ξ is bounded and the range of L_ξ is closed in $L^2(\xi)$.

(A4) There is a proper closed supporting hyperplane to R at each of its boundary points.

(A5) For each $\theta \neq 0$ $\sup_X \|L_x\theta\| > 0$.

Let $v_0 = \inf_{\Xi} d(\tau,\xi)$.

<u>Theorem 1.4.1</u>. Under assumptions (A1)-(A5) if $\tau \in R(M(\xi))$ for some $\xi \in \Xi$ then $d(\tau,\xi_0) = v_0$ and $\xi_0 \in \Xi$ if and only if there is a function $\phi \in F$ such that $\|\phi(x)\|_V \equiv 1$ and $\int L_x^*\phi(x)d\xi_0(x)$ is i) proportional to τ and ii) in $R \cap \partial R$.

One can also show that

$$(\tau, M^{\#}(\xi)\tau) = \sup_{\theta \in N} \frac{(\tau,\theta)^2}{\int ||L_x\theta||_V^2 \, d\xi(x)}$$

where $N = N(\xi) = \{\theta: \int ||L_x\theta||^2 d\xi(x) > 0\}$. Let $\Delta = \{\theta \in \Theta: (\tau,\theta) = 1\}$.

Theorem 1.4.2. Suppose there is a point $\delta_0 \in \Delta$ and a design $\xi_0 \in \Xi$ satisfying

i) $S(\xi_0) \subset \{x: ||L_x\delta_0||_V = \sup_X ||L_x\xi_0||_V\}$,

ii) $\int L_x^* L_x \delta_0 d\xi_0(x) = \alpha\tau$ for some $\alpha > 0$, and

iii) $\int ||L_x\theta||_V^2 d\xi_0(x) = 0$ entails $(\tau,\theta) = 0$.

Then ξ_0 satisfies $d(\tau,\xi_0) = \inf d(\tau,\xi)$ and

iv) $\inf_\Delta \sup_X ||L_x\delta||_V^2 = \sup_X ||L_x\delta_0||_V^2$.

The conditions required are (A1) and (A2). Conversely if (A1) - (A5) hold and there is a $\xi_0 \in \Xi$ satisfying $d(\tau,\xi_0) = v_0 < \infty$ then a point $\delta_0 \in \Delta$ may be found satisfying conditions i) through iv).

Proof. Assume i) through iii), (A1), and (A2). We have for $\xi \in \Xi$

$$d(\tau,\xi) \geq [\inf_{\theta \in N \cap \Delta} \int ||L_x\theta||_V^2 d\xi(x)]^{-1},$$

and since

$$\inf_{N\cap\Delta} \int ||L_x\theta||^2 d\xi \leq \inf_{N\cap\Delta} \sup_X ||L_x\theta||^2 \leq \sup_X ||L_x\delta_0||^2 = s,$$

we have $d(\tau,\xi) \geq s^{-1}$. Now, using ii)

$$d(\tau,\xi_0) = \sup_N \frac{\alpha^{-2}[\int (L_x\theta, L_x\delta_0)_v d\xi_0(x)]^2}{\int \|L_x\theta\|_v^2 d\xi_0(x)}$$

$$\leq \sup_N S \frac{\alpha^{-2}(\int \|L_x\theta\| d\xi_0(x))^2}{\int \|L_x\theta\|^2 d\xi_0(x)} \leq s\alpha^{-2} = s^{-1} .$$

Since $s^{-1} = d(\tau,\xi_0) \geq [\inf_{N\cap\Delta} \sup_X \|L_x\theta\|^2]^{-1}$ we have for $\theta \in N \cap \Delta$,

$$\sup_X \|L_x\theta\|^2 \geq \inf_{N\cap\Delta} \sup_X \|L_x\theta\|^2 \geq \sup_X \|L_x\delta_0\|^2 .$$

By iii) $\Delta = N \cap \Delta$ so we conclude that $d(\tau,\xi_0) = \inf_{\Xi} d(\tau,\xi)$ and iv) holds.

Now assume A1) - A5) and that $d(\tau,\xi_0) = \inf_{\Xi} d(\tau,\xi) < \infty$. By Theorem 1.4.1. there is a function $\phi: X \to \mathbb{R} \times H$ such that $\|\phi(x)\| \equiv 1$,

$$\int L_x^*\phi(x) d\xi_0(x) = \beta\tau,$$

and $\beta\tau \in \partial R$. By (A4) there is a $\lambda \neq 0$, $\lambda \in \Theta$, such that $(\lambda,r) \leq \beta(\lambda,r)$ for all $r \in R$. Since by (A5) $\sup_X \|L_x\lambda\| > 0$ we can find a sequence of points $\{x_n\}$ in X satisfying $\|L_{x_n}\lambda\| \uparrow \sup_X \|L_x\lambda\|$ and $\|L_{x_n}\lambda\| > 0$. Set

$$r_n = L_{x_n}^* \frac{L_{x_n}\lambda}{\|L_{x_n}\lambda\|} .$$

Then $r_n \in R$ for all n and since $(\lambda,r) \leq \beta(\lambda,r)$ for all r,

$$\lim_{n\to\infty} \|L_{x_n}\lambda\| \leq \int (L_x^*\phi(x),\lambda) d\xi_0(x) \leq \sup_X \|L_x\lambda\|$$

with strict inequality unless $\|L_x\lambda\| \equiv \sup_X \|L_x\lambda\|$ a.e. ξ_0.

Set $\delta_0 = \frac{\lambda}{(\tau,\lambda)}$ $((\tau,\lambda) \neq 0$ since $\beta(\tau,\lambda) > 0)$. Clearly i) is satisfied. From above we also conclude that $\phi(x) = k_x L_x \lambda$ a.e. ξ_0. This in turn implies that $\phi(x) = \frac{L_x\lambda}{||L_x\lambda||}$ a.e. ξ_0. Therefore

$$\int L_x^* \phi(x) d\xi_0(x) = [\int L_x^* L_x \delta_0 d\xi_0(x)][\sup_X ||L_x \lambda||]^{-1}$$

and we see that ii) is also satisfied. If iii) is not satisfied then there is a sequence θ_n such that $\int ||L_x\theta_n||^2 d\xi_0(x) \to 0$ and $(\tau,\theta_n) \to t \neq 0$. This implies $d(\tau,\xi_0) = +\infty$ which contradicts our assumptions. We conclude that iii) is satisfied and consequently that iv) also is satisfied. □

One can find the appropriate arguments to prove the next theorem in [Spruill, 1980].

<u>Theorem 1.4.3</u>. If assumptions (A1) and (A2) hold and if there is a constant $p > 0$ such that for all θ

$$\sup_X ||L_x\theta||_V \geq p\,||\theta|| \tag{1.2}$$

then (A4) and (A5) hold.

Spruill [1985a] proves that for the extrapolation problems with $X = [a,b]$, $\Theta = W_m^2[a,c]$, the inequality is satisfied.

Optimal Designs for Interpolation

2.1. Introduction

Much is known about the extrapolation of an unknown function from an interval when the observed values are subject to random errors. If the function is a polynomial of known degree then the variance of the estimated value is minimized by measuring the function's value at certain points determined by a Chebyshev polynomial. See Hoel and Levine (1964) and Karlin and Studden (1966). Even when the unknown function is not a polynomial, but close, this is a good choice of points and similar statements apply to the extrapolation of derivatives. See Spruill (1984, 1985a, 1985b).

The topic of concern below is interpolation. Roughly speaking, the optimal designs for polynomial interpolation are obtained like those for extrapolation. There is a collection of polynomials which play a role similar to the Chebyshev polynomials in the extrapolation problems, determining the supports of the optimal designs at their points of proper oscillation. The optimal masses are determined by these points. Again, even when the function is not a polynomial, but close, these designs perform well.

2.2. Optimal Designs for Polynomials

The value $\theta(c)$ of an unknown polynomial θ of degree m-1 is to be estimated based upon uncorrelated observations

$\{Y(x_i)\}_{i=1}^N$ taken with the restriction that the x_i's lie in the set $X = [A,B] \cup [D,E]$, where $B < c < D$, $Y(x_i) = \theta(x_i) + \gamma_i$ for $i = 1,\ldots,N$ and the γ_i are zero mean errors of constant variance σ^2. The variance of the least squares estimator of $\theta(c)$ depends upon the selection of points x_i in X and it is their optimal selection which is the subject of this paper. All but a few proofs are omitted below since the methods above apply.

Let the $m\text{-}1^{st}$ degree polynomial p_m^c solve the problem

$$P_{m-1}(c,X)\text{: minimize } \|p\|_{\infty,X} \text{ over all polynomials of degree } m-1 \text{ such that } p(c) = 1,$$

where $\|p_m\|_{\infty,X} = \max\limits_{x\in X}|p_m(x)|$. In Chapter 3 below it is shown that the solution p_n^c to $P_{m-1}(c,X)$ is unique and that there are at least m points

$$A \le x_1 < x_2 < \ldots < x_L \le B, \quad D \le x_{L+1} < \ldots < x_m \le E,$$

and numbers q_R and q_S in $\{0,1\}$ for which

$$p_m^c(x_i) = \begin{cases} (-1)^{q_R-i}\|p_m^c\|_{\infty,X} & i = 1,\ldots,L \\ (-1)^{q_S-i}\|p_m^c\|_{\infty,X} & i = L+1,\ldots,m. \end{cases}$$

We say that p_m^c oscillates properly at the points $x_1,\ldots,x_m$. It is also shown that these points do not depend upon the point c and sufficient information is there discovered to enable the rapid identification of the solutions p_m^c by a computer. In the case of symmetric intervals, B-A = E-D, formulas are given for these polynomials involving the trigonometric functions. The importance of these polynomials is that the optimal design's support is contained in the set of proper oscillation points.

Theorem 2.2.1. The optimal design ξ_c for estimating $\theta(c)$ using the least squares estimator is supported at the proper oscillation points of p_m^c, a set of either m or m+1 points, independent of c, and containing B, D and at least one of A and E. If there are m points $x_1 < x_2 < \ldots < x_m$ then the optimal proportions are

$$\xi_c(x_i) = |\phi_{x_i}(c)| / \sum_{j=1}^{m} |\phi_{x_j}(c)|,$$

where $\{\phi_{x_j}\}_{j=1}^{m}$ are the Lagrange interpolation polynomials of degree m-1 to the points $\{x_1,\ldots,x_m\}$. If there are m+1 points then every optimal design is a convex combination of the two m point designs formed as above from $\{x_1,\ldots,x_m\}$ and $\{x_2,\ldots,x_{m+1}\}$.

Methods of proof previously employed yield the m point case without difficulty. To see the m+1 point case utilize the fact that for all polynomials $\theta \in P_{m-1}$, $\int p_m^c(x)\theta(x)d\xi_c(x) \stackrel{\theta}{\equiv} \gamma\theta(c)$ for some $\gamma > 0$. This implies that for constants c_{uv} and γ, depending only on the points $x_1,\ldots,x_{m+1}$, we have $c_{jk}\xi_c(x_j) + c_{kj}\xi_c(x_k) \equiv \gamma$ so that $\xi_c(x_{m+1})$ determines the remaining weights. Now it is easily seen that the largest mass assigned by an optimal design to x_{m+1} is assigned by the one which places zero mass at x_1. This verifies the claim.

For m = 1 or 2 one can see that $p_m(x) \equiv 1$. All designs are optimal for m = 1 while for m = 2 any design ξ for which $\int x d\xi(x) = c$ is optimal. When m = 3 or 4 the optimal masses are determined from the supports in the usual way. For m = 4, the support is {A,B,D,E}. When m = 3, if B-A > E-D then the support is {A,B,D}, while if E-D > B-A it is {B,D,E}. Some illustrative optimal designs, their variances, and plots of p_m^c can be found in Figures 2.1 through 2.3.

When B-A = E-D and m > 2 explicit formulas for the optimal designs are available below. They are a consequence of the fact found below that for m even, m ≥ 2, A = -1 = -E, B = -D,

$$p_m^0(x) = M_m S_{m/2}(x),$$

where for $x \in X = [-1,-D] \cup [D,1]$

$$S_k(x) = \cos((2k-2)\tan^{-1}((\frac{x^2 - D^2}{1 - x^2})^{1/2})).$$

One can also find there the facts that

$$M_m = ||p_m^0||_{\infty,X} = 2[(\frac{1+D}{1-D})^{\frac{m}{2}-1} + (\frac{1-D}{1+D})^{\frac{m}{2}-1}]^{-1},$$

and for all x, $S_k(x)$ satisfy the difference equation

$$S_{k+1}(x) + S_{k-1}(x) = 2S_k(x)S_2(x), \quad k = 2,3,..., \tag{2.1}$$

where $S_1(x) \equiv 1$ and

$$S_2(x) = \frac{2x^2}{D^2-1} + \frac{1+D^2}{1-D^2}.$$

A simple transformation can be made from [A,E] to [-1,1] so assume $A = -1 = -E$ and let $m > 2$.

<u>Corollary 2.2.1</u>. For m even the unique optimal design for estimating $\theta(c)$ is supported on the set $I_m = \{-x_{m/2},\ldots,-x_1, x_1,\ldots,x_{m/2}\}$ where $x_j = ((z_j^2 + D^2)/(z_j^2+1))^{1/2}$ and $z_j = \tan(\frac{(j-1)\pi}{m-2})$, for $j = 1,\ldots,m/2$. The masses are determined as usual and the resulting variance is

$$V_m = \sigma^2 N^{-1} S^2_{m/2}(c) = \frac{N^{-1}\sigma^2}{4}[\rho_+^{\frac{m}{2}-1} + \rho_-^{\frac{m}{2}-1}]^2,$$

where $\rho_\pm = S_2(c) \pm (S_2^2(-) - 1)^{1/2}$.

For m odd let $v_1 < \ldots < v_{m+1}$ be the points of I_{m+1}. The optimal designs are convex combinations of the two obtained as above from $\{v_1,\ldots,v_m\}$ and $\{v_2,\ldots,v_{m+1}\}$ and the variance is V_{m+1}.

<u>Proof</u>. The only part which requires some explanation is the formula for V_m. Since $V_m = \sigma^2 N^{-1} \|p_m^c\|^{-2}_{\infty,X}$ and $p_m^c(x) = p_m^0(x)/p_m^0(c)$ we have $V_m = \sigma^2 N^{-1} S^2_{m/2}(c)$. Solving the difference equation (2.1) yields the formula claimed. □

2.3. <u>Optimal Minimax Designs</u>

In this section it is assumed that the mean function θ is an unknown member of the set of functions $W_m^2[A,E]$ having absolutely continuous $m-1^{st}$ derivatives and square integrable m^{th} derivatives $\theta^{(m)}$. It is assumed that for some fixed known $\varepsilon > 0$ and all θ $\|\theta^{(m)}\|_2^2 = \int_A^E (\theta^{(m)}(t))^2 dt \le \varepsilon^2$. A linear estimator $\ell_0'Y$ is employed to estimate $\theta(c)$ which minimizes the maximum mean square error

$$\text{mmse} = \sup_\theta E_\theta(\ell'Y - \theta(c))^2 ,$$

where the supremum is over the set of θ's described above. The reader is referred to Speckman (1979) and Spruill (1984, 1985a, 1985b) for some further information about these estimators. If $\varepsilon = 0$ then the set of mean functions θ is the polynomials of degree no more than m-1 and the estimator becomes the usual least squares estimator. Let $A \le x_1 < \ldots < x_L \le B$ and $D \le x_{L+1} < \ldots < x_m \le E$. Define the associated function

$$\delta(x) = [R^2+\eta Q^2]^{-1}[\eta Q(\sum_{i=1}^{L} (-1)^{L-i}\phi_{x_i}(x) + \sum_{i=L+1}^{m} (-1)^{L+1-i}\phi_{x_i}(x))$$
$$+ \int_A^E h_c(t)h_x(t)dt]$$

where

$$(m-1)!h_x(t) = (x-t)_+^{m-1} - \sum_{i=1}^{m} \phi_{x_i}(x)(x_i-t)_+^{m-1} ,$$

$$R^2 = \|h_c\|_2^2 , \quad \text{and} \quad Q = \sum_{i=1}^{m} |\phi_{x_i}(c)| .$$

Let

$$\eta = \sigma^2/N\varepsilon^2 \in (0,\infty) .$$

<u>Theorem 2.3.1</u>. There is an optimal m-point design $\xi_{\eta,c}$ for estimating $\theta(c)$ if and only if the function δ determined

by the support of $\xi_{\eta,c}$ oscillates properly at the points x_i. In that case

$$\xi_{\eta,c}(x_i) = |\phi_{x_i}(c)| / \sum_{j=1}^{m} |\phi_{x_j}(c)|$$

for $i = 1,\ldots,m$, $\xi_{\eta,c}$ is unique among m point optimal designs, and δ minimizes

$$\rho(\theta) = ||\theta||_{\infty,X}^2 + \eta\, ||\theta^{(m)}||_2^2$$

among all functions $\theta \in W_m^2[A,E]$ which satisfy $\theta(c) = 1$.

In the symmetric case $X = [-1,-D] \cup [D,1]$, $c = 0$, there can be no m point optimal minimax designs if m is odd. This can be seen in the following way. Introduce the problems

P_η: minimize $\rho(\theta)$ over $\theta \in W_m^2[-1,1]$ such that $\theta(c) = 1$.

It can be shown that P_η has a unique solution θ_0 and it oscillates properly in at least m points. In the symmetric case $\theta_0(-x)$ solves P_η if θ_0 does, so $\theta_0(x) \equiv \theta_0(-x)$ and θ_0 must oscillate an even number of times numbering at least m+1. However, since every optimal design's support is contained in θ_0's oscillation set, the reflection of any m point optimal design being optimal entails the identity $\delta_1^{(m)}(x) \equiv \delta_2^{(m)}(x) \equiv \delta_0^{(m)}(x)$ of the two associated δ functions, an identity which one can easily determine to be impossible. In the symmetric case whenever m is odd the following theorem can be used. Let $x_1 < \ldots < x_{m+1}$ be points of X. Define the function $f_x(t)$ by

$$(m-1)!f_x(t) = (x-t)_+^{m-1} - \Sigma_{i=1}^{m+1} \phi_{x_i}(x)(x_i-t)_+^{m-1},$$

where $\{\phi_{x_i}\}_{i=1}^{m+1}$ are the Lagrange interpolation polynomials of degree m to $x_1,\ldots,x_{m+1}$. Taking $Q = \Sigma_{i=1}^{m+1}|\phi_{x_i}(0)|$ and $R^2 = \|f_0\|_2^2$, the associated function δ is

$$\delta(x) = [\eta Q^2+R^2]^{-1}[\eta Q(\sum_{i=1}^{\frac{m+1}{2}} (-1)^{\frac{m+1}{2}-i} \phi_{x_i}(x)$$

$$+ \sum_{i=\frac{m+3}{2}}^{m+1} (-1)^{\frac{m+3}{2}-i} \phi_{x_i}(x)) + \int_{-1}^{1} f_0(t)f_x(t)dt].$$

<u>Theorem 2.3.2.</u> There is an optimal design on m+1 points $\{x_1,\ldots,x_{m+1}\}$ if and only if the associated function δ oscillates properly at these points. If δ oscillates at $\{x_i\}_{i=1}^{m+1}$ then $x_{m+2-i} = -x_i$ for all i, the unique m+1 point optimal design ξ_0 places masses

$$\xi_0(x_i) = |\phi_{x_i}(0)|/\sum_{j=1}^{m+1}|\phi_{x_j}(0)|,$$

for $i = 1,\ldots,m+1$, δ solves P_η, and the mmse for the optimal design is $\frac{\sigma^2}{N}[Q^2 + \eta^{-1}R^2]$.

The proof of Theorem 2.3.2 can be carried out in the same way as that of Theorem 4.1 of Spruill (1984). The key facts in proving the present theorem are first, that for all $\theta \in W_m^2[-1,1]$

$$\theta(0) = \sum_{i=1}^{m+1} \phi_{x_i}(0)\theta(x_i) + \int_{-1}^{1} f_0(t)\theta^{(m)}(t)dt$$

and second, that, because of the symmetry, $\delta^{(m)}(x) = f_0(x)$.

In the preceeding theorems, the design ξ_0 is optimal if $d_\eta(\xi_0)$ is the minimum value of

$$d_\eta(\xi) = \sup_\theta\left(\frac{\theta^2(c)}{\int\theta^2(x)d\xi(x) + \eta\int(\theta^{(m)}(t))^2dt}\right)$$

over all Borel probability measures on X. For exact designs ξ the mmse of Speckman's estimator is $\frac{\sigma^2}{N} d_\eta(\xi)$, so $\frac{\sigma^2}{N} d_\eta(\xi_0)$ is approximately the smallest possible mmse. The characterizations given in the theorem can be used to find optimal designs, and were used to find those listed in Figures 2.4 through 2.7.

One can show that as $\eta \to \infty$ the optimal minimax designs are always m or m+1 point designs and that there is some optimal design $\bar{\xi}$ for polynomial interpolation which performs nearly as well as the optimal in that $d_\eta(\bar{\xi}) - d_\eta(\xi_{\eta,c}) \to 0$ as $\eta \to \infty$. This just means that when η is sufficiently large the optimal interpolation procedure for a polynomial mean, using the least squares estimator, will perform nearly as well as the optimal procedure, where knowledge of ε and σ is required.

Similar behavior is exhibited as c converges to B (or D) when η is fixed, there being optimal designs ξ_c for polynomial interpolation for which $d_\eta(\xi_c) - d_\eta(\xi_{\eta,c}) \to 0$ as $c \to B$.

2.4. Implementation

Suppose one is interested in interpolating a function θ whose behavior is basically that of a polynomial of degree m-1, but some protection is desired against deviations from this model, say $\|\theta^{(m)}\|_2 \leq \varepsilon$. Compute the optimal designs for polynomial interpolation whose supports will always contain B and D, take the approximate proportions of observations (see below also) required at these two points, and use them to estimate σ^2. Now find the optimal minimax design for the estimated value of η. We cannot at present prove that the optimal minimax design will include both B and D in its support even for ε small, but we have never found it to be otherwise computationally. If B and D are in the support then, since for ε small the masses assigned to B and D will not differ greatly from those already employed, the remaining points of the minimax design can be observed in their required proportions.

The optimal designs given in Sections 2 and 3 are only approximate. Their utilization in constructing actual experiments can be accomplished as described below, following Fedorov (1972). Let ξ be any design on $r < N$ points. Define a design $\tilde{\xi}$ obtained from ξ as any one obtained in the following way. With [x] = smallest integer greater than or equal to x, first assign $[(N-r)\xi(x_i)]$ of the N observations at x_i, $i = 1,\ldots,r$, then assign the remainder in any manner and call the resulting design $\tilde{\xi}$.

The inequalities (2.2) show that the exact design $\tilde{\xi}_{\eta,c}$ constructed from the optimal minimax design performs well in comparison to the optimal exact design, a fact which in the polynomial case is a consequence of Fedorov's arguments. Let Ξ_N be the exact designs supported on N or fewer points of X. An optimal exact design is $\xi_N(\eta,c)$ satisfying $d_\eta(\xi_N(\eta,c)) = \min_{\Xi_N} d_\eta(\xi)$. The inequalities are

$$0 \le 1 - d_\eta(\xi_N(\eta,c))/d_\eta(\tilde{\xi}_{\eta,c}) \le \frac{m}{N}$$

and

$$0 \le 1 - d_\eta(\xi_{\eta,c})/d_\eta(\xi_N(\eta,c)) \le \frac{m}{N} \qquad (2.2)$$

whenever $\xi_{\eta,c}$ is supported on m points. They are a consequence of the obvious inequalities $d_\eta(\xi_{\eta,c}) \le d_\eta(\xi_N(\eta,c)) \le d_\eta(\tilde{\xi}_{\eta,c})$ and the following lemma.

<u>Lemma 2.4.1</u>. Let ξ be any discrete design supported on $r < N$ points. Then

$$d_\eta(\tilde{\xi}) \le (1 - \frac{r}{N})^{-1} d_\eta(\xi).$$

<u>Proof</u>. The function d_η is well defined for any positive Borel measure. If ψ_1 and ψ_2 are two such measures and $k \ge 1$ is a constant then $kd_\eta(k\psi_1) \ge d_\eta(\psi_1)$ and $d_\eta(\psi_1) \ge d_\eta(\psi_1 + \psi_2)$. Since $\frac{N}{N-r}\tilde{\xi} = \xi + \psi$ where $\psi \ge 0$ we have $d_\eta(\xi) \ge d_\eta(\frac{N}{N-r}\tilde{\xi}) \ge (1 - \frac{r}{N}) d_\eta(\tilde{\xi})$. $\square$

As c tends to B or as $\eta \to \infty$ we should be able to use the designs ξ_c. How does $\tilde{\xi}_c$ compare with $\xi_N(\eta,c)$? Assume ξ_c and $\xi_{\eta,c}$ are supported on m points, otherwise replace m by m+1 where appropriate.

Lemma 2.4.2. $d_\eta(\tilde{\xi}_c) - d_\eta(\xi_N(\eta,c)) \le \gamma^2 d_\eta(\xi_c) - \gamma d_\eta(\xi_\eta,c)$ where $\gamma = (1 - \frac{m}{N})^{-1}$.

Proof. If $e = d_\eta(\xi_{\eta,c})/d_\eta(\xi_c)$ then

$$d_\eta(\tilde{\xi}_c) \le \gamma d_\eta(\xi_c) = \frac{\gamma d_\eta(\xi_{\eta,c})}{e} \le \frac{\gamma}{e} d_\eta(\xi_N(\eta,c))$$

so

$$d_\eta(\tilde{\xi}_c) - d_\eta(\xi_N(\eta,c)) \le (\frac{\gamma}{e} - 1) d_\eta(\xi_N(\eta,c)).$$

Since $d_\eta(\xi_N(\eta,c)) \le \gamma d_\eta(\xi_{\eta,c})$, one has

$$d_\eta(\tilde{\xi}_c) - d_\eta(\xi_N(\eta,c)) \le \gamma^2 d_\eta(\xi_c) - \gamma d_\eta(\xi_{\eta,c}). \qquad \square$$

When η is large or c is close B or D $\tilde{\xi}_c$ will be nearly as good an approximation to $\xi_N(\eta,c)$ as $\tilde{\xi}_{\eta,c}$. Even for reasonable values of N the bound can be small. For example, when m = 5 and $\eta = 10^{-5}$ with N = 20 we find that for $X = [-1,0] \cup [.7,1]$, c = .3, $d_\eta(\xi_c) = 3.06$ so the following design $\tilde{\xi}_c$, 1 at -1, 2 at -.62, 7 at 0, 5 at .7 and 2 at 1 with 3 anywhere has a resulting estimator whose maximum mean square error over $\|\theta^{(m)}\|_2 \le 70.71$ is $d_\eta(\tilde{\xi}_c)/20 \le d_\eta(\xi_{20}(\eta,c))/20 + .068$.

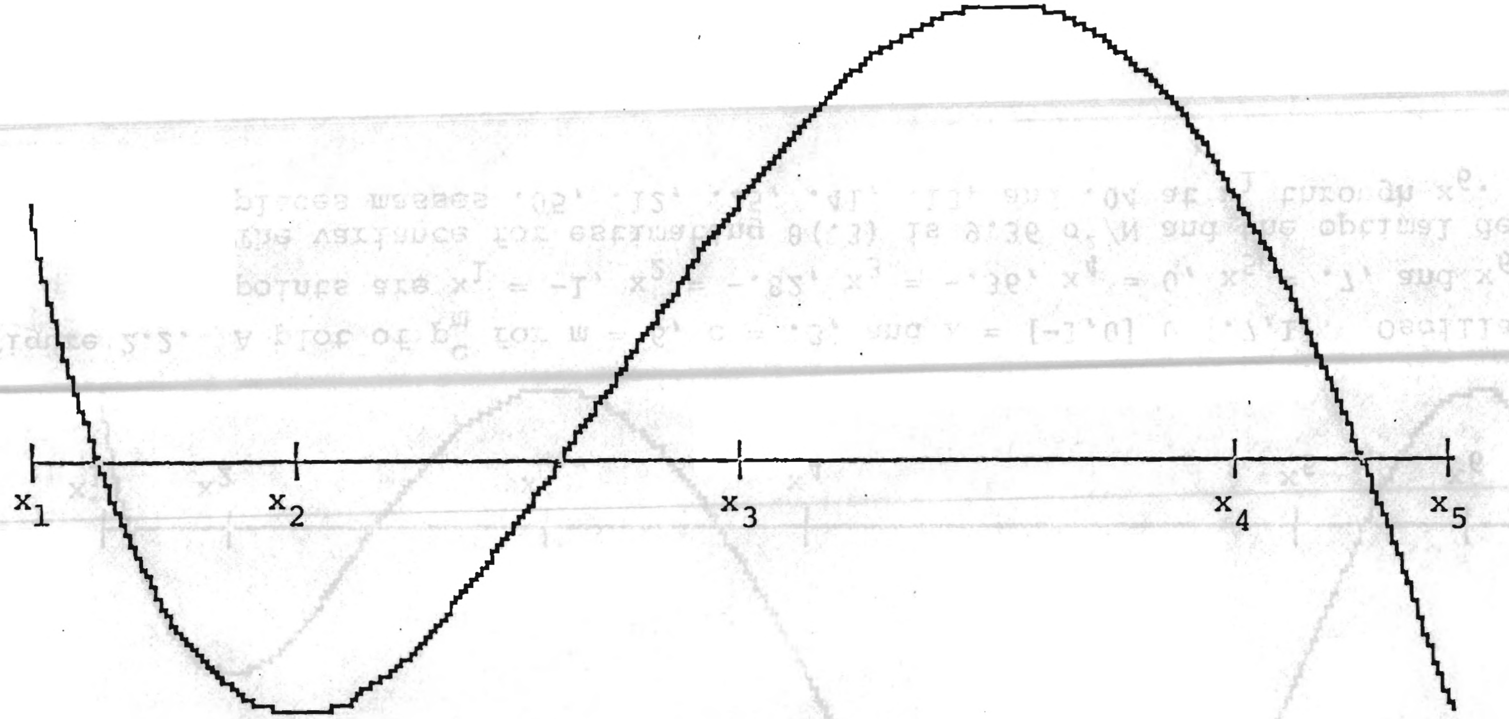

Figure 2.1. A plot of p_m^c for $m = 5$, $c = .3$, and $X = [-1,0] \cup [.7,1]$. Oscillation points are $x_1 = -1$, $x_2 = -.62$, $x_3 = 0$, $x_4 = .7$, and $x_5 = 1$. The variance for estimating $\theta(.3)$ is $2.99\ \sigma^2/N$ and the optimal design places masses .03, .13, .44, .31, and .09 at x_1 through x_5.

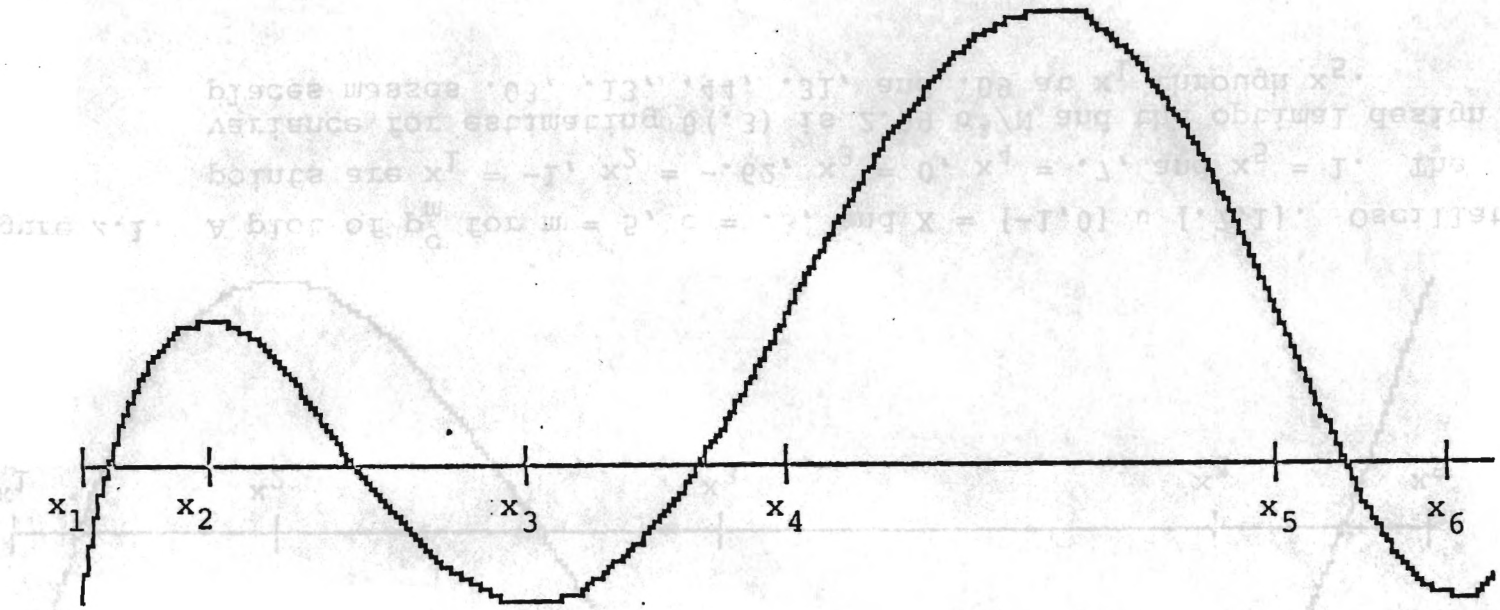

Figure 2.2. A plot of p_m^c for $m = 6$, $c = .3$, and $X = [-1,0] \cup [.7,1]$. Oscillation points are $x_1 = -1$, $x_2 = -.82$, $x_3 = -.36$, $x_4 = 0$, $x_5 = .7$, and $x_6 = .95$. The variance for estimating $\theta(.3)$ is $9.36\ \sigma^2/N$ and the optimal design places masses .05, .12, .25, .41, .13, and .04 at x_1 through x_6.

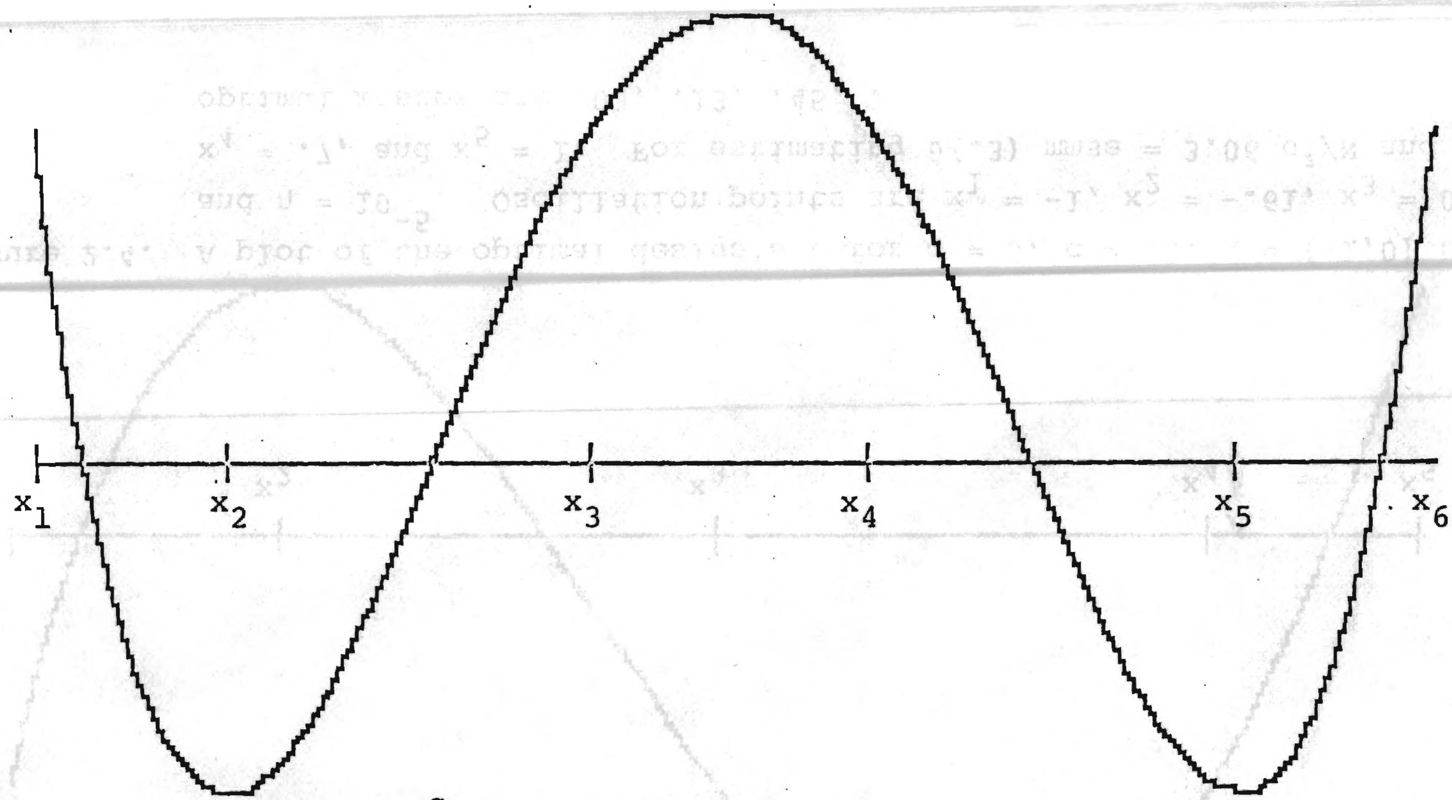

Figure 2.3. A plot of p_m^c for m = 5, c = 0, and X = [-1,-.2] ∪ [.2,1]. Oscillation points are $x_1 = -1 = -x_6$, $x_2 = -.72 = -x_5$, and $x_4 = -.2 = -x_4$. The variance for estimating $\theta(0)$ is 1.82 σ^2/N and optimal masses are .03, .11, .50, .34, and .02 at x_1 through x_5. With equal weight on this and its reflection through 0 one obtains the unique optimal design for m = 6.

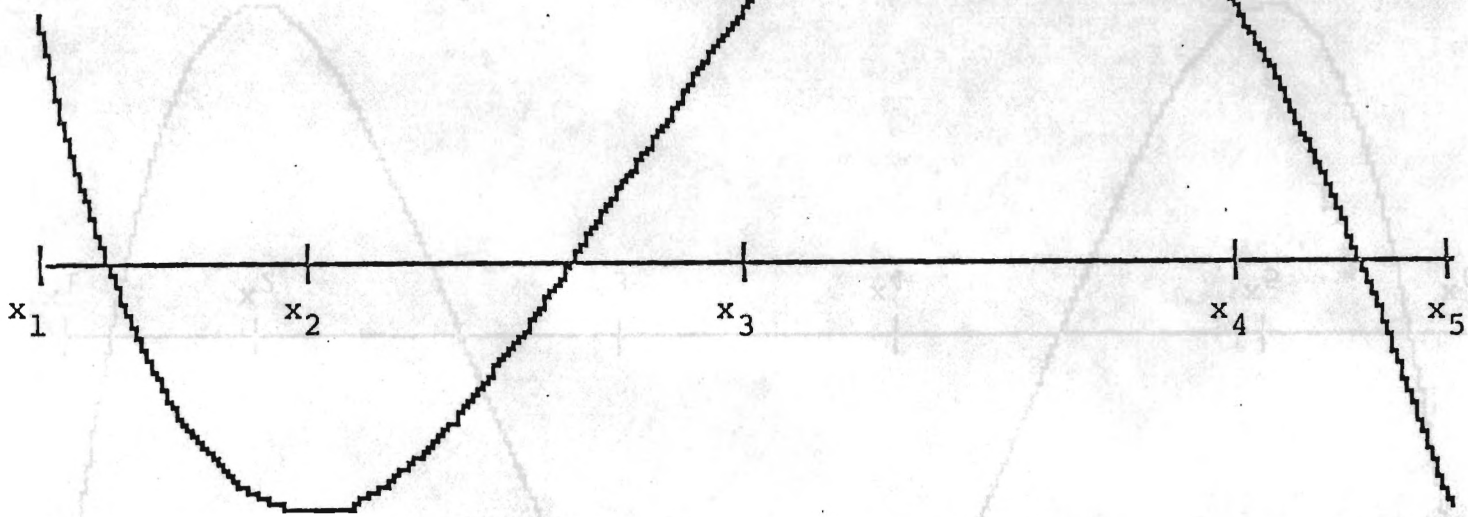

Figure 2.4. A plot of the optimal design's δ for $m = 5$, $c = .3$, $X = [-1,0] \cup [.7,1]$, and $\eta = 10^{-5}$. Oscillation points are $x_1 = -1$, $x_2 = -.61$, $x_3 = 0$, $x_4 = .7$, and $x_5 = 1$. For estimating $\theta(.3)$ mmse $= 3.06\ \sigma^2/N$ and the optimal masses are .03, .13, .45, .31, and .08.

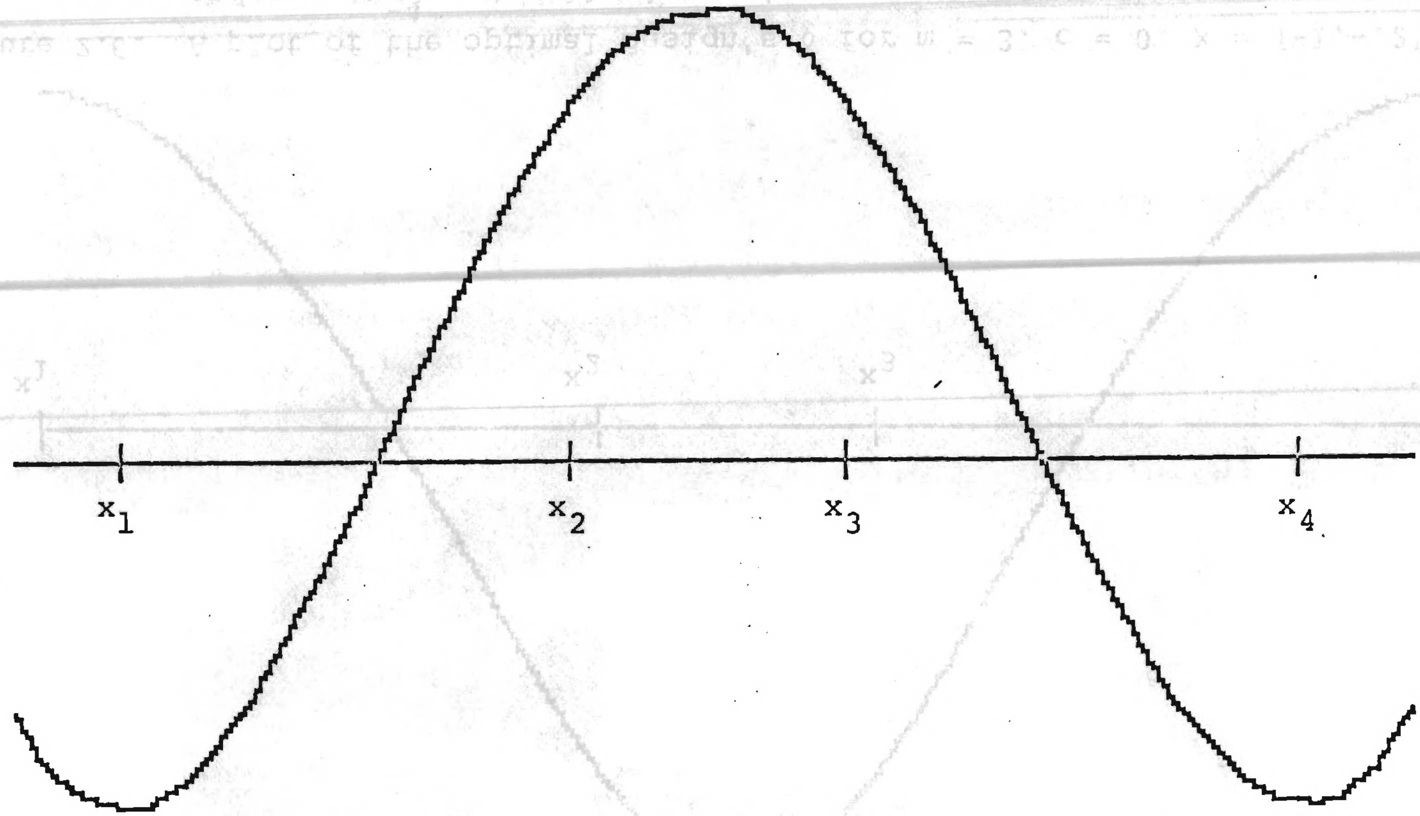

Figure 2.5. A plot of the optimal design's δ for $m = 4$, $c = 0$, $X = [-1,-.2] \cup [.2,1]$, and $\eta = 10^{-5}$. Oscillation points are $x_1 = -x_4 = -.85$ and $x_2 = -x_3 = -.2$. For estimating $\theta(0)$ mmse $= 1.43\ \sigma^2/N$ and the optimal masses at x_1 and x_2 are .03 and .47 with those at x_3 and x_4 determined by symmetry.

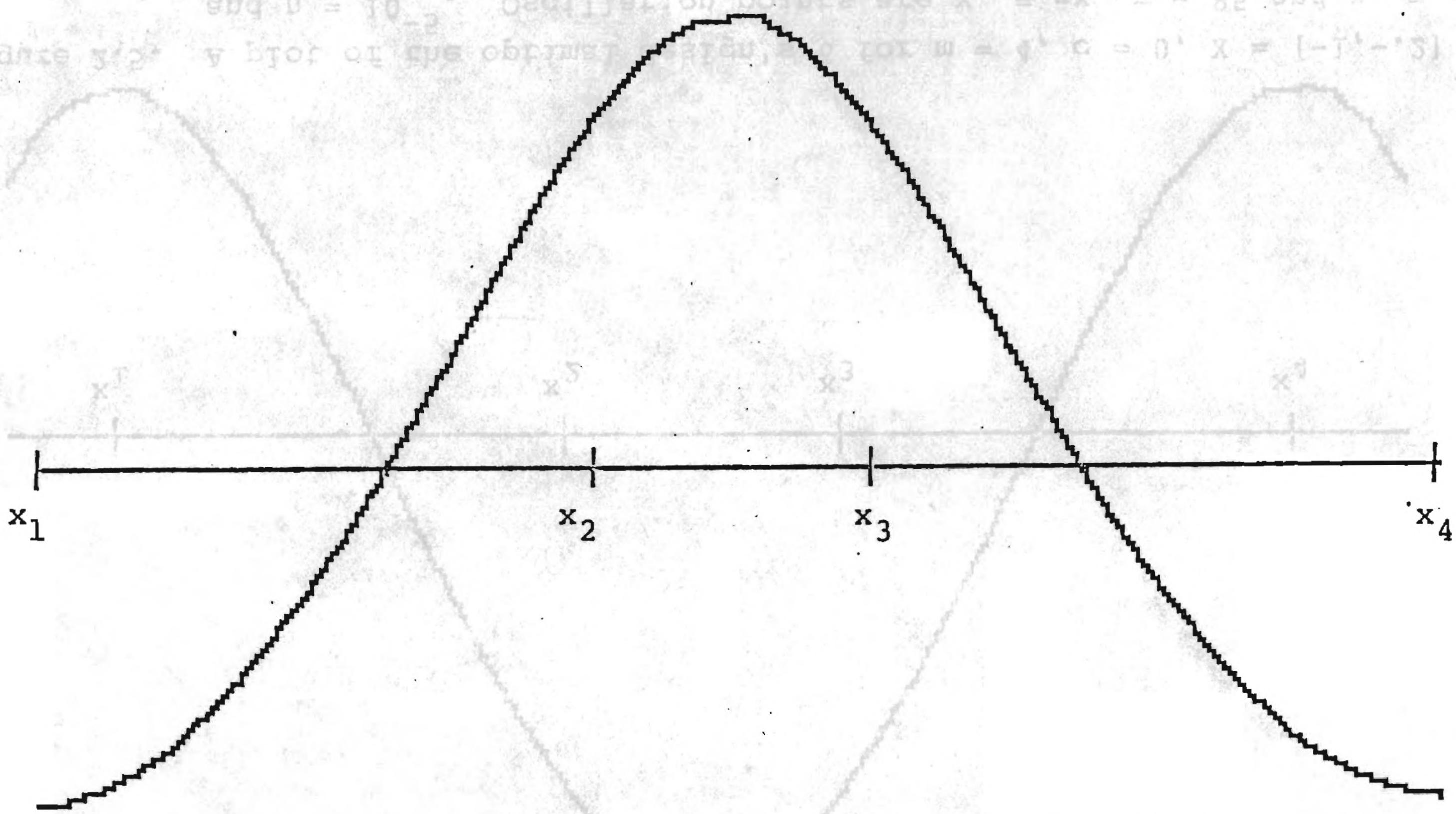

Figure 2.6. A plot of the optimal design's δ for $m = 3$, $c = 0$, $X = [-1,-.2] \cup [.2,1]$, and $\eta = 10^{-4}$. Oscillation points are $x_1 = -x_4 = -1$ and $x_2 = -x_3 = -.2$. For estimating $\theta(0)$ mmse $= 1.43\ \sigma^2/N$ and the optimal masses at x_1 and x_2 are .03 and .47.

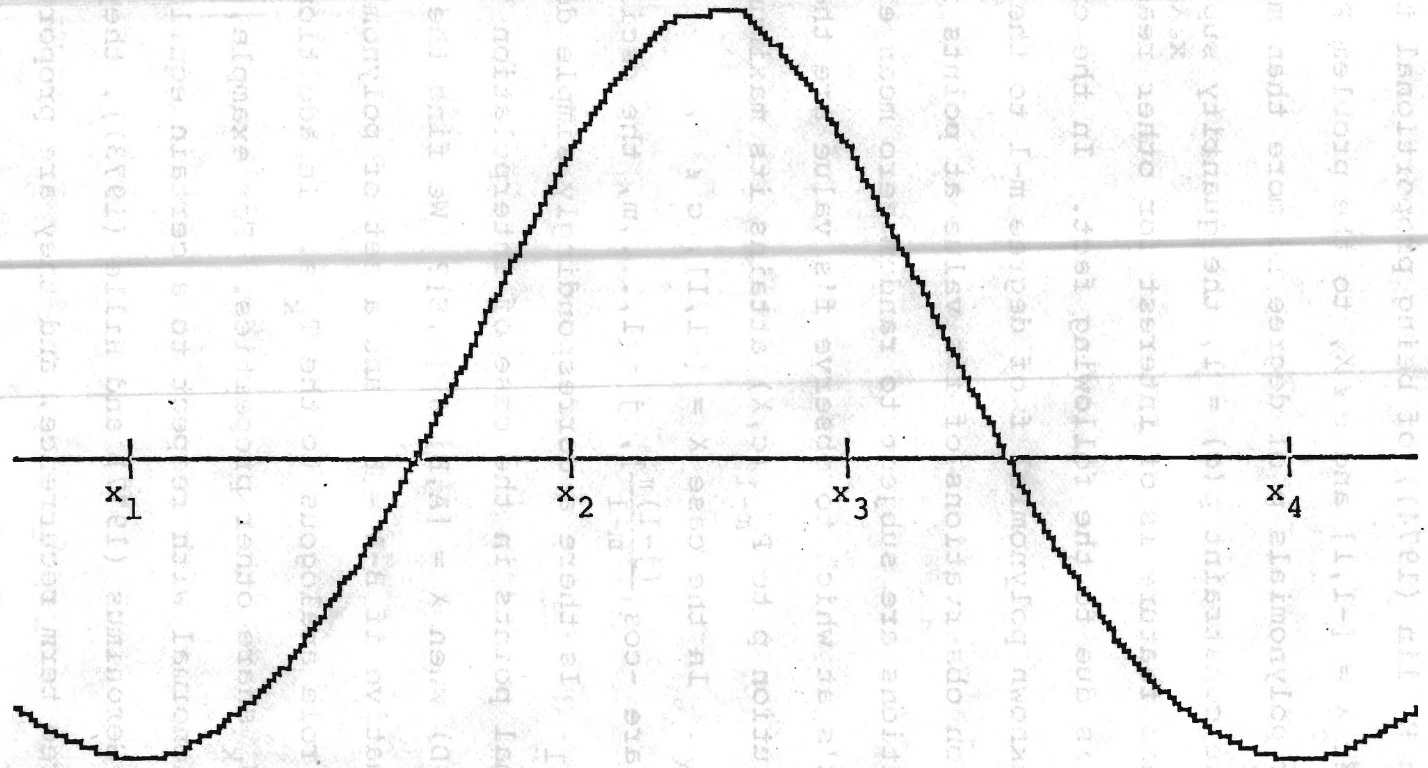

Figure 2.7. A plot of the optimal design's δ for $m = 3$, $c = 0$, $X = [-1,-.2] \cup [.2,1]$, and $\eta = 5\times10^{-5}$. Oscillation points are $x_1 = -x_4 = -.83$ and $x_2 = -x_3 = -.2$. For estimating $\theta(0)$ mmse $= 1.66\ \sigma^2/N$ and optimal masses for x_1 and x_2 are .03 and .47.

Some Chebyshev-Like Polynomials

3.1. Introduction

A Chebyshev polynomial $T_{m-1}(x) = \cos((m-1)\arccos(x))$ has the property (see Rivlin (1974)) of being proportional to the solution, whenever $X = [-1,1]$ and $c \notin X$, to the problem $P_{m-1}(c,X)$: minimize over polynomials p of degree no more than m-1, and subject to the constraint $p(c) = 1$, the quantity $\sup_{x \in X}|p(x)| = \|p\|_{\infty,X}$. This feature is of interest for other reasons, but our interest is due to the following fact. In the extrapolation of an unknown polynomial f of degree m-1 to the point $c \notin X$, based on observations of f's value at points x in X, which observations are subject to random zero mean error, the optimal x's at which to observe f's value are those for which the solution p to $P_{m-1}(c,X)$ attains its maximum absolute value $\|p\|_{\infty,X}$. In the case $X = [-1,1]$, $c \notin X$, those points are $-\cos[\frac{(j-1)\pi}{m-1}]$, $j = 1,\ldots,m$, the oscillation points of T_{m-1}. Is there a correspondingly simple determination of optimal points in the case of interpolation to a point $c \in (B,D)$ when $X = [A,B] \cup [D,E]$? We find the answer in the affirmative if $B-A = E-D$ and a set of polynomials S_k which play a role analogous to the T_k's. In addition, the polynomials S_k share other properties. For example, the S_k's are orthonormal with respect to a certain equilibrium measure (see Geronimus (1977) and Hille (1973)), they satisfy a similar three term recurrence, and they are proportional, when the degree is even, to the monic polynomials minimizing $\|q\|_{\infty,X}$.

Even when the two intervals are not of the same length the solutions, which are shown to be cosines of elliptic integrals, can be characterized by their oscillation properties providing a reasonably rapid method, not involving quadratures, for numerically finding solutions. Solutions to the monic minimizer of $\|q\|_{\infty,X}$ for odd degrees are also characterized.

3.2. Equioscillation

Consider the problem, for a given continuous function f on $X = [A,B] \cup [D,E]$, $Z_{m-1}(f)$: find $\bar{p} \in F_{m-1}$ minimizing $\|f-p\|_{\infty,X}$. Here $F_{m-1} = \{p \in P_{m-1}: p(c) = 0\}$, P_{m-1} is the collection of polynomials of degree no more than m-1, and $B < C < D$. Let $m \geq 1$.

Lemma 3.2.1. There is a solution $\bar{p}$ to $Z_{m-1}(f)$. Furthermore, for any solution there are subsets $R \subset [A,B]$ and $S \subset [D,E]$ with $\#(R) + \#(S) \geq m$, q_R, q_S each in $\{0,1\}$, and points $x_1 < x_2 < \ldots < x_L$ in R and $x_{L+1} < \ldots < x_m$ in S such that

$$f(x_j) - \bar{p}(x_j) = \begin{cases} (-1)^{j-q_R} \|f-\bar{p}\|_{\infty,X}, & j \in \{1,\ldots,L\} \\ (-1)^{j-q_S} \|f-\bar{p}\|_{\infty,X}, & j \in \{L+1,\ldots,m\} \end{cases} \tag{3.2.1}$$

Proof. Existence of the solution $\bar{p}$ is clear. The full proof then proceeds by examining the several possible cases, two of which are presented here. Assume $\#(R) + \#(S) = n < m$.

The first case has $\#(R)\#(S) > 0$ and $f(x_L) - \bar{p}(x_L) = f(x_{L+1}) - \bar{p}(x_{L+1})$. In this case consider

$$\tilde{p}(x) = \alpha(x-c)^2 \prod_{\substack{j=1 \\ j \neq L}}^{n-1} (x-z_j)$$

where $x_1 < z_1 < x_2 < \ldots < z_{L-1} < x_L < c < x_{L+1} < z_{L+1} < \ldots < x_n$ and $z_1,\ldots,z_{n-1}$ are zeros of $f-\bar{p}$. Then $\tilde{p} \in F_{m-1}$ and choosing α so that the sign of $\tilde{p}(x_1)$ is the same as $f(x_1) - \bar{p}(x_1)$ shows that for $\varepsilon > 0$ sufficiently small $\|f-(\bar{p}+\varepsilon\tilde{p})\|_{\infty,X} < \|f-\bar{p}\|_{\infty,X}$. The contradiction eliminates this case.

In the second case $\#(R)\#(S) > 0$ and $f(x_L) - \bar{p}(x_L) = -(f(x_{L+1}) - \bar{p}(x_{L+1}))$. Set

$$\tilde{p}(x) = \alpha(x-c) \prod_{\substack{i=1 \\ i\neq L}}^{n-1} (x-z_i),$$

where the z_j's are again zeros of $f-\bar{p}$. Again, for $\varepsilon > 0$ sufficiently small, $\bar{p} + \varepsilon\tilde{p}$ does better than $\bar{p}$ and is in F_{m-1}. □

Henceforth, we shall refer to the oscillation described by (2.1) as proper oscillation and the points x_j at which $|p(x)| = \|p\|_{\infty,X}$ as oscillation points. When $m = 1$ or 2 the solutions are $p_1(x) = p_2(x) = 1$. Let $m > 2$.

<u>Theorem 3.2.1</u>. There is a unique solution $p_m \in P_{m-1}$ to the problem $P_{m-1}(c,X)$. Furthermore, p is the solution if and only if p oscillates properly on X in at least m points, $p(c)=1$ and has B and D among the oscillation points with $p(B) = p(D) = \|p\|_{\infty,X}$.

<u>Proof</u>. Consider the solutions $\bar{p}$ to $Z_{m-1}(f)$. They form a convex set and must oscillate properly in at least m points. It follows that they all share at least m common points of oscillation and therefore that any two agree in at least m

points. The solutions to $Z_{m-1}(f)$ are thus unique. It is clear that solutions to $P_{m-1}(c,X)$ exist. Let $\tilde{p}$ be one and consider the solution $\bar{p}$ to $Z_{m-1}(\tilde{p})$. We must have $||\tilde{p}-\bar{p}||_{\infty,X} = ||\tilde{p}||_{\infty,X}$ and $\tilde{p}-\bar{p}$ oscillating properly. We also have 0 solving $Z_{m-1}(\tilde{p})$ so that $\bar{p} = 0$ and $\tilde{p}$ oscillates properly in at least m points. By the same arguments that applied to the solutions to $Z_{m-1}(f)$ it follows that $\tilde{p}$ is unique.

It is easy to see that at least three of {A,B,D,E} are in the oscillation set; otherwise p_m' would have at least m-1 zeros in [A,E] which is impossible since $m > 2$.

Assume that B is not in the oscillation set. Then p_m' has at least m-1 zeros in [A,E] unless $p_m(x_L) = -||p_m||_{\infty,X}$ and $p_m(D) = ||p_m||_{\infty,X}$. There are now two possibilities; either p_m has precisely m oscillation points or has m+1. If it has m then the proof of the second case in Lemma 3.2.1 shows we need $p_m(x_L) = ||p_m||_{\infty,X} = p_m(D)$; otherwise we obtain a contradiction to the definition of p_m. We conclude that if D only is in, then p_m oscillates at precisely m+1, since it can certainly oscillate in no more. Now there are several cases. Considering them each shows, as we show for just one, that each is impossible, and hence that B and D are both in the oscillation set. The case we show is that for which m is even. Then the oscillation pattern exhibited by p_m, where $M = ||p_m||_{\infty,X}$ and for #(R) even and #(S) odd, is

$$+M\ldots, +M, -M; +M, -M, \ldots, +M,$$

where the semicolon separates the two intervals [A,B] and [D,E]. We see that p_m has at least m-1 zeros and so must be degree m-1. Since this is odd it forces another zero so that p_m has at least m zeros. This is not possible.

Completing the scheme above shows that p_m must oscillate properly, have $p_m(c) = 1$, and $p_m(B) = p_m(D) = \|p\|_{\infty,X}$. If a polynomial p satisfies $p(c) = 1$, and $p(B) = p(D) = \|p\|_{\infty,X}$ then (see Lemma 1.2.2) it follows that $p-p_m$ has at least m+1 zeros on [A,E], so $p = p_m$. □

Using this characterization we found p_m's numerically for several cases. Graphs of these can be found in Figures 3.1 through 3.4.

3.3. Differential Equations

When $m > 2$ the solutions to $P_{m-1}(c,X)$ for different values of c must be scalar multiples. This is a consequence of the fact that if γ_j solves $P_{m-1}(c_j,X)$, $j = 1,2$, then $\alpha\gamma_1(B) = \gamma_2(B)$ for some $\alpha > 0$ and the characterization then shows $\alpha\gamma_1-\gamma_2$ has at least m zeros in [A,E]. If $\rho_{m-1} = \|p_m\|_{\infty,X}$, then $\rho_{m-1} \leq \rho_{n-1}$ for $m \geq n$, and since $\rho_2 < 1$ p_m' has a zero in (B,D). Henceforth, when we write p_m we shall mean the solution to the problem $P_{m-1}(c,X)$, where c has been chosen so that $p_m'(c) = 0$. Let $m > 2$.

Theorem 3.3.1. For some K, M, F, G, and Q, p_m satisfies one of the equations

$$(p_m')^2(x-v)(x-B)(x-D)(x-F) = K(M^2-p_m^2)(x-c)^2, \qquad (3.3.1)$$

where $v = A$ and $F \geq E$ or $v = E$ and $F \leq A$, or

$$(p_m')^2(x-A)(x-B)(x-D)(x-E)(x-F)(x-G) = K(M^2-p_m^2)(x-c)^2(x-Q)^2, \qquad (3.3.2)$$

where $E < Q < F < G$ or $G < F < Q < A$.

<u>Proof</u>. The complete proof proceeds by consideration of several cases. Two are presented below, the remainder being handled similarly. Certainly $m-2 \leq \deg(p_m) \leq m-1$.

If $\deg(p_m) = m-2$ and m is odd, then $M^2-p_m^2$ must have simple zeros at A, B, D, and E. If $L = \#(R)$ is odd and $U = \#(S)$ is even, then the sign pattern is

$$+M,\ldots,-M,\ +M\ ;\ +M,\ -M,\ldots,-M$$

and $p_m(\pm\infty) = \mp\infty$ (the case L even and U odd is similar). It follows that $M^2-p_m^2$, where $M = \|p_m\|_{\infty,X}$, has $2(m-4) + 4$ zeros and p_m' has $L - 2 + U - 2 + 1 = m-3$. Therefore, $(M^2-p_m^2)(x-c)^2$ has $2m-2$ and $(p_m')^2(x-A)(x-B)(x-D)(x-E)$ has $2m-6+4 = 2m-2$. Each is of degree $2m-2$ and their zeros agree, so for some constant K, p_m satisfies (3.3.1) with $v = A$ and $F = E$.

When $\deg(p_m) = m-2$ and m is even one can again show that $M^2-p_m^2$ has simple zeros at A, B, D, and E and that p_m satisfies (3.3.1) with $v = A$ and $F = E$.

Similarly, if $\deg(p_m) = m-1$ and $M^2-p_m^2$ has simple zeros at precisely A,B,D or B,D,E, then (3.3.1) is satisfied.

When $\deg(p_m) = m-1$ and $M^2-p_m^2$ has simple zeros at A, B, D, and E, then, as we show in the case m odd, p_m satisfies (3.3.2).

In the case m odd, $\deg(p_m) = m-1$, and $M^2-p_m^2$ with simple zeros at A, B, D, and E we have in case L is odd and U is even, the sign pattern

$$+M,\ldots,-M,+M \ ; \ +M, -M,\ldots,-M.$$

Since p_m is of even degree, either $p_m(\pm\infty) = +\infty$ or $p_m(\pm\infty) = -\infty$. Only the former case is treated since the argument for the latter is the same. In this case clearly p_m+M has a zero at $F > E$, p_m-M a zero at $G > F$, and p_m' a zero at $Q \in (E,F)$. Then $M^2-p_m^2$ has $2(m-4) + 4 = 2m-2$ zeros, p_m' has $L - 2 + U - 2 + 2 = m-2$ zeros, and the polynomials on both sides of (3.3.2) have 2m+2 zeros, are of degree 2m+2, and their zeros coincide. The equality, for some K, follows. The remaining subcases again show that p_m satisfies (3.3.2). □

We remark here that none of the cases listed in the theorem can be eliminated; all can be shown computationally to be possible depending on the configuration determined by A, B, D, and E. It also follows from the form of the equations that the solutions p_m can be expressed as cosines of elliptic integrals.

In general, the equations (3.3.1) and (3.3.2) seem to reveal little useful information since they contain several unknown quantities. An exception occurs in case B-A = E-D as will now be shown. We can assume A = -1, B = -D, and E = 1, where

$D \in (0,1)$, and it is not difficult to prove that in this case the special choice $c = 0$ leads to $p_m'(c) = 0$, $p_m(c) = 1$. Let $k \geq 2$.

Lemma 3.3.1. If $m = 2k-1$ then $p_m = p_{m+1}$ and $\deg(p_{m+1}) = 2k-2$.

Proof. Surely, if $p_{m+1}(x)$ is a solution then so is $p_{m+1}(-x)$. Uniqueness shows $p_{m+1}(x) - p_{m+1}(-x) = 0$ and since $p_{m+1}(x) = \alpha_m x^m + \alpha_{m-1} x^{m-1} + \ldots + \alpha_0$ we see that the coefficients of all odd powers vanish. Therefore $\deg(p_{m+1}) \leq 2k-2$. Since $p_{m+1}(x)$ has at least $m+1-2$ zeros and $p_{m+1}(0) = 1$, $\deg(p_{m+1}) \geq 2k-2$. $\square$

It now follows that for m even, $m > 2$, p_m is of degree $m-2$ and must satisfy (see the proof of Theorem 3.1)

$$(p_m')^2(1-x^2)(x^2-D^2) = K(M^2-p_m^2)x^2 .$$

It follows by integrating this that $p_m(x) = MS_{\frac{m}{2}}(x)$, where

$$S_k(x) = \cos((2k-2)\tan^{-1}(\sqrt{\frac{x^2-D^2}{1-x^2}})),$$

for $x \in [D,1)$.

Since S_k satisfies the three term recurrence

$$S_{k+1}(x) + S_{k-1}(x) = 2S_k(x)S_2(x), \quad k = 2,3,\ldots$$

on $[D,1)$, $S_1(x) \equiv 1$, and

$$S_2(x) = \frac{2x^2}{D^2-1} + \frac{1+D^2}{1-D^2}$$

there, we can extend the S_k's to $(-\infty,+\infty)$ where they can be seen to be polynomials. For m even, determining M_m so that $M_m S_{\frac{m}{2}}(0) = 1$, and using the oscillation characterization, confirms that $p_m = M_m S_{\frac{m}{2}}$ solves $P_{m-1}(0,X)$. Using the three term recurrence shows

$$M_m = 2\left[\left(\frac{1+D}{1-D}\right)^{\frac{m}{2}-1} + \left(\frac{1-D}{1+D}\right)^{\frac{m}{2}-1}\right]^{-1}.$$

Like the Chebyshev polynomials, the polynomials $\{S_k\}_{k\geq 1}$ form an orthogonal collection; specifically,

$$\frac{2}{\pi}\int_X S_k(x)S_{k'}(x)\frac{|x|dx}{\sqrt{(1-x^2)(x^2-D^2)}} = \delta_{kk'}$$

whenever k and k' are in $\{1,2,...\}$. The polynomials S_k are also related to the monic minimizers q of $||q||_{\infty,X}$. Some details are given in the next section.

3.4. The Approximation Problem

For an arbitrary $X = [A,B] \cup [D,E]$, what are the monic polynomials q which minimize $||q||_{\infty,X}$, and what is their relationship to the solutions of the problems $P_{m-1}(c,X)$? Introduce the problems, for f an arbitrary continuous function on $[A,E]$,

$$A_{m-1}(f)\text{: minimize } ||f-p||_{\infty,X} \text{ over all polynomials } p \in P_{m-1}.$$

The proof of the following lemma parallels that of Lemma 3.2.1.

Lemma 3.4.1. There is a solution $\bar{p}$ to $A_{m-1}(f)$. For each solution $\bar{p}$ there are subsets $R \subset [A,B]$ and $S \subset [D,E]$ such that

$\#(R) + \#(S) \geq m+1$, points q_R and q_S in $\{0,1\}$, and points $x_1 < x_2 < \ldots < x_L$ in R and $x_{L+1} < \ldots < x_{m+1}$ in S, such that

$$f(x_j) - \bar{p}(x_j) = \begin{cases} (-1)^{j-q_R} \|f-\bar{p}\|_{\infty,X} & j = 1,\ldots,L \\ (-1)^{j-q_S} \|f-\bar{p}\|_{\infty,X} & j = L+1,\ldots,m+1 \end{cases} .$$

This proper oscillation at m+1 points again guarantees the uniqueness of the solutions to $A_{m-1}(f)$. The monic minimizer q of interest is a solution, where $m \geq 2$, of

$$MP_{m-1}: \text{ minimize } \|q\|_{\infty,X} \text{ over all polynomials } q(x) = x^{m-1} - v(x), \text{ such that } v \in P_{m-2}.$$

A solution q to MP_{m-1} satisfies $q(x) = x^{m-1} - \bar{v}(x)$, where $\bar{v}$ solves $A_{m-2}(x^{m-1})$, so q must oscillate properly in at least $m-2+2 = m$ points. Again this implies the unicity of the solution. The following can be proven using no new techniques.

<u>Theorem 3.4.1</u>. The monic polynomial $\bar{q}$ of degree m-1 solves MP_{m-1} if either

a) $\bar{q}$ oscillates properly at m+1 points or

b) $\bar{q}$ oscillates properly at m points, A and E are in the oscillation set, and $\bar{q}(x_L) = -\bar{q}(x_{L+1})$.

Conversely, if $\bar{q}$ solves MP_{m-1} then $\bar{q}$ satisfies either a) or b).

We note that if a monic polynomial $q \in P_{m-1}$ oscillates properly at m+1 points then, necessarily, m is odd and both B and D are in the oscillation set. This implies that if m is odd, q_m solves MP_{m-1}, and q_m oscillates properly at m+1 points then q_m is a multiple of p_m, the solution to $P_{m-1}(c)$. If m is even then the solutions to MP_{m-1} and $P_{m-1}(c)$ cannot be multiples since their oscillation patterns are unlike. One can prove as in Theorem 3.3.1 that the minimizer q_m must satisfy one of the equations

$$(q_m')^2(x-A)(x-E) = K(M^2-q_m^2) \tag{3.4.1}$$

or

$$(q_m')^2(x-A)(x-B)(x-D)(x-E)(x-H)(x-I) = K(M^2-q_m^2)(x-F)^2(x-G)^2 \tag{3.4.2}$$

for some values K,M, and $B < F < H < I < G < D$. Even in the symmetric case we have no explicit solution for m even. When m is odd however we know that $p_m = p_{m+1}$ so that p_m oscillates m+1 times and B and D are in the oscillation set. Using that fact and the three term recurrence one can prove the following. Let $m \geq 3$ be odd.

<u>Lemma 3.4.2</u>. The solution q_m of MP_{m-1} satisfies

$$q_m(x) = \frac{1}{2}\left(\frac{4}{D^2-1}\right)^{\frac{m-1}{2}} S_{\frac{m+1}{2}}(x), \quad x \in (-\infty,+\infty)$$

and

$$||q_m||_{\infty,X} = 2\left(\frac{1-D^2}{4}\right)^{\frac{m-1}{2}} .$$

Standard recurrence formulas for orthogonal polynomials generate odd degree polynomials to complete the q set. However, we cannot show that they are q_m's for m even.

The last lemma shows that the logarithmic capacity for two intervals of equal length L separated by a distance Δ is $\frac{1}{2}\sqrt{L(L+\Delta)}$. In case the intervals are $[-1,-D] \cup [D,1]$ this is $\frac{\sqrt{1-D^2}}{2}$, a value which is known and can be found in (VI.33) of Geronimus (1977). There also the measure (see also II.15 of Geronimus)

$$\frac{|x|dx}{\sqrt{(1-x^2)(x^2-D^2)}}$$

occurs as a potential theoretic equilibrium distribution. I am indebted to Professor J. Geronimo for informing me of the connections with potential theory.

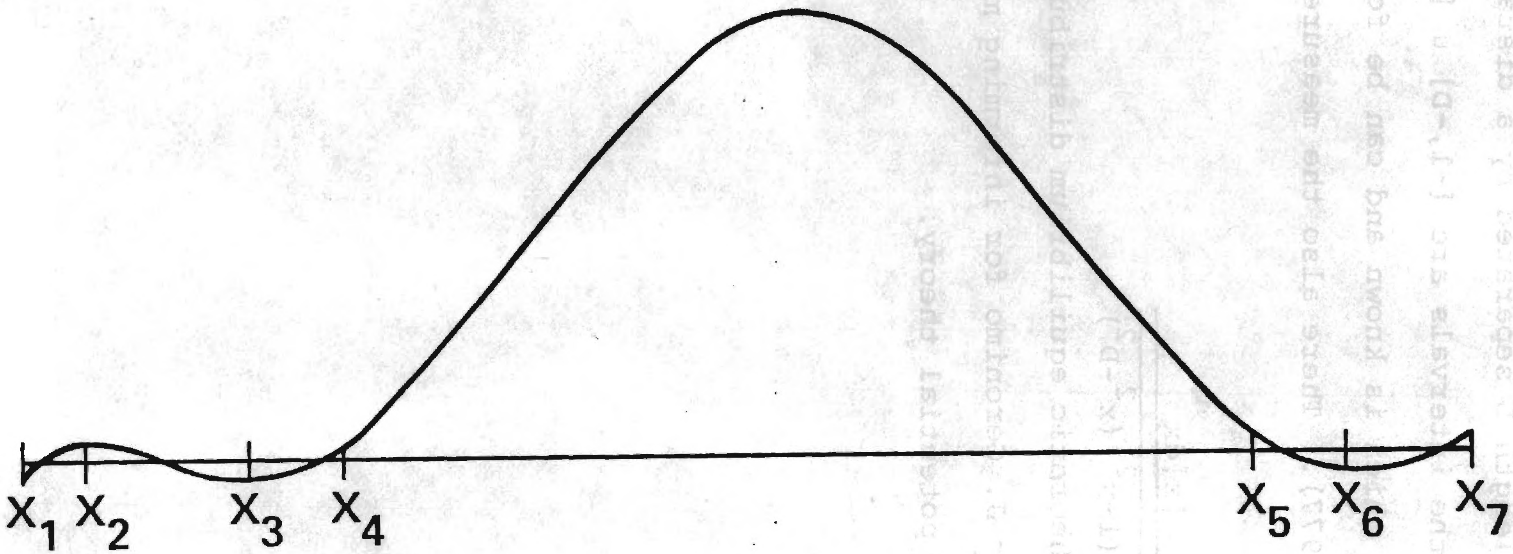

Figure 3.1. Plot of p_m for $m = 7$, $X = [-1,-.55] \cup [.7,1]$. Oscillation points are $x_1 = -1$, $x_2 = -.907$, $x_3 = -.688$, $x_4 = -.550$, $x_5 = .7$, $x_6 = .834$, $x_7 = 1$.

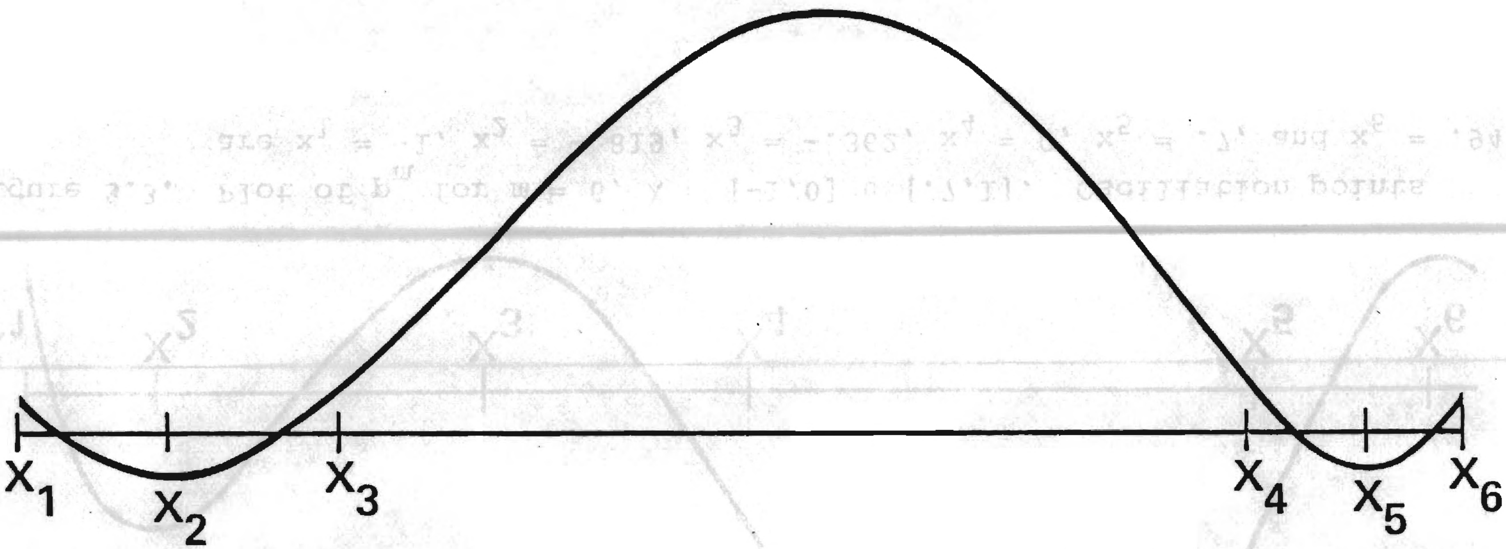

Figure 3.2. Plot of p_m for $m = 6$, $X = [-1,-.55] \cup [.7,1]$. Oscillation points are $x_1 = -x_6 = -1$, $x_2 = -.787$, $x_3 = -.55$, $x_4 = .7$, $x_5 = .868$.

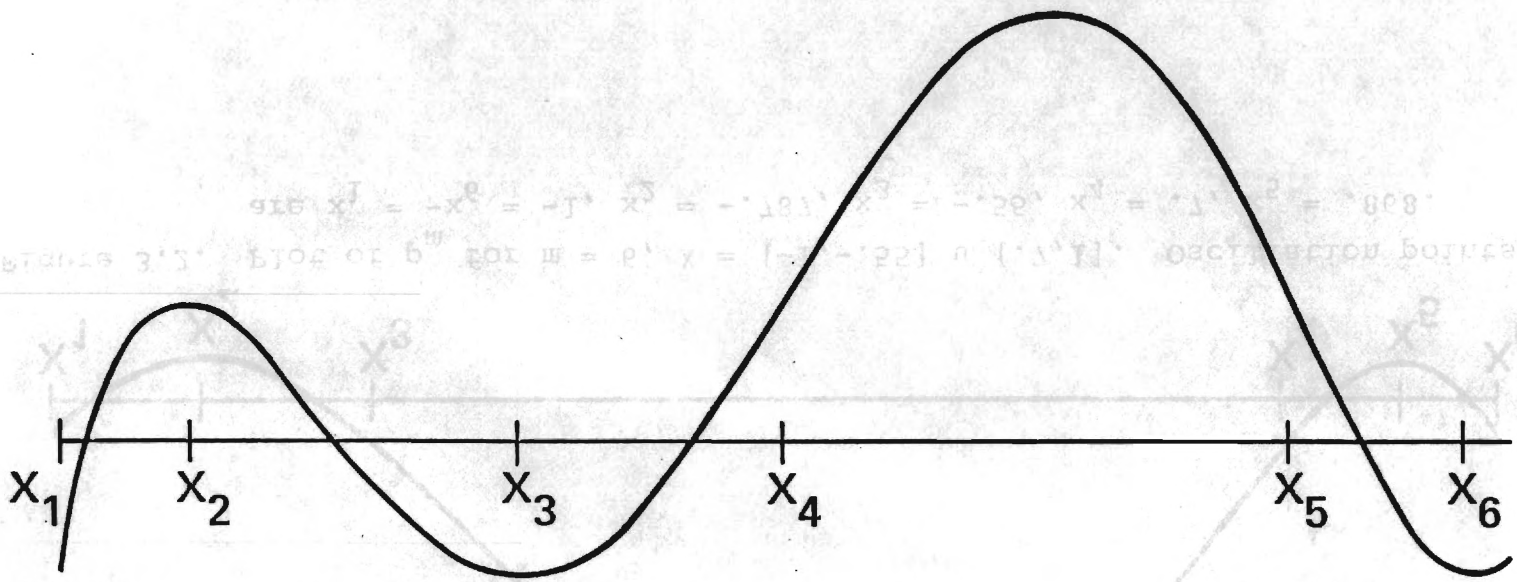

Figure 3.3. Plot of p_m for $m = 6$, $X = [-1,0] \cup [.7,1]$. Oscillation points are $x_1 = -1$, $x_2 = -.819$, $x_3 = -.362$, $x_4 = 0$, $x_5 = .7$, and $x_6 = .948$.

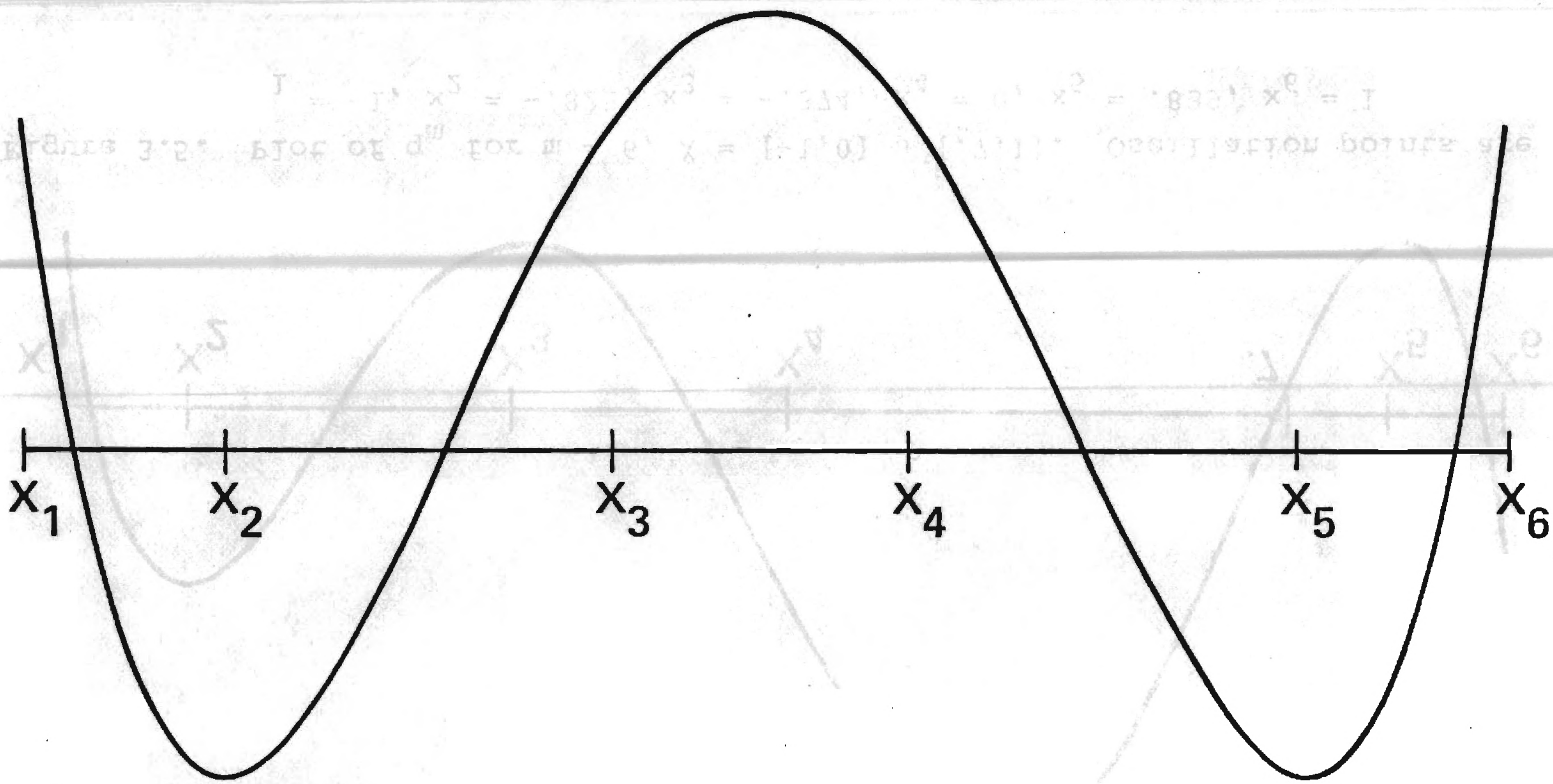

Figure 3.4. Plot of p_m, $m = 5$ for $X = [-1,-.2] \cup [.2,1]$. Oscillation points are $x_6 = -x_1 = 1$, $x_5 = -x_2 = .721$, $x_4 = -x_3 = .2$.

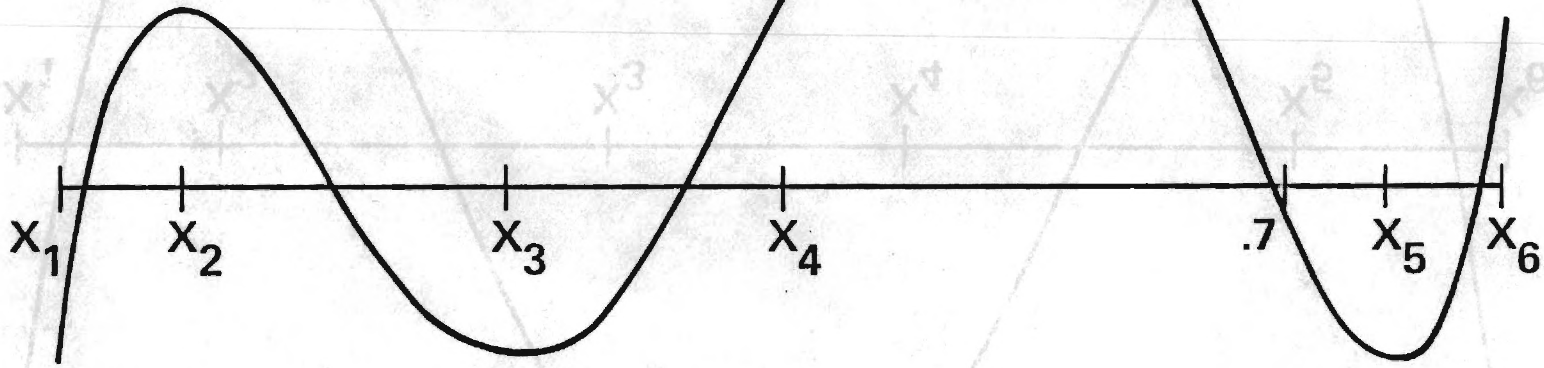

Figure 3.5. Plot of q_m for $m = 6$, $X = [-1,0] \cup [.7,1]$. Oscillation points are $x_1 = -1$, $x_2 = -.823$, $x_3 = -.374$, $x_4 = 0$, $x_5 = .839$, $x_6 = 1$.

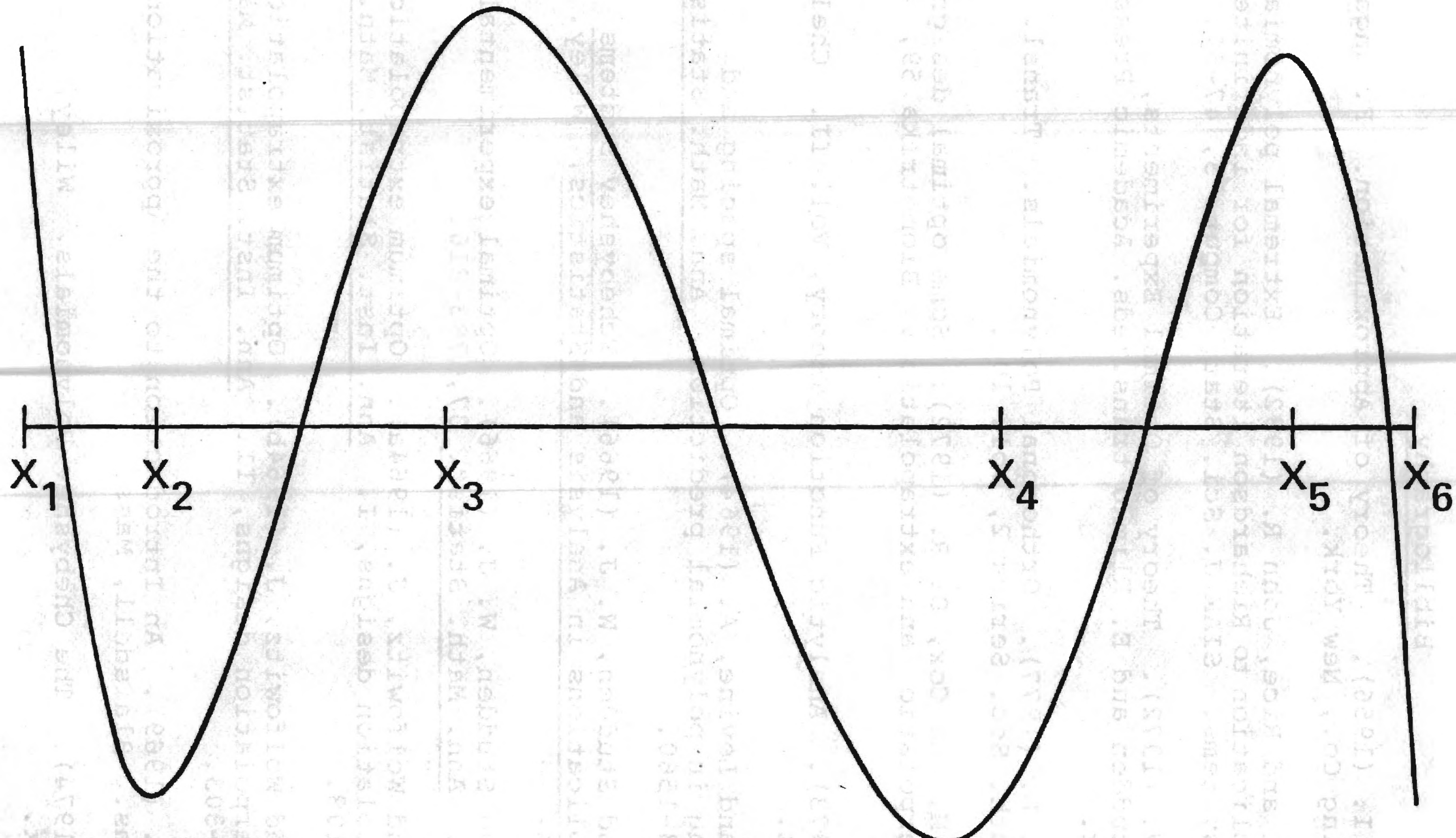

Figure 3.6. Plot of q_m for $m = 6$, X $[-1,-.4] \cup [.4,1]$. Oscillation points are $x_6 = -x_1 = 1$, $x_5 = -x_2 = .813$, $x_4 = -x_3 = .4$.

Bibliography

Achieser, N. I. (1956). Theory of Approximation. F. Ungar Publishing Co., New York.

De Boor, Carl and Rice, John R. (1982). Extremal polynomials with application to Richardson iteration for indefinite linear systems. SIAM J. Sci. Stat. Comput. 3, 47-57.

Fedorov, V. V. (1972). Theory of Optimal Experiments, W. J. Studden and E. Klimko trans./eds. Academic Press, New York.

Geronimus, Y. L. (1977). Orthogonal Polynomials. Transl. Amer. Math. Soc. Series 2, Vol. 108.

Herzberg, A. M. and Cox, D. R. (1972). Some optimal designs for interpolation and extrapolation. Biometrika 59, 551-561.

Hille, E. (1973). Analytic Function Theory, Vol. II. Chelsea, New York.

Hoel, P. G. and Levine, A. (1964). Optimal spacing and weighting in polynomial prediction. Ann. Math. Statist. 33, 1553-1560.

Karlin, S. and Studden, W. J. (1966). Tchebyshev Systems with Applications in Analysis and Statistics. Wiley. New York.

Karlin, S. and Studden, W. J. (1966). Optimal experimental designs. Ann. Math. Statist. 37, 783-816.

Kiefer, J. and Wolfowitz, J. (1964a). Optimum extrapolation and interpolation designs, I. Ann. Inst. Statist. Math. 16, 79-108.

Kiefer, J. and Wolfowitz, J. (1964b). Optimum extrapolation and interpolation designs, II. Ann. Inst. Statist. Math. 16, 295-303.

Rivlin, T. J. (1969). An Introduction to the Approximation of Functions. Blaisdell, Mass.

Rivlin, T. (1974). The Chebyshev Polynomials. Wiley, New York.

Speckman, P. (1979). Minimax estimates of linear functionals in a Hilbert space. (manuscript).

Spruill, C. (1980). Optimal designs for second order processes with general linear means. Ann. Statist. 8, 652-663.

Spruill, M. C. (1984). Optimal designs for minimax extrapolation. J. Multivariate Anal. 15, 52-62.

Spruill, Carl (1985a). Model robustness of Hoel-Levine optimal designs. J. Statist. Plann. Inference 11, 217-225.

Spruill, M. C. (1985b). Optimal extrapolation of derivatives. Metrika (to appear).

Spruill, M. C. (1987). Optimal designs for interpolation. J. Statist. Plann. Inference (to appear).

Studden, W. J. (1968). Optimal designs on Tchebycheff points. Ann. Math. Statist. 39, 1435-1477.

Wahba, G. (1978). Improper priors, spline smoothing and the problems of guarding against model errors in regression. J. R. Statist. Soc. B 40, 364-372.