

## OPTIMAL DESIGNS FOR RATIONAL MODELS AND WEIGHTED POLYNOMIAL REGRESSION

BY HOLGER DETTE, LINDA M. HAINES<sup>1</sup> AND LORENS IMHOF

*Ruhr-Universität Bochum, University of Natal and RWTH Aachen*

In this paper  $D$ -optimal designs for the weighted polynomial regression model of degree  $p$  with efficiency function  $(1 + x^2)^{-n}$  are presented. Interest in these designs stems from the fact that they are equivalent to locally  $D$ -optimal designs for inverse quadratic polynomial models. For the unrestricted design space  $\mathbb{R}$  and  $p < n$ , the  $D$ -optimal designs put equal masses on  $p + 1$  points which coincide with the zeros of an ultraspherical polynomial, while for  $p = n$  they are equivalent to  $D$ -optimal designs for certain trigonometric regression models and exhibit all the curious and interesting features of those designs. For the restricted design space  $[-1, 1]$  sufficient, but not necessary, conditions for the  $D$ -optimal designs to be based on  $p + 1$  points are developed. In this case the problem of constructing  $(p + 1)$ -point  $D$ -optimal designs is equivalent to an eigenvalue problem and the designs can be found numerically. For  $n = 1$  and  $2$ , the problem is solved analytically and, specifically, the  $D$ -optimal designs put equal masses at the points  $\pm 1$  and at the  $p - 1$  zeros of a sum of  $n + 1$  ultraspherical polynomials. A conjecture which extends these analytical results to cases with  $n$  an integer greater than  $2$  is given and is examined empirically.

**1. Introduction.** Weighted polynomial regression models with variance functions which depend on the explanatory variable have played, and continue to play, an important role in the development of classical optimal design theory. The reasons for this are essentially twofold. First, there is a wealth of elegant mathematics associated with the construction of  $D$ -optimal designs for many of these models. In particular for certain classes of variance functions, it is possible to show that the  $D$ -optimal designs put equal masses on  $p + 1$  points of support, where  $p$  is the degree of the polynomial embedded in the model, and, furthermore, it is possible to use tools from the theory of differential equations and of canonical moments in order to establish that these support points coincide with the zeros of classical orthogonal polynomials [see, e.g., Fedorov (1972), pages 85–91; Studden (1980); Dette and Studden (1997), Section 5.5]. Second, there are certain regression models, both linear and nonlinear, which are particularly important in practice and for which the problem of constructing  $D$ -optimal and locally  $D$ -optimal designs, respectively, can be recast as that of finding  $D$ -optimal designs for specific

---

Received March 1999; revised June 1999.

<sup>1</sup>Supported in part by the National Research Foundation, South Africa.

AMS 1991 *subject classifications*. Primary 62K05; secondary 34L40.

*Key words and phrases*.  $D$ -optimal design, weighted polynomial regression, rational models, Schrödinger equation.

weighted polynomial regression models. In such cases the optimal designs can be readily obtained numerically but analytic results are, at least in general, difficult to derive and this is clearly demonstrated in the papers of Huang, Chang and Wong (1995), He, Studden and Sun (1996) Chang and Lin (1997), Ortiz and Rodríguez (1998) or Imhof, Krafft and Schaefer (1998).

The problem of constructing locally  $D$ -optimal designs for rational or inverse polynomial models with constant error terms was shown by He, Studden and Sun (1996) to be equivalent to that of finding  $D$ -optimal designs for weighted polynomial regression models with variance functions which are polynomials in the explanatory variable, and some general results for these designs were also presented in that paper. As a counterpoint to this, explicit algebraic expressions for the support points of the  $D$ -optimal designs for a particular inverse quadratic model were derived by Haines (1992), following the numerical results of Cobby, Chapman and Pike (1986). The aim of the present study is to draw these two strands of research together and in particular to search for analytic solutions to the problem of constructing  $D$ -optimal designs for inverse quadratic models and their weighted polynomial regression model counterparts.

This paper is organized in the following way. In Section 2 the problem of constructing  $D$ -optimal designs for a polynomial regression model of degree  $p$  with efficiency function of the form  $(1 + x^2)^{-n}$ , where  $x$  is the explanatory variable,  $n \in \mathbb{R}$  and the efficiency function is the reciprocal of the variance function, is introduced and is shown to be equivalent to that of finding  $D$ -optimal designs for inverse quadratic models. Some results relating to numbers of support points for these  $D$ -optimal designs are also presented there. The construction of  $D$ -optimal designs for this weighted polynomial regression model with an unrestricted design space  $\mathbb{R}$  and  $p \leq n$  are considered in Section 3 and those for a restricted design space  $[-1, 1]$  in Section 4. Some broad conclusions and pointers for future research are given in Section 5.

## 2. Preliminary remarks.

2.1. *The design problem.* Consider the polynomial regression model of degree  $p$  defined by

$$(2.1) \quad y = \beta_0 + \beta_1 x + \cdots + \beta_p x^p + \varepsilon,$$

where  $y$  is the response corresponding to an explanatory variable  $x$ , with  $x$  taken from the design space  $\mathcal{X}$ ,  $\beta = (\beta_0, \dots, \beta_p)$  is a vector of unknown parameters, and the term  $\varepsilon$  represents a random error with mean 0 and variance  $\sigma^2/\lambda(x)$ , where  $\lambda(x)$  is an efficiency function of the form

$$(2.2) \quad (1 + x^2)^{-n}, \quad n \in \mathbb{R}.$$

Then for an approximate design  $\xi$ , which is a probability measure on the design space  $\mathcal{X}$ , the Fisher information matrix for the parameters  $\beta$  can be

expressed as

$$(2.3) \quad M(\xi) = \int_{\mathcal{X}} \lambda(x) f_p(x) f_p^T(x) d\xi(x),$$

where  $f_p(x)$  denotes the vector of monomials up to order  $p$ , that is,  $f_p(x) = (1, x, \dots, x^p)^T$ . The problem addressed in the present study is that of constructing  $D$ -optimal designs for the polynomial regression model (2.1) with independent heteroscedastic error terms specified by (2.2) and thus of finding designs which maximize the determinant of the information matrix (2.3).

2.2. *A related locally D-optimal design problem for rational models.* Consider a rational model with homoscedastic error terms and with an expected response given by

$$(2.4) \quad \eta(u, \theta) = \frac{\theta_0 + \theta_1 u + \dots + \theta_q u^q}{(1 + \theta_{q+1} u + \theta_{q+2} u^2)^m},$$

where the explanatory variable  $u$  is taken from a design space  $\mathcal{X} \subset \mathbb{R}$ ,  $\theta = (\theta_0, \dots, \theta_{q+2})^T$  denotes a vector of unknown parameters, and the polynomial  $Q(u) = 1 + \theta_{q+1} u + \theta_{q+2} u^2$  is assumed to be bounded away from 0 and positive. Then, following He, Studden and Sun (1996), it is readily shown that the Fisher information matrix for  $\theta$  at the point  $u$  can be expressed as

$$(2.5) \quad I(u, \theta) = \frac{1}{Q^{2m+2}(u)} B(\theta) f_{q+2}(u) f_{q+2}^T(u) B^T(\theta),$$

where  $B(\theta)$  is the  $(q + 3) \times (q + 3)$ -matrix defined by

$$(2.6) \quad B(\theta) = \begin{bmatrix} 1 & \theta_{q+1} & \theta_{q+2} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \theta_{q+1} & \theta_{q+2} & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 & \theta_{q+1} & \theta_{q+2} \\ 0 & -m\theta_0 & -m\theta_1 & -m\theta_2 & -m\theta_3 & \dots & -m\theta_{q-1} & -m\theta_q & 0 \\ 0 & 0 & -m\theta_0 & -m\theta_1 & -m\theta_2 & \dots & -m\theta_{q-2} & -m\theta_{q-1} & -m\theta_q \end{bmatrix}.$$

The information matrix (2.5) depends on the unknown parameters  $\theta$  and a locally optimal design  $\zeta$  on the design space  $\mathcal{X}$  maximizes a concave function of the information matrix

$$(2.7) \quad M(\zeta, \theta) = \int_{\mathcal{X}} I(u, \theta) d\zeta(u)$$

for a fixed value or “best guess” of  $\theta$  [see Chernoff (1953)]. Now it is immediately clear that maximizing the determinant of the matrix (2.7) with respect to the design measure  $\zeta$  is equivalent to maximizing the determinant of the information matrix for a weighted polynomial regression model of degree  $p = q + 2$  with efficiency function of the form (2.2) and  $n = 2m + 2$ , and hence that the problem of constructing  $D$ -optimal designs for the two

models is the same. Moreover, the support points of the locally  $D$ -optimal design for the rational model are independent of the parameters in the numerator of the rational function (2.4) and can be obtained from those for the  $D$ -optimal design of the corresponding weighted regression model by the linear transformation

$$(2.8) \quad x \rightarrow u = t(x) = \frac{1}{\sqrt{\theta_{q+2}}} \left\{ x \left( 1 - \frac{\theta_{q+1}^2}{4\theta_{q+2}} \right)^{1/2} - \frac{\theta_{q+1}}{2\sqrt{\theta_{q+2}}} \right\}.$$

For  $m \geq q/2$  locally  $D$ -optimal designs for the rational model with design space  $\mathcal{X}$  equal to  $\mathbb{R}$  exist and the transformation (2.8) is thus a straightforward one [see, e.g. Haines (1992)]. However, for  $m < q/2$ , the locally  $D$ -optimal designs exist only for design spaces which are bounded intervals, say  $[a, b]$ , and some care is therefore needed when invoking the transformation (2.8) in that the endpoints of the interval have to be transformed from those relating to the weighted polynomial regression model to give  $[t^{-1}(a), t^{-1}(b)]$ .

*2.3. The number of support points.* Consider again the weighted polynomial regression model (2.1) with efficiency function (2.2). The results of this section give a partial solution to the problem of determining the number of support points of the  $D$ -optimal designs for these models.

**LEMMA 2.1.** *A  $D$ -optimal design for the weighted polynomial regression model (2.1) with efficiency function (2.2) has exactly  $p + 1$  points of support if any one of the conditions:*

- (i)  $p < n$ ;
- (ii)  $n = 0, 1, \dots, p - 1$  with  $\mathcal{X}$  a bounded interval;
- (iii)  $p > n$  and  $n(n - 1) \cdots (n - p) > 0$  with  $\mathcal{X}$  a bounded interval symmetric about zero, holds.

*Moreover, in the cases (ii) and (iii) the support of the optimal design contains the boundary points of the design space.*

**PROOF.** The proof for condition (i) follows that of Haines (1992). In particular the Kiefer–Wolfowitz equivalence theorem (1960) states that a design  $\xi^*$  is  $D$ -optimal if and only if the directional derivative  $d(x, \xi^*)$  at any point  $x$  in the design space  $\mathcal{X}$  is less than or equal to zero, with equality holding at the support points of  $\xi^*$ . In the present case the directional derivative at  $x$  for any design  $\xi$  is given by

$$(2.9) \quad d(x, \xi) = \frac{f_p^T(x) M^{-1}(\xi) f_p(x)}{(1 + x^2)^n} - (p + 1)$$

and thus has the form

$$\frac{P_{2p}(x)}{(1 + x^2)^n} - (p + 1),$$

where  $P_{2p}(x)$  is a polynomial of degree  $2p$ . A simple calculation shows that

$$\frac{d}{dx}d(x, \xi) = \frac{(1 + x^2)P'_{2p}(x) - 2nxP_{2p}(x)}{(1 + x^2)^{n+1}}$$

and hence that  $d(x, \xi)$  has no more than  $2p + 1$  stationary points. Now for  $n > p$ ,  $\lim_{|x| \rightarrow \infty} d(x, \xi) = -(p + 1)$  and as a consequence the function  $d(x, \xi)$  has at most  $p + 1$  maxima and  $p$  minima. Thus since the directional derivative  $d(x, \xi^*)$  of the  $D$ -optimal design has maxima at the support points it follows that the optimal design has at most  $p + 1$  such points and furthermore, since the information matrix  $M(\xi^*)$  is necessarily nonsingular, that the design  $\xi^*$  is in fact based on exactly  $p + 1$  points of support.

For the proof under assumption (ii) consider the necessary and sufficient condition for  $\xi^*$  to be  $D$ -optimal,  $d(x, \xi^*) \leq 0$ , rewritten as

$$(2.10) \quad g(x) = f_p^T(x)M^{-1}(\xi^*)f_p(x) - (p + 1)(1 + x^2)^n \leq 0$$

for all  $x \in \mathcal{X} = [a, b]$ , where  $-\infty < a < b < \infty$ , with equality holding at the support points of  $\xi^*$ . Then for  $n$  a nonnegative integer strictly less than  $p$ , the function  $g(x)$  is a polynomial of exact degree  $2p$  and it follows immediately that  $g(x)$  has at most  $p + 1$  zeros and further that, if it does indeed have  $p + 1$  zeros, then these include the two endpoints,  $a$  and  $b$ . Thus, since  $M(\xi^*)$  is necessarily nonsingular, it is clear that  $\xi^*$  has exactly  $p + 1$  points of support including the endpoints.

The proof assuming that condition (iii) holds is more intricate. Suppose that  $p$  is odd, the case for  $p$  even being similar. Let  $\mathcal{X} = [-b, b]$ ,  $0 < b < \infty$ . Then the elements of the information matrix of a  $D$ -optimal design  $\xi^*$  must satisfy  $[M(\xi^*)]_{ij} = 0$  for  $i + j$  odd. Hence  $[M^{-1}(\xi^*)]_{ij} = 0$  for  $i + j$  odd, and it follows that the function  $g(x)$  is even, that is,  $g(x) = h(x^2)$ , where  $h(t) = \sum_{i=0}^p a_i t^i - (p + 1)(1 + t)^n$ . Now let  $H(t) = h(t)$  for  $t \geq 0$ , and let  $H(t)$  be equal to the Taylor expansion of  $h(t)$  about 0 up to degree  $p + 1$  for  $t < 0$ . Then  $H(t) \leq 0$  for all  $t \in [0, b^2]$  and the function  $H(t)$  has as many zeros in  $[0, b^2]$  as the function  $g(x)$  in  $[0, b]$ . Let  $q$  denote the number of zeros of  $H(t)$  in  $\mathbb{R}$  counted according to their multiplicities. Then the assumption that  $n(n - 1) \cdots (n - p) > 0$  ensures that  $H^{(p+1)}(t) < 0$  for all  $t \in \mathbb{R}$  and it therefore follows from Rolle's theorem that  $q \leq p + 1$ . Furthermore, for  $t < 0$   $H(t)$  is equal to a polynomial of even degree  $p + 1$  and the leading coefficient of this polynomial is negative. Thus  $\lim_{t \rightarrow -\infty} H(t) = -\infty$ . On the other hand,  $\lim_{t \rightarrow \infty} H(t) = \infty$ , since  $n < p$ . Hence  $q$  must be odd, and so  $q \leq p$ . Since every zero of  $H(t)$  in  $(0, b^2)$  is at least a double zero,  $H(t)$  vanishes at most  $(p - 1)/2$  times in  $(0, b^2)$  and at most  $(p + 1)/2$  times in  $[0, b^2]$ . Therefore,  $g(x)$  has at most  $p - 1$  distinct zeros in  $(-b, 0) \cup (0, b)$  and at most  $p + 1$  distinct zeros in  $[-b, b]$ . Thus  $g(x)$  has exactly  $p + 1$  zeros in  $[-b, b]$  and the zeros must include the endpoints.  $\square$

Thus  $D$ -optimal designs for the weighted polynomial model (2.1) with efficiency function (2.2) and  $p < n$  are based on exactly  $p + 1$  points. For

$p = n$ , the problem of determining the number of support points for the  $D$ -optimal designs is inextricably associated with the nature of those points and is thus discussed in detail in the ensuing section. In particular it is shown there that for  $p = n$  there exists an infinite number of  $D$ -optimal designs with  $p + 1$  and with more than  $p + 1$  points of support. If  $p > n$  and  $n \notin \mathbb{N}_0$ , neither of the assumptions (ii) or (iii) of Lemma 2.1 are necessarily satisfied, but the following result shows that in such cases the  $D$ -optimal designs are based on at most  $p + 2$  support points.

**LEMMA 2.2.** *If  $p > n$ , the  $D$ -optimal design for the weighted polynomial regression model (2.1) with efficiency function (2.2) and design space  $\mathcal{X}$  a bounded interval has at most  $p + 2$  support points.*

**PROOF.** For  $p > n$  it is clear from the proof of Lemma 2.1 that the directional derivative  $d(x, \xi^*)$  has no more than  $2p + 1$  stationary points and also that  $\lim_{|x| \rightarrow \infty} d(x, \xi^*) = \infty$ . Thus because there is equality in (2.10) for all  $p_0 \geq p + 1$  support points of the  $D$ -optimal design,  $d(x, \xi^*)$  has either  $p - 1$  or  $p$  zeros in the interval  $(a, b)$ . The first case corresponds to  $p + 1$  and the second to  $p + 2$  points of support in  $[a, b]$ .  $\square$

The following example illustrates the fact that  $D$ -optimal designs based on  $p + 2$  points do exist but that condition (iii) of Lemma 2.1 is not a necessary one.

**EXAMPLE 2.3.** Consider the polynomial regression model (2.1) with efficiency function (2.2) and design space  $\mathcal{X} = [-1, 1]$ . Suppose further that the regression is quadratic, that is,  $p = 2$ , and consider values of  $n < p$  for which condition (iii) of Lemma 2.1 is violated. Then for  $n = -2$ , the  $D$ -optimal design comprises four points of support and, specifically, puts masses of 0.1675 on each of the points  $\pm 0.1895$  and of 0.3325 on each of the points  $\pm 1$ . In contrast, the  $D$ -optimal design for  $n = \frac{3}{2}$  puts equal masses on the three support points,  $-1, 0$  and  $1$ . In fact empirical evidence suggests that there is a cut-off value for  $n$  satisfying the equation  $2^{n-1} = n + 2$ , that is,  $n \approx -1.8625$ , below which the requisite  $D$ -optimal design is based on four points and above which it is based on three points of support.

**3. Unrestricted design spaces.** Consider the weighted polynomial regression model (2.1) with efficiency function (2.2) and unrestricted design space  $\mathbb{R}$ . Then for  $p \leq n$ ,  $D$ -optimal designs exist and can be determined explicitly. This is explored in the following series of results.

**THEOREM 3.1.** *Let  $p < n$  and  $\mathcal{X} = \mathbb{R}$ . Then the  $D$ -optimal design for the polynomial regression model (2.1) with efficiency function (2.2) puts equal masses at the  $p + 1$  zeros of the polynomial*

$$C_{p+1}^{(-n-1/2)}(\sqrt{-x^2}),$$

where  $C_m^{(\alpha)}(u)$  denotes the  $m$ th ultraspherical polynomial defined by the recursive relations  $C_{-1}^{(\alpha)}(u) = 0$ ,  $C_0^{(\alpha)}(u) = 1$  and

$$(3.1) \quad mC_m^{(\alpha)}(u) = 2(m + \alpha - 1)uC_{m-1}^{(\alpha)}(u) - (m + 2\alpha - 2)C_{m-2}^{(\alpha)}(u)$$

for  $m \geq 1$ .

PROOF. Observe first that by Lemma 2.1 the required  $D$ -optimal design has exactly  $p + 1$  support points, say  $-\infty < x_0 < \dots < x_p < \infty$ . Then it follows from a standard argument in design theory [see, e.g., Karlin and Studden (1966), page 329; Silvey (1980), page 43] that the  $D$ -optimal design  $\xi^*$  puts equal masses at these points and that the determinant of the information matrix  $M(\xi^*)$  is given by

$$(3.2) \quad |M(\xi^*)| = (p + 1)^{-(p+1)} \prod_{j=0}^p (1 + x_j^2)^{-n} \prod_{0 \leq i < j \leq p} (x_i - x_j)^2.$$

Furthermore by invoking arguments similar to those presented in Szegö (1975), page 141, it is clear that the polynomial  $k(x) = \prod_{j=0}^p (x - x_j)$  which has roots equal to the support points of the  $D$ -optimal design  $\xi^*$  satisfies the differential equation

$$(3.3) \quad (1 + x^2)k''(x) - 2nxk'(x) + c_0k(x) = 0,$$

where the term  $c_0 = (p + 1)/(2n - p)$  is obtained by comparing the coefficients of  $x^p$  on both sides of the equation. For ease of reference the polynomial  $k(x)$  is termed the support polynomial.

The differential equation (3.3) is second order and has a fundamental set of solutions comprising the two hypergeometric functions  ${}_2F_1(-p/2, -n + p/2, 1/2; -x^2)$  and  $x {}_2F_1(-p/2, -n + (p + 1)/2, 3/2; -x^2)$  [see Boyce and DiPrima (1997), page 135]. For  $p$  a nonnegative integer, it now follows that exactly one of these solutions is a polynomial of degree  $p + 1$  and, moreover, that it coincides with the ultraspherical polynomial  $C_{p+1}^{(-n-1/2)}(\sqrt{-x^2})$  of degree  $p + 1$  [see Szegö (1975), equations (4.7.29) and (4.7.30)]. This result is further corroborated by comparing the differential equation (3.3) with equation (4.7.5) of Szegö (1975).  $\square$

The polynomial  $C_{p+1}^{(-n-1/2)}(\sqrt{-x^2})$ , with superscripted parameter  $\alpha < -\frac{1}{2}$  and an imaginary argument, is a generalized form of an ultraspherical polynomial and as a consequence enjoys some, but not necessarily all, of the properties of its classical counterparts [Szegö (1975), page 63]. Thus, for example, the Rodrigues' formula for ultraspherical polynomials as given in Abramowitz and Stegun [(1964), page 785, (22.11.2)] holds and, specifically, the polynomial  $C_{p+1}^{(-n-1/2)}(\sqrt{-x^2})$  is proportional to

$$(1 + x^2)^{n+1} \frac{d^{p+1}}{dx^{p+1}} \frac{1}{(1 + x^2)^{n-p}}.$$

Note that this result was also obtained in the context of optimal design by Antille (1977). In contrast, the orthogonality properties of the classical ultraspherical polynomials and also results pertaining to the location of their zeros [see, e.g., Szegő (1975), page 145] do not extend to the polynomials of the form  $C_{p+1}^{(-n-1/2)}(\sqrt{-x^2})$ . It is interesting to note however that the family of polynomials of the form  $C_{p+1}^{(-n-1/2)}(\sqrt{-x^2})$  with  $n$  fixed and  $p < n$  are orthogonal with respect to the weight function  $1/(1+x^2)^{n+1}$  on the interval  $(-\infty, \infty)$ , a property which is unrelated to the orthogonality property of the classical ultraspherical polynomials. Finally it is observed, as an unexpected byproduct of Theorem 3.1, that the roots of the ultraspherical polynomials of the form  $C_{p+1}^{(-n-1/2)}(u)$  are all imaginary.

The following result gives the limiting form of the  $D$ -optimal designs derived in Theorem 3.1 for the case of  $p = \lfloor cn \rfloor$  with  $0 < c < 1$  as  $n$  tends to infinity.

**THEOREM 3.2.** *The  $D$ -optimal design for the weighted polynomial regression model (2.1) with efficiency function (2.2), design space  $\mathbb{R}$  and  $p = p_n = \lfloor cn \rfloor$ , where  $0 < c < 1$ , converges weakly to the distribution with density*

$$(3.4) \quad g(x) = \frac{1}{c\pi} \frac{\sqrt{c(2-c) - (1-c)^2 x^2}}{1+x^2} I\left\{|x| \leq \frac{\sqrt{c(2-c)}}{1-c}\right\}$$

as  $n \rightarrow \infty$ .

**PROOF.** Recurrence formulas for the moments of the uniform distribution on the zeros of the ultraspherical polynomial  $C_{p_n+1}^{(-n-1/2)}(\sqrt{-x^2})$  follow immediately from the results of Dette and Wong [(1995), Theorem 3.1(b)]. In particular, let  $\mu_{k,n}$  denote the  $k$ th moment of this distribution for given  $n$  and let  $p_n = \lfloor cn \rfloor$ . Then, for  $r = 1, 2, \dots$ , the odd moments satisfy  $\mu_{2r-1,n} = 0$  and the even moments can be expressed as

$$(3.5) \quad \mu_{2r,n} = \frac{1}{2n - 2p_n - 1 + 2r} \left\{ (p_n + 1) \sum_{j=0}^{r-1} \mu_{2r-2-2j,n} \mu_{2j,n} + (p_n + 1) \sum_{j=1}^{r-1} \mu_{2r-2j,n} \mu_{2j,n} - (2r-1) \mu_{2r-2,n} \right\}.$$

Thus, in the limit as  $n$  tends to infinity, the ratio  $p_n/n$  tends to  $c$  and the limiting moments  $\mu_k = \lim_{n \rightarrow \infty} \mu_{k,n}$  are described by  $\mu_{2r-1} = 0$  and by the recursive relation

$$(3.6) \quad \mu_{2r} = \frac{c}{2(1-c)} \left\{ \sum_{j=0}^{r-1} \mu_{2r-2j-2} \mu_{2j} + \sum_{j=1}^{r-1} \mu_{2r-2j} \mu_{2j} \right\}$$



for  $r = 1, 2, \dots$ . It now follows that the Stieltjes transform of the limiting distribution  $\xi$ , defined by

$$\begin{aligned}
 (3.7) \quad S(z) &= \int_{\mathbb{R}} \frac{d\xi(x)}{z-x} \\
 &= \sum_{r=0}^{\infty} \frac{\mu_{2r}}{z^{2r+1}} \quad \text{for } z \in \mathbb{C} \setminus \mathbb{R} \quad \text{and} \quad |z| > \frac{\sqrt{c(2-c)}}{1-c},
 \end{aligned}$$

can be readily derived from these moments. More precisely, observe that the identities

$$\begin{aligned}
 \sum_{r=1}^{\infty} \sum_{j=0}^{r-1} \mu_{2r-2-2j} \mu_{2j} z^{-2r+2} &= z^2 S^2(z), \\
 \sum_{r=1}^{\infty} \sum_{j=1}^{r-1} \mu_{2r-2j} \mu_{2j} z^{-2r+2} &= z^4 S^2(z) - 2z^3 S(z) + z^2, \\
 \sum_{r=1}^{\infty} \mu_{2r} z^{-2r+2} &= z^3 S(z) - z^2
 \end{aligned}$$

hold and hence, on multiplying expression (3.6) by  $z^{-2r+2}$ , summing and rearranging, that

$$(1 + z^2)S^2(z) - \frac{2}{c}zS(z) + \frac{2-c}{c} = 0.$$

Thus the Stieltjes transform is given by

$$S(z) = \frac{z + \sqrt{(1-c)^2 z^2 - c(2-c)}}{c(1+z^2)},$$

where the branch of the square root is defined such that  $\lim_{|z| \rightarrow \infty} S(z) = 0$  and the required limiting distribution is obtained by invoking an inversion formula of Chihara (1978), page 90.  $\square$

It is interesting to observe that for  $c = 1$  the limiting distribution (3.4) is Cauchy. This result suggests that the Cauchy distribution plays a particular role in describing  $D$ -optimal designs for the case of  $p = n$  and this idea is made more precise in the following theorem.

**THEOREM 3.3.** *Consider the weighted polynomial regression model (2.1) with efficiency function (2.2),  $p = n$  and  $\mathcal{X} = \mathbb{R}$ . Then every design which puts equal masses at the support points*

$$(3.8) \quad x_j = \tan\left(-\frac{\pi}{2} + \frac{\pi}{k}j + \alpha\right), \quad j = 0, \dots, k-1,$$

where  $\alpha \in (0, \pi/k)$  and  $k \geq p + 1$ , is  $D$ -optimal. Moreover, the design with continuous support described by the Cauchy density  $g(x) = 1/\{\pi(1+x^2)\}$  is also  $D$ -optimal.

PROOF. A straightforward calculation shows that the information matrix (2.3) of the design  $\xi$  for the weighted polynomial regression model in  $x$  can be expressed as

$$M(\nu) = \int_{-\pi/2}^{\pi/2} r(t)r^T(t) d\nu(t),$$

where  $\nu$  denotes the measure on  $(-\pi/2, \pi/2)$  induced from  $\xi$  by the transformation  $x = \tan t$  and the vector  $r(t)$  is given by

$$(3.9) \quad \begin{aligned} r(t) &= (\cos^p t, \cos^{p-1}t \sin t, \dots, \cos t \sin^{p-1} t, \sin^p t)^T \\ &= (r_0(t), \dots, r_p(t))^T. \end{aligned}$$

Now it follows from Euler's identities, which relate trigonometric functions to complex exponential functions, that

$$(3.10) \quad \begin{aligned} &\text{span}\{r_0(t), \dots, r_p(t)\} \\ &= \begin{cases} \text{span}\{1, \sin 2t, \cos 2t, \dots, \sin pt, \cos pt\}, & \text{if } p \text{ is even,} \\ \text{span}\{\sin t, \cos t, \sin 3t, \cos 3t, \dots, \sin pt \cos pt\}, & \text{if } p \text{ is odd.} \end{cases} \end{aligned}$$

Thus, since  $D$ -optimal designs are invariant with respect to the choice of the basis spanned by the regression functions, it is clear that the  $D$ -optimal designs for the model of interest are equivalent to those for homoscedastic trigonometric regression involving the functions (3.10). Specifically, the  $D$ -optimal designs associated with these trigonometric regression models comprise  $k \geq p + 1$  equally spaced and equally weighted points on the interval  $(-\pi/2, \pi/2)$  [see, e.g., Pukelsheim (1993), Section 9.16] and the required designs in terms of the  $x$  variables are obtained by invoking the transformation  $x = \tan t$ .

The information matrix of any  $D$ -optimal design associated with the trigonometric regression functions of the form (3.10) is given by

$$\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2}I_p \end{pmatrix}$$

for  $p$  even and by  $\frac{1}{2}I_{p+1}$  for  $p$  odd, where  $I_d$  denotes the identity matrix of order  $d$ , and it is easily shown that the information matrix for a design with continuous support described by the uniform distribution on the interval  $(-\pi/2, \pi/2)$  is the same [Fedorov (1972), page 94–95]. Thus the second assertion of the theorem follows by again invoking the transformation  $x = \tan t$ .  $\square$

It should be noted that all the curious and interesting features of  $D$ -optimal designs for trigonometric regression hold in the present case and, in particular, that the directional derivatives are equal to zero for all  $x$  in the design space  $\mathbb{R}$ . The following example illustrates some of the above findings.

EXAMPLE 3.4. Consider a rational model with deterministic component

$$(3.11) \quad \eta(u, \theta) = \frac{\theta_0 + \theta_1 u + \theta_2 u^2}{1 + \theta_3 u + \theta_4 u^2}, \quad u \in \mathbb{R},$$

where  $\theta_3^2 < 4\theta_4$ . Then the problem of constructing a locally  $D$ -optimal design for this model is equivalent to that of finding a  $D$ -optimal design for the weighted polynomial regression model (2.1) of degree 4 with efficiency function  $\lambda(x) = (1 + x^2)^{-4}$  and  $\mathcal{X} = \mathbb{R}$ . It now follows immediately from Theorem 3.3 that there exists an infinite number of  $D$ -optimal designs for this weighted polynomial regression model and, by applying the transformation (2.8), that there exists an infinite number of locally  $D$ -optimal designs for the rational model (3.11). Thus, for example, a symmetric five-point  $D$ -optimal design for the weighted polynomial regression model puts equal masses at

$$\left\{ -\tan \frac{2\pi}{5}, -\tan \frac{\pi}{5}, 0, \tan \frac{\pi}{5}, \tan \frac{2\pi}{5} \right\} = \{-3.078, -0.727, 0, 0.727, 3.078\}$$

and the  $D$ -optimal design with infinite support is described by the Cauchy distribution with density  $1/(\pi(1 + x^2))$  for  $x \in \mathbb{R}$ . For the rational model (3.11) with  $\theta_3 = 1$  and  $\theta_4 = 1$ , the corresponding five-point locally  $D$ -optimal design puts equal masses at the support point  $-3.165, -1.129, -\frac{1}{2}, 0.129$  and  $2.165$  and the  $D$ -optimal design with infinite support is given by the Cauchy distribution with density function

$$g(u) = \frac{2}{\sqrt{3}\pi} \frac{1}{\{1 + (2u + 1)^2/3\}} \quad \text{for } u \in \mathbb{R}.$$

REMARK 3.5. For the weighted polynomial model (2.1) with efficiency function (2.2) and  $p = n$ , the  $D$ -optimal designs based on  $p + 1$  points and given in Theorem 3.3 can also be derived by invoking arguments related to those in Theorem 3.1. To be precise, for  $p = n$  the fundamental set of solutions to the differential equation (3.3) comprises the two ultraspherical polynomials  $C_p^{(-p-1/2)}(\sqrt{-x^2})$  and  $C_{p+1}^{(-p-1/2)}(\sqrt{-x^2})$  of degrees  $p$  and  $p + 1$ , respectively (see the discussion in the proof of Theorem 3.1]. It thus follows that every linear combination,

$$(3.12) \quad k_c(x) = C_{p+1}^{(-p-1/2)}(ix) + icC_p^{(-p-1/2)}(ix) \quad \text{for } c \in \mathbb{R}$$

satisfies the differential equation (3.3) and furthermore, since such a combination is a polynomial of degree  $p + 1$ , that its zeros yield the support points of the required  $D$ -optimal designs. Now observe that for  $x = \tan t$ ,

$$\begin{aligned} & C_{p+1}^{(-p-1/2)}(i \tan t) \\ &= C_{p+1}^{(-p-1/2)}\left(i \cot\left(\frac{\pi}{2} - t\right)\right) \end{aligned}$$

$$\begin{aligned}
&= \frac{(2i)^{p+1} \Gamma(\frac{1}{2})}{\Gamma(p+2) \Gamma(-p-\frac{1}{2})} \sum_{m=0}^{\lfloor (p+1)/2 \rfloor} (-1)^m \binom{p+1}{2m} \\
&\quad \times \left( \cos\left(\frac{\pi}{2} - t\right) \right)^{p+1-2m} \left( \sin\left(\frac{\pi}{2} - t\right) \right)^{2m-p-1} \\
&= \left( \frac{2i}{\cos t} \right)^{p+1} \frac{\Gamma(\frac{1}{2})}{\Gamma(p+2) \Gamma(-p-\frac{1}{2})} \cos\left((p+1)\left(\frac{\pi}{2} - t\right)\right),
\end{aligned}$$

where the second equality follows from the explicit representation of the ultraspherical polynomials [see, e.g., Abramowitz and Stegun (1964), page 775 (22.3.4)] and the third equality is a consequence of certain basic properties of trigonometric functions [see, e.g., Jolley (1961), (643)]. Similarly

$$C_p^{(-p-1/2)}(i \tan t) = \frac{(2i)^p}{(\cos t)^{p+1}} \frac{\Gamma(-\frac{1}{2})}{\Gamma(p+2) \Gamma(-p-\frac{1}{2})} \sin\left((p+1)\left(\frac{\pi}{2} - t\right)\right).$$

Thus finding the zeros of the polynomial (3.12) is equivalent, through the transformation  $x = \tan t$ , to solving the trigonometric equation

$$(3.13) \quad \tan\left((p+1)\left(\frac{\pi}{2} - t\right)\right) = c^*,$$

where  $t \in (-\pi/2, \pi/2)$  and  $c^* \in \mathbb{R}$ . Furthermore it is readily seen that (3.13) has exactly  $p+1$  solutions equally spaced in the interval  $(-\pi/2, \pi/2)$ , and thus that these solutions, transformed according to the relation  $x = \tan t$ , are precisely the support points of the  $(p+1)$ -point  $D$ -optimal designs described in Theorem 3.3.

**4. Optimal designs on restricted design spaces.** Consider the weighted polynomial regression model (2.1) with efficiency function (2.2) and constrained design space  $\mathcal{X} = [-1, 1]$ . Assume also that one of the conditions of Lemma 2.1 is satisfied and hence that the  $D$ -optimal designs for this model setting are supported on exactly  $p+1$  points. Then the first result presented in this section provides a partial solution to the problem of constructing such designs for the particular case of  $p < n$ .

**THEOREM 4.1.** *Let  $x_0(p, n)$  denote the largest root of the ultraspherical polynomial  $C_{p+1}^{(-n-1/2)}(\sqrt{-x^2})$ . Then if  $x_0(p, n) \leq 1$ , the  $D$ -optimal design for the polynomial regression model (2.1) with efficiency function (2.2),  $p < n$  where  $n \in \mathbb{R}$ , and  $\mathcal{X} = [-1, 1]$  puts equal masses on the zeros of that polynomial. Moreover, a sufficient condition for  $x_0(p, n) \leq 1$  is given by*

$$n \geq \frac{1}{2}(-2 + 3p + \sqrt{1 + 3p^2}).$$

**PROOF.** The first part of the theorem follows immediately from Theorem 3.1. To prove the second part, observe that the recurrence relation for the

monic version of an ultraspherical polynomial (i.e., the form with leading coefficient equal to 1) is given from the results in the Appendix of Chihara (1978) by  $\hat{C}_{-1}^{(-n-1/2)}(u) = 0$ ,  $\hat{C}_0^{(-n-1/2)}(u) = 1$  and

$$\hat{C}_j^{(-n-1/2)}(u) = x\hat{C}_{j-1}^{(n-1/2)}(u) - \lambda_j\hat{C}_{j-2}^{(-n-1/2)}(u) \quad \text{for } j \geq 1,$$

where

$$\lambda_j = \frac{(j-1)(j-3-2n)}{4(j-n-5/2)(j-n-3/2)}.$$

Moreover by Proposition 10.6 of Lorentzen and Waadeland (1991), page 461, all zeros of the polynomial  $\hat{C}_{p+1}^{(-n-1/2)}(u)$  are located within the disc  $\{z: |z| < 2M\}$  where  $M = \max_{j=2}^{p+1} |\lambda_j|$ . Now it is readily shown that  $|\lambda_j|$  is an increasing function of  $j$  for  $1 \leq j \leq p+1$  with  $p < n$  and hence that

$$M = |\lambda_{p+1}| = \frac{p(2n+2-p)}{(2n-2p+3)(2n-2p+1)}.$$

Thus  $x_0(p, n) \leq 1$  provided  $M \leq 1/2$  and thus provided  $n \geq \frac{1}{2}(-2 + 3p + \sqrt{1 + 3p^2})$ .  $\square$

Furthermore if  $p < n$  and  $x_0(p, n) > 1$  the support of the  $D$ -optimal design must include the boundary points  $-1$  and  $+1$  since, if this were not the case, a second  $D$ -optimal design on the design space  $\mathcal{X} = \mathbb{R}$  with  $p+1$  support points would be obtained and Theorem 3.1 contradicted. In addition it is clear from Lemma 2.2 that for  $p > n$  the required  $D$ -optimal design contains the end-points of the design space in its support.

For  $p = 0, 1, \dots, 4$  explicit expressions for the support points of  $D$ -optimal designs based on  $p+1$  points can be readily obtained by using the above proposition and by invoking symmetry arguments, and this is illustrated in the following example.

EXAMPLE 4.2. Suppose that  $p = 4$ . Then the largest root of the ultraspherical polynomial  $C_{p+1}^{(-n-1/2)}(\sqrt{-x^2})$  is given by

$$x_0(4, n) = \left\{ \frac{1}{2n-7} \left( 5 + 2\sqrt{5} \sqrt{\frac{n-1}{2n-5}} \right) \right\}^{1/2}$$

and is strictly decreasing with increasing  $n$ . It thus follows that  $x_0(4, n) \leq 1$  for  $n \geq \frac{1}{2}(11 + \sqrt{21})$  and hence that for such values of  $n$  the support points of the  $D$ -optimal designs on  $[-1, 1]$  coincide with the zeros of the polynomial  $C_5^{(-n-1/2)}(\sqrt{-x^2})$ . The bound from Theorem 4.1 is given by  $n \geq 17/2$ . For values of  $n < \frac{1}{2}(11 + \sqrt{21})$  with  $n(n-1)(n-2)(n-3)(n-4) \geq 0$ , condi-

tion (iii) of Lemma 2.1 holds and the  $D$ -optimal design on  $[-1, 1]$  is symmetric and puts equal masses on the support points  $0, \pm a, \pm 1$ , where  $0 < a < 1$  and  $a$  is obtained algebraically by maximizing the determinant of the information matrix as

$$a = \left\{ \frac{(n+2) - \sqrt{n^2 - 2n + 25}}{2n - 7} \right\}^{1/2}.$$

More generally  $D$ -optimal designs for the weighted polynomial regression model (2.1) with efficiency function (2.2) and design space  $\mathcal{X} = [-1, 1]$  which are based on  $p+1$  points of support can be obtained as follows. The determinant of the information matrix of a design with  $x_0 = -1$ , interior points  $-1 < x_1 < x_2 < \dots < x_{p-1} < 1$ , and  $x_p = 1$  is given by

$$(4.1) \quad |M(\xi)| = 2^{2-2n} (p+1)^{-(p+1)} \prod_{j=1}^{p-1} \frac{(1-x_j^2)^2}{(1+x_j^2)^n} \prod_{1 \leq i < j \leq p-1} (x_i - x_j)^2.$$

On differentiating this expression with respect to  $x_i$ ,  $i = 1, \dots, p-1$ , and invoking arguments similar to those of Szegő (1975), page 141, it is immediately clear that the support polynomial for the interior points  $k(x) = \prod_{j=1}^{p-1} (x - x_j)$  is a solution to the second-order differential equation

$$(4.2) \quad (1-x^4)k''(x) + 2x\{(n-2)x^2 - (n+2)\}k'(x) + (c_1x^2 + c_2)k(x) = 0,$$

where the terms  $c_1$  and  $c_2$  are constants with  $c_1 = (p-1)(p+2-2n)$ . This equation can in turn be written in Sturm-Liouville form as

$$\frac{d}{dx} \left[ \frac{(1-x^2)^2}{(1+x^2)^n} \frac{d}{dx} k(x) \right] + \frac{1-x^2}{(1+x^2)^{n+1}} (c_2 + c_1x^2) k(x) = 0$$

[see Birkhoff and Rota (1989), page 256]. It then follows immediately from arguments similar to those of Huang, Chang and Wong (1995) and Chang and Lin (1997) that the support polynomial has the form

$$(4.3) \quad k(x) = \begin{cases} \sum_{j=0}^{(p-1)/2} a_{2j} x^{2j}, & \text{for } p \text{ odd,} \\ \sum_{j=0}^{p/2-1} a_{2j+1} x^{2j+1}, & \text{for } p \text{ even} \end{cases}$$

with coefficients given by the elements of the eigenvector corresponding to the smallest eigenvalue satisfying the equation

$$(4.4) \quad Aa = (-c_2)a,$$

where

$$(4.5) \quad a = \begin{cases} (a_0, a_2, \dots, a_{p-1})^T, & \text{for } p \text{ odd,} \\ (a_1, a_3, \dots, a_{p-1})^T, & \text{for } p \text{ even,} \end{cases}$$

$$(4.6) \quad A = \begin{bmatrix} \beta_0 & \gamma_0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \alpha_1 & \beta_1 & \gamma_1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \alpha_2 & \beta_2 & \gamma_2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \alpha_{q-1} & \beta_{q-1} & \gamma_{q-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & \alpha_q & \beta_q \end{bmatrix},$$

$$(4.7) \quad \alpha_j = \begin{cases} (p - 2j + 1)(p - 2n + 2j), & \text{for } p \text{ odd,} \\ (p - 2j)(p + 1 - 2n + 2j), & \text{for } p \text{ even,} \end{cases}$$

$$(4.8) \quad \beta_j = \begin{cases} -4(n + 2)j, & \text{for } p \text{ odd,} \\ -2(n + 2)(2j + 1), & \text{for } p \text{ even,} \end{cases}$$

$$(4.9) \quad \gamma_j = \begin{cases} (2j + 1)(2j + 2), & \text{for } p \text{ odd,} \\ (2j + 2)(2j + 3), & \text{for } p \text{ even,} \end{cases}$$

where  $q = [(p - 1)/2]$ . The above formulation provides a complete, albeit numeric, solution to the problem of finding  $(p + 1)$ -point  $D$ -optimal designs for the weighted polynomial regression model (2.1) with efficiency function (2.2) on the design space  $[-1, 1]$ . Furthermore, it is readily extended to any design space which is a symmetric, compact interval. Analytic solutions to this type of optimal design problem are, however, elusive. It is therefore particularly interesting that for the present model setting with  $n = 1$  and  $n = 2$  the eigenvalue problem (4.4) can be solved explicitly. The remainder of this section is devoted to these special cases.

Observe firstly that by Descartes' rule of signs all coefficients of the support polynomial defined in (4.3), and hence all elements of the corresponding eigenvector of the matrix  $A$ , are nonzero and must alternate in sign. The following result shows that for any  $n$  and  $p$  sufficiently large there is in fact exactly one eigenvector of  $A$  with this property.

LEMMA 4.3. *Assume that*

$$p > \begin{cases} 2(n - 1), & \text{if } p = 2q + 1 \text{ is odd,} \\ 2(n - 1) - 1, & \text{if } p = 2q + 2 \text{ is even.} \end{cases}$$

Then there exists exactly one eigenvector of the matrix  $A$  defined in (4.5) with components which have precisely  $q$  changes of sign. Furthermore, this eigenvector corresponds to the smallest eigenvalue of  $A$ .

PROOF. Recall the definition of the matrix  $A$  in (4.5), (4.7)–(4.9) and choose  $\delta \in \mathbb{R}$  sufficiently large so that every principal minor of the matrix  $B = \delta I + A$  is positive. It then follows from this choice of  $\delta$  and from the fact that  $\alpha_1, \dots, \alpha_q$  and  $\gamma_0, \dots, \gamma_{q-1}$  are positive that  $B$  is an oscillatory matrix [see Gantmacher (1959), page 103]. The assertion now follows from Theorem XIII.13 in Gantmacher (1959).  $\square$

For  $n = 1$  and  $2$  with  $p > n$ , the eigenvector of  $A$  with alternating components described in Lemma 4.3, and hence the associated  $D$ -optimal design, can be derived explicitly as follows.

THEOREM 4.4. (a) *The  $D$ -optimal design for the weighted polynomial regression model (2.1) with efficiency function  $\lambda(x) = (1 + x^2)^{-1}$  and  $p > 1$  puts equal masses  $1/(p + 1)$  at the  $p + 1$  zeros of the polynomial*

$$(4.10) \quad (1 - x^2) \left\{ x C_{p-2}^{(3/2)}(x) + \sqrt{\frac{p(p-1)}{2}} C_{p-1}^{(1/2)}(x) \right\},$$

where  $C_k^{(\alpha)}(x)$  denotes the  $k$ th ultraspherical polynomial defined in (3.1).

(b) *The  $D$ -optimal design for the weighted polynomial regression model (2.1) with efficiency function  $\lambda(x) = (1 + x^2)^{-2}$  and  $p > 2$  puts equal masses  $1/(p + 1)$  at the  $p + 1$  zeros of the polynomial*

$$(4.11) \quad (1 - x^2) \left\{ x^2 C_{p-3}^{(3/2)}(x) + \frac{x}{2} \left( \sqrt{8(p-1)(p-2) + 1} + 1 \right) C_{p-2}^{(1/2)}(x) - \frac{(p-1)(p-2)}{2} C_{p-1}^{(-1/2)}(x) \right\}.$$

PROOF. The proofs for parts (a) and (b) of the theorem are similar and only that for the more complicated part (b) is therefore presented here. It follows from Lemma 2.1 that the  $D$ -optimal design has  $p + 1$  support points including both endpoints. Furthermore, since  $p > 2$ , it follows from Lemma 4.3 that there exists exactly one eigenvector of the matrix  $A$  with precisely  $\lfloor (p-1)/2 \rfloor$  sign changes. For definiteness, assume that  $p$  is even with  $p = 2q + 2$ . Then it is not hard to verify that

$$(4.12) \quad c_2 = 6 + (p-1)(p-2) + 2\sqrt{8(p-1)(p-2) + 1}$$

is a solution to (4.4) and that the corresponding eigenvector  $(a_1^*,$



$\alpha_3^*, \dots, \alpha_{2q+1}^*)^T$  has components which have alternating signs and are given by

$$(4.13) \quad \alpha_{2i+1}^* = (-1)^{i+1} \left\{ 4i + \frac{(p-1)(p-2)}{(2i+1)} + \left[ \sqrt{8(p-1)(p-2)+1} + 1 \right] \right\} \\ \times \binom{q}{i} \prod_{j=1}^i \frac{2q+2j-1}{2j-1} = \alpha_{2i+1}^{*(1)} + \alpha_{2i+1}^{*(2)} + \alpha_{2i+1}^{*(3)}$$

for  $i = 0, \dots, q$ . Note the exact correspondence between the three terms specified algebraically in the two summations. Consequently, it now suffices to show that the polynomial  $(1-x^2) \sum_{i=0}^q \alpha_{2i+1}^* x^{2i+1}$  is proportional to (4.11). To this end observe that

$$(4.14) \quad C_{2q}^{(1/2)}(x) = \frac{1}{2} \sum_{i=0}^q (-1)^{q-i} \frac{\Gamma(q+i+\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(2i+1)\Gamma(q-i+1)} (2x)^{2i} \\ = \frac{\Gamma(q+\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(q+1)} \sum_{i=0}^q (-1)^{q-i} \binom{q}{i} \prod_{j=1}^i \frac{2q+2j-1}{2j-i} x^{2i},$$

where the first equality follows from equation (22.3.4) in Abramowitz and Stegen (1964) and the second equality is obtained by straightforward but tedious computation. This immediately implies that

$$(4.15) \quad \sum_{i=0}^q \alpha_{2i+1}^{*(3)} x^{2i+1} = q_0 \left\{ \sqrt{8(p-1)(p-2)+1} + 1 \right\} x C_{2q}^{(1/2)}(x),$$

where  $q_0 = (-1)^{q+1} \Gamma(\frac{1}{2})\Gamma(q+1)/\Gamma(q+\frac{1}{2})$ . Furthermore, by differentiating and integrating (4.14), and in so doing invoking formula (4.7.14) in Szegő (1975), it follows that

$$(4.16) \quad \sum_{i=0}^q \alpha_{2i+1}^{*(1)} x^{2i+1} = 2q_0 x^2 C_{2q-1}^{(3/2)}(x)$$

and that

$$(4.17) \quad \sum_{i=0}^q \alpha_{2i+1}^{*(2)} x^{2i+1} = -q_0 (p-1)(p-2) C_{2q+1}^{(-1/2)},$$

respectively. Thus the assertion for  $p = 2q + 2$  follows by adding (4.15)–(4.17). The proof for the case of  $p$  odd with  $p = 2q + 1$  is similar, with the eigenvector corresponding to (4.12) being given by  $(\alpha_0^*, \dots, \alpha_{2q}^*)^T$  where

$$\alpha_{2i}^* = (-1)^i \left\{ i \left( 4i - 1 + \sqrt{8(p-1)(p-2)+1} \right) + \frac{(p-1)(p-2)}{2} \right\} \\ \times \binom{q}{i} \prod_{j=1}^i \frac{2q+2j-3}{2j-1}$$

for  $i = 0, \dots, q$ .  $\square$

REMARK 4.5. Hoel (1958) has shown that the  $D$ -optimal design for the homoscedastic polynomial regression model of degree  $p$  puts equal masses at the  $p + 1$  zeros of the polynomial  $(1 - x^2)P'_p(x)$  where  $P_p(x)$  is the  $p$ th Legendre polynomial. Now this model corresponds to the weighted polynomial regression model (2.1) with efficiency function (2.2) and  $n = 0$  and furthermore it is well known that  $P'_p(x)$  is proportional to the ultraspherical polynomial  $C_{p-1}^{(3/2)}(x)$  [see Szegö (1975), page 81]. Thus the result of Hoel (1958) for efficiency function (2.2) with  $n = 0$  and the results of Theorem 4.4 for  $n = 1$  and 2 lead naturally to the following conjecture.

CONJECTURE. *There exist constants  $\delta_0, \dots, \delta_n$  such that the  $D$ -optimal design for the weighted polynomial regression model (2.1) with efficiency function (2.2), where  $n \in \mathbb{N}_0$  with  $n < p$  and  $\mathcal{X} = [-1, 1]$ , puts equal masses at the  $p + 1$  zeros of the polynomial  $(1 - x^2)k(x)$ , where*

$$(4.18) \quad k(x) = \sum_{j=0}^n \delta_j x^j C_{p-1-j}^{(3/2-n+j)}(x).$$

A proof of this general conjecture requires the explicit determination of the smallest eigenvalue of the matrix  $A$  in order to find the coefficients in (4.18) and this would seem to be intractable for all  $n$  greater than 2. There is, however, numerical evidence to suggest that the conjecture is indeed correct. In particular for given  $n$  and  $p$  it is possible to find the smallest eigenvalue of the matrix  $A$  numerically, to substitute the sum of ultraspherical polynomials (4.18) for  $k(x)$  in the differential equation (4.2) and to solve for the coefficients  $\delta_0, \dots, \delta_n$ , either by equating coefficients of powers of  $x$  or by evaluating (4.2) at selected values of  $x$ . In all cases examined, the resultant expression for  $k(x)$  was verified as a solution to the differential equation (4.2) and was further shown to have coefficients with alternating signs. Thus, for  $p$  sufficiently large, it follows immediately from Lemma 4.3 that the design which puts equal masses on support points given by the zeros of the polynomial  $(1 - x^2)k(x)$  is indeed  $D$ -optimal. Otherwise for  $p > n$  the  $D$ -optimality of this latter design can be established by invoking the Kiefer–Wolfowitz equivalence theorem (1960). The following example illustrates these ideas.

EXAMPLE 4.6. Consider the weighted polynomial regression model (2.1) of degree  $p = 9$  with efficiency function (2.2) and  $n = 3$ . Then the differential equation (4.18) is given by

$$(4.19) \quad (1 - x^4)k''(x) + 2x(x^2 - 5)k'(x) + (40x^2 + c_2)k(x) = 0$$

and the term  $c_2$ , which is equal to minus the smallest eigenvalue of the matrix  $A$ , is computed as 113.1363. Substituting the sum of four ultraspherical polynomials defined by (4.18) for  $k(x)$  into this differential equation,

setting  $\delta_3 = 1$  for definiteness, evaluating the resultant expression at  $x = 0.2, 0.4$  and  $0.6$ , and solving for  $\delta_0, \delta_1$  and  $\delta_2$  yields the polynomial

$$\begin{aligned} k(x) &= 31.1023C_8^{(-3/2)}(x) - 70.0175xC_7^{(-1/2)}(x) \\ &\quad + 15.2841x^2C_6^{(1/2)}(x) + x^3C_5^{(3/2)}(x) \\ &= -0.007942 + 0.449276x^2 - 0.346052x^4 + 0.780734x^6 - 0.518347x^8. \end{aligned}$$

It is straightforward to verify by direct substitution that this expression does indeed satisfy the differential equation (4.19) for all  $x$ . Furthermore, since Lemma 4.3 holds for  $n = 3$  and  $p = 9$ , the  $D$ -optimal design is based on 10 points and puts equal masses on the zeros of the support polynomial  $(1 - x^2)k(x)$ , that is, on the points  $\pm 0.1445, \pm 0.4308, \pm 0.6969, \pm 0.9022$  and  $\pm 1$ . It is also interesting to note that for  $n = 3$  and  $p = 5$  and  $6$ , the coefficient  $\delta_3$  can be taken to be 0 and the required support polynomials expressed as  $0.1289C_4^{(-3/2)}(x) - 0.8921xC_3^{(-1/2)}(x) + x^2C_2^{(1/2)}(x)$  and as  $0.6235C_5^{(-3/2)}(x) - 0.4564xC_4^{(-1/2)}(x) + x^2C_3^{(1/2)}(x)$ , respectively, while for  $p = 4$  the coefficients  $\delta_2$  and  $\delta_3$  can both be taken to be 0 and  $k(x)$  is simply proportional to  $0.2616C_3^{(-3/2)}(x) + xC_2^{(-1/2)}(x)$ .

REMARK 4.7. The  $D$ -optimal designs derived in Theorem 4.4 converge weakly to the arcsine distribution on the interval  $[-1, 1]$  as  $p$  approaches infinity. The proof of this result follows closely that of Theorem 3.2 and is therefore only outlined here. In particular by invoking arguments similar to those presented in Dette and Wong (1995) and by observing that

$$(4.20) \quad \lim_{p \rightarrow \infty} \frac{c_2}{(p + 1)^2} = 1 \quad \text{for } n = 1, 2$$

it can be shown that the moments  $\mu_r$  of the required limiting distribution are given by  $\mu_{2r-1} = 0, \mu_2 = 1/2$  and

$$\mu_{2r+2} = \frac{1}{2} \left\{ \sum_{u=0}^{r-1} \mu_{2r-2-2u} \mu_{2u} - \sum_{u=1}^r \mu_{2r+2-2u} \mu_{2u} \right\}$$

for all  $r \geq 1$ . It now follows by simple induction that these moments are precisely the moments of the arcsine distribution on the interval  $[-1, 1]$ , that is,  $\mu_{2r} = 2^{-2r} \binom{2r}{r}$  for  $r \geq 1$ . Finally note that the result also holds for the general case with  $n < p$  provided the  $D$ -optimal design is supported at  $p + 1$  points and provided condition (4.20) relating to the smallest eigenvalue of the matrix  $A$  can be established.

**5. Conclusions.** The main aim of the present study has been to develop a cohesive approach to the problem of constructing  $D$ -optimal designs for the

weighted polynomial regression model with efficiency function of the form  $(1 + x^2)^{-n}$  and thus, equivalently, to the problem of constructing locally  $D$ -optimal designs for inverse quadratic models. The results have been derived using tools from the theory of differential equations and are both powerful and interesting. For the unrestricted design space  $\mathbb{R}$  and  $p \leq n$ , the problem of finding  $D$ -optimal designs can be completely solved analytically. Thus for  $p < n$ , the  $D$ -optimal designs put equal masses on  $p + 1$  points which coincide with the zeros of the ultraspherical polynomial  $C_{p+1}^{(-n-1/2)}(\sqrt{-x^2})$  and also exhibit an attractive limiting form for  $p = \lfloor cn \rfloor$  with  $0 < c < 1$  as  $n$  approaches infinity. The case of  $p = n$  is particularly interesting in that a full gamut of unusual  $D$ -optimal designs based on  $p + 1$  and more than  $p + 1$  points and on a continuous support base can be derived from the simple but elegant insight that these designs are equivalent to  $D$ -optimal designs for certain trigonometric regression models. For the weighted polynomial regression model with restricted design space  $[-1, 1]$ ,  $D$ -optimal designs exist for all values of  $p$  and  $n$  but the problem of deriving analytical results for the construction these designs has proved to be a particularly challenging one. Thus if  $n$  and  $p$  satisfy one of the conditions specified in Lemma 2.1, then the  $D$ -optimal designs are based on  $p + 1$  points of support and can be constructed elegantly and effectively by solving a specific eigenvalue problem. This procedure is nevertheless a numerical one and indeed the required  $D$ -optimal designs can be obtained straightforwardly by using an optimization routine to maximize the determinant of the appropriate information matrix. Thus a key feature of the present study is the development of analytical results for constructing  $D$ -optimal designs for the cases of  $n = 1$  and  $2$  and, in particular, the proof that these designs put equal masses on the  $p + 1$  points of support given by  $\pm 1$  and the  $p - 1$  zeros of a sum of  $n + 1$  ultraspherical polynomials.

A number of open problems emerge immediately and naturally from the present study. Thus, for example, it would be worthwhile to attempt to sharpen the conditions derived in Lemma 2.1 concerning the number of support points for the  $D$ -optimal designs of interest with  $p > n$  and in addition, and in a related context, to derive explicit expressions for  $D$ -optimal designs based on  $p + 2$  points of support. Both of these problems are, however, difficult ones to solve and this is further demonstrated in a different model setting by the work of Huang, Chang and Wong (1995). The conjecture that the support points of  $D$ -optimal designs for the restricted design space  $[-1, 1]$  and  $p > n$  with  $n \in \mathbb{N}_0$  are based on the zeros of a sum of ultraspherical polynomials is an extremely powerful and far-reaching one and it would clearly be particularly satisfying to find a general proof for it. Such a proof, while seeming to be tantalizingly close, has not as yet been found. In a broader context it would be interesting to develop  $D$ -optimal designs for design spaces which are not symmetric about zero and, more importantly, to extend the current methodology to accommodate rational models other than inverse quadratic polynomials and to derive analytic results for the construc-

tion of Bayesian  $D$ -optimal designs for those models. Such ideas are, however, beyond the scope of the present paper.

**Acknowledgments.** The authors thank Professor W. J. Studden for many helpful discussions and I. Gottschlich for typing parts of this paper with considerable technical expertise.

## REFERENCES

- ABRAMOWITZ, M. and STEGUN, I. A. (1964). *Handbook of Mathematical Functions*. Dover, New York.
- ANTILLE, G. (1977). Plans d'expérience. Mesures et plans d'expériences  $D$ -optimaux pour les polynômes de régression. *C.R. Acad. Sci. Paris* **285** 577–580.
- BIRKHOFF, G., and ROTA, G. C. (1989). *Ordinary Differential Equations*. Wiley, New York.
- BOYCE, W. E. and DiPRIMA, R. C. (1997). *Elementary Differential Equations and Boundary Value Problems*. 6th ed. Wiley, New York.
- CHANG, F.-C. and LIN, G.-C. (1997).  $D$ -optimal designs for weighted polynomial regression. *J. Statist. Plann. Inference* **62**, 317–331.
- CHERNOFF, H. (1953). Locally optimal designs for estimating parameters. *Ann. Math. Statist.* **24**, 586–602.
- CHIHARA, T. S. (1978). *Introduction to Orthogonal Polynomials*. Gordon and Breach, New York.
- COBBY, J. M., CHAPMAN, P. F. and PIKE, D. J. (1986). Design of experiments for estimating inverse quadratic polynomial responses. *Biometrics* **42** 659–664.
- DETTE, H. and STUDDEN, W. J. (1997). *The Theory of Canonical Moments with Applications in Statistics, Probability and Analysis*. Wiley, New York.
- DETTE, H. and WONG, W. K. (1995). Recurrence moment formulas for  $D$ -optimal designs. *Scand. J. Statist.* **22** 505–512.
- FEDOROV, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York.
- GANTMACHER, F. R. (1959). *The Theory of Matrices*. **2**. Chelsea, New York.
- HAINES, L. M. (1992). Optimal design for inverse quadratic polynomials. *South African Statist. J.* **26** 25–41.
- HE, Z., STUDDEN, W. J. and SUN, D. (1996). Optimal designs for rational models. *Ann. Statist.* **24** 2128–2147.
- HOEL, P. G. (1958). Efficiency problems in polynomial estimation. *Ann. Math. Statist.* **29** 1134–1145.
- HUANG, M.-N. L., CHANG, F.-C. and WONG, W. K. (1995).  $D$ -optimal designs for polynomial regression without an intercept. *Statist. Sinica* **5** 441–458.
- IMHOF, L., KRAFFT, O. and SCHAEFER, M. (1998).  $D$ -optimal designs for polynomial regression with weight function  $w(x) = x/(1+x)$ . *Statist. Sinica* **8** 1271–1274.
- JOLLEY, L. B. W. (1961). *Summation of Series*. Dover, New York.
- KARLIN, S. and STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics*. Wiley, New York.
- KIEFER, J. and WOLFOWITZ, J. (1960). The equivalence of two extremum problems. *Canad. J. Math.* **12** 363–366.
- LORENTZEN, L. and WAADELAND, H. (1991). *Continued Fractions with Applications*. North-Holland, New York.
- ORTIZ, I. and RODRIGUEZ, C. (1998).  $D$ -optimal designs for weighted polynomial regression without any initial terms. In *MODA 5. Advances in Model-Oriented Data Analysis and Experimental Design*. (A. C. Atkinson, L. Pronzato and H. P. Wynn, eds.) 67–74.
- PUKELSHEIM, F. (1993). *Optimal Design of Experiments*. Wiley, New York.
- SILVEY, S. D. (1980). *Optimal Design*. Chapman and Hall, London.

- STUDDEN, W. J. (1980).  $D_s$ -optimal design for polynomial regression using continued fractions. *Ann. Statist.* **8** 1132–1141.
- SZEGÖ, G. (1975). *Orthogonal Polynomials*. Amer. Math. Soc., Providence, RI.

H. DETTE  
RUHR-UNIVERSITÄT BOCHUM  
FAKULTÄT FÜR MATHEMATIK  
44780 BOCHUM  
GERMANY  
E-MAIL: Holger.dette@ruhr-uni-bochum.de

L. M. HALMES  
SCHOOL OF MATHEMATICS, STATISTICS  
AND INFORMATION TECHNOLOGY  
UNIVERSITY OF NATAL  
DIETERMARITZBURG  
SOUTH AFRICA  
E-MAIL: haines@stat.unp.ac.za

L. IMHOF  
RWTH AACHEN  
INSTITUT FÜR STATISTIK  
UND WIRTSCHAFTSMATHEMATIK  
52056 AACHEN  
GERMANY  
E-MAIL: imhof@stochastik.rwth-aachen.de