

# Optimal Approximation in Hilbert Spaces with Reproducing Kernel Functions

By F. M. Larkin

**Abstract.** Characterisations of optimal linear estimation rules are given in terms of the reproducing kernel function of a suitable Hilbert space. The results are illustrated by means of three different, useful function spaces, showing, among other things, how Gaussian quadrature rules, and the Whittaker Cardinal Function, relate to optimal linear estimation rules in particular spaces.

**1. Introduction.** For the purpose of numerically estimating the value of a linear functional  $Lf$ , of some function  $f(x)$ , it is very often convenient to compute an approximation in the form of a linear combination of certain function values. Thus, we might consider that

$$(1) \quad Lf \simeq \sum_{i=1}^n w_i f(x_i)$$

and, by judicious choice of the abscissae  $\{x_j; j = 1, 2, \dots, n\}$  and the weights  $\{w_j; j = 1, 2, \dots, n\}$ , hope to keep the magnitude of the error

$$(2) \quad Rf = Lf - \sum_{i=1}^n w_i f(x_i)$$

tolerably small. A great many approximations of the form of (1) have been suggested in the case when  $L$  represents a process of definite integration; the methods of Chebyshev, Newton-Cotes and Gauss are classic examples, while Sard [7] has derived other formulae.

Taking the problem of numerical quadrature as the typical example of linear estimation, two separate "best" approaches, which might be associated with the names of Gauss and Sard respectively, are discernible. In the Gaussian approach one considers a set of basis functions, for example  $\{x^j; j = 0, 1, 2, \dots\}$ , and tries to choose the  $2n$  parameters  $\{w_j, x_j; j = 1, 2, \dots, n\}$  so as to make approximation (1) exact for the first  $2n$  members of this set. For functions not treated exactly by the resulting approximation process it is a separate problem to discover bounds on the value of  $|Rf|$ .

In the approach of Sard one considers the totality of functions, for example, a real Hilbert space, to which the approximate formulae (1) might be applied; for any  $f$  in this class one derives a bound on  $|Rf|$  in algebraic form, which is then minimised by appropriate choice of the parameters  $\{w_j$  and/or  $x_j; j = 1, 2, \dots, n\}$ . The resulting approximation formula is said to be "optimal"; with respect to weights and/or abscissae.

---

Received July 31, 1969, revised February 19, 1970.

AMS 1969 subject classifications. Primary 4160, 4140, 4144, 6520; Secondary 4630, 4615.

Key words and phrases. Optimal approximation, reproducing kernel functions, Gaussian quadrature, optimal quadrature, Szegő kernel, Paley-Wiener-Hilbert space, Whittaker cardinal function.

Copyright © 1971, American Mathematical Society

The purpose of this paper is to examine the relationship between the two above approaches to the problem of approximating bounded linear functionals in Hilbert spaces possessing reproducing kernel functions. Questions such as "Which functions are treated exactly by an approximation formula derived by means of Sard's approach?" and "Is there a space of functions in which a Gaussian formula is optimal?" will be of interest. Also, simultaneous equations for the optimal values of  $\{w_j, x_j, j = 1, 2, \dots, n\}$  will be derived.

**2. Some General Results.** Let  $D$  denote a point set contained in the space of a real, or complex, variable, and let  $\mathcal{H}$  be a Hilbert space of functions  $f(x)$ , the domain of each  $f \in \mathcal{H}$  being  $D$ . If there exists a function  $K(x, \bar{y})$ , of two variables  $x, y \in D$ , which satisfies the inner product relation

$$(3) \quad h(y) = (h(x), K(x, \bar{y})), \quad \text{for all } h \in \mathcal{H}$$

and also, for any fixed  $y \in D$ ,  $K(x, \bar{y}) \in \mathcal{H}$  is regarded as a function of  $x$ , then  $K$  is said to be a *reproducing kernel function* for  $\mathcal{H}$ . Here, the bar denotes "complex conjugate".

Not all Hilbert spaces possess reproducing kernel functions; those which do are characterised by the following property (Aronszajn, [1]):—A necessary and sufficient condition that  $\mathcal{H}$  possess a reproducing kernel function is that, for every fixed  $x \in D$ , the linear functional  $Lf = f(x)$  is bounded. That is, there exists a finite constant  $C_x$ , depending upon  $x$ , such that

$$|f(x)| \leq C_x \cdot \|f\|, \quad \text{for all } f \in \mathcal{H}.$$

Furthermore, it turns out that the reproducing kernel function, if it exists, is unique and satisfies the relation

$$(4) \quad K(y, \bar{x}) \equiv \overline{K(x, \bar{y})}, \quad \text{for all } x, y \in D.$$

The importance of Hilbert spaces possessing reproducing kernel functions stems from the desirability of estimating a function (and hence, functionals) from values of its ordinates at given abscissae. In order to localise the required function to a bounded region in the space, one can conveniently make use of values of the norm of the function and of bounded, linear functionals (Golomb and Weinberger, [2]); thus, it is important that ordinate values be bounded linear functionals, which implies that the space should possess a reproducing kernel function.

Let  $L$  be a bounded linear functional in  $\mathcal{H}$  with a representer  $g(\cdot)$  and assume that the nodes  $\{x_j; j = 1, 2, \dots, n\}$  always remain within  $D$ . From Eq. (2), applying (3) and the Riesz representation theorem, we see that

$$Rf = \left( f(y), g(y) - \sum_{i=1}^n w_i K(y, \bar{x}_i) \right),$$

hence, by Schwarz's theorem,

$$|R|^2 \leq \|f\|^2 \cdot \left\| g(y) - \sum_{i=1}^n w_i K(y, \bar{x}_i) \right\|^2.$$

By definition, the approximation rule will be optimal if the parameters  $\{w_j, x_j; j = 1, 2, \dots, n\}$  are chosen so as to minimise the quantity

$$(5) \quad S = \left\| g(y) - \sum_{i=1}^n \bar{w}_i K(y, \bar{x}_i) \right\|^2.$$

We now minimise the positive definite quadratic form  $S$  using a standard variational technique. Letting  $w_k$  vary by a small amount  $\delta w_k$ , the corresponding small change in  $S$  is given by

$$(\delta S)_{w_k} = \left. \begin{aligned} & - \left( \delta \bar{w}_k K(y, \bar{x}_k), g(y) - \sum_{i=1}^n \bar{w}_i K(y, \bar{x}_i) \right) \\ & - \left( g(y) - \sum_{i=1}^n \bar{w}_i K(y, \bar{x}_i), \delta \bar{w}_k K(y, \bar{x}_k) \right) \end{aligned} \right\}$$

i.e.,

$$(\delta S)_{w_k} = \left. \begin{aligned} & - \delta \bar{w}_k \left( \overline{g(x_k)} - \sum_{i=1}^n \overline{w_i K(x_k, \bar{x}_i)} \right) \\ & - \delta w_k \left( g(x_k) - \sum_{i=1}^n \bar{w}_i K(x_k, \bar{x}_i) \right) \end{aligned} \right\}.$$

Hence,  $(\delta S)_{w_k}$  vanishes identically for any small  $\delta w_k$ , provided that

$$(6) \quad \sum_{i=1}^n \bar{w}_i K(x_k, \bar{x}_i) = g(x_k).$$

But

$$g(x_k) = (g(y), K(y, \bar{x}_k)) = \overline{(K(y, \bar{x}_k), g(y))} = \overline{L^y K(y, \bar{x}_k)}$$

so that

$$(7) \quad \sum_{i=1}^n w_i K(x_i, \bar{x}_k) = L^y K(y, \bar{x}_k),$$

where the superfix  $y$  indicates that  $L$  operates on  $K(y, \bar{x}_k)$  regarded as a function of its first argument.

A stationary point of  $S$  with respect to small changes in the abscissae may be found in a similar fashion. The small change in  $S$  corresponding to a small change  $\delta x_k$  in  $x_k$  may be expressed as

$$(8) \quad (\delta S)_{x_k} = \left. \begin{aligned} & - \overline{\delta x_k} \bar{w}_k \left( \frac{\partial}{\partial \bar{x}_k} \overline{g(x_k)} - \sum_{i=1}^n w_i \frac{\partial}{\partial \bar{x}_k} K(x_i, \bar{x}_k) \right) \\ & - \delta x_k w_k \left( \frac{\partial}{\partial x_k} g(x_k) - \sum_{i=1}^n w_i \frac{\partial}{\partial x_k} \overline{K(x_i, \bar{x}_k)} \right) \end{aligned} \right\}$$

if the appropriate derivatives exist. This expression vanishes identically for any  $\delta x_k$  provided that

$$(9) \quad \sum_{i=1}^n w_i \frac{\partial}{\partial \bar{x}_k} K(x_i, \bar{x}_k) = L^y \frac{\partial}{\partial \bar{x}_k} K(y, \bar{x}_k) = \overline{\frac{d}{dx_k} g(x_k)}.$$

The matrix of the linear equations (7) is a Gram matrix, and so is nonsingular if the  $\{x_i\}$  are distinct; we have thus proved the following results:

**THEOREM 1.** *If the distinct abscissae  $\{x_i; j = 1, 2, \dots, n\}$  within  $D$  are prescribed,*

the optimal approximation rule of the form (1) is characterised by the property that it treats the functions  $\{K(y, \bar{x}_i); j = 1, 2, \dots, n\}$  exactly.

COROLLARY 1. Given the ordinates  $\{f_i; j = 1, 2, \dots, n\}$  corresponding to distinct abscissae  $\{x_i; j = 1, 2, \dots, n\}$  within  $D$ , the optimal estimate of a bounded linear functional may be found by applying the functional to the function  $\hat{f}(y)$  defined as that unique linear combination of the functions  $\{K(y, \bar{x}_i); j = 1, 2, \dots, n\}$  which interpolates the points  $\{(x_i, f_i); j = 1, 2, \dots, n\}$ .

This result follows from noting that Eqs. (1) and (7) may be expressed in the determinant form

$$\begin{vmatrix} f_1 & f_2 & \cdots & f_n & (\hat{f}(y), g(y)) \\ K(x_1, \bar{x}_1) & K(x_2, \bar{x}_1) & \cdots & K(x_n, \bar{x}_1) & (K(y, \bar{x}_1), g(y)) \\ K(x_1, \bar{x}_2) & K(x_2, \bar{x}_2) & \cdots & K(x_n, \bar{x}_2) & (K(y, \bar{x}_2), g(y)) \\ \vdots & \vdots & & \vdots & \vdots \\ K(x_1, \bar{x}_n) & K(x_2, \bar{x}_n) & \cdots & K(x_n, \bar{x}_n) & (K(y, \bar{x}_n), g(y)) \end{vmatrix} = 0,$$

for any representer  $g(y)$ . Clearly, this relation results from forming an inner product with  $g(y)$  and the relation

$$\begin{vmatrix} f_1 & f_2 & \cdots & f_n & \hat{f}(y) \\ K(x_1, \bar{x}_1) & K(x_2, \bar{x}_1) & \cdots & K(x_n, \bar{x}_1) & K(y, \bar{x}_1) \\ K(x_1, \bar{x}_2) & K(x_2, \bar{x}_2) & \cdots & K(x_n, \bar{x}_2) & K(y, \bar{x}_2) \\ \vdots & \vdots & & \vdots & \vdots \\ K(x_1, \bar{x}_n) & K(x_2, \bar{x}_n) & \cdots & K(x_n, \bar{x}_n) & K(y, \bar{x}_n) \end{vmatrix} = 0,$$

which is the defining relation for  $\hat{f}(y)$ .

COROLLARY 2. The representer  $r(y)$  of the bounded linear functional  $R$  in Eq. (2), vanishes at the distinct abscissae  $\{x_i; j = 1, 2, \dots, n\}$ , within  $D$ , if the weights are chosen optimally.

This follows by inspection of Eq. (6), since

$$(10) \quad r(y) = g(y) - \sum_{i=1}^n \bar{w}_i K(y, \bar{x}_i).$$

THEOREM 2. If the weights  $\{w_k; k = 1, 2, \dots, n\}$  are prescribed, the optimal approximation of the form (1) has the property that it treats the functions,

$$\{(\partial/\partial \bar{x}_k)K(y, \bar{x}_k); k = 1, 2, \dots, n\}$$

exactly, provided that the derivatives and distinct abscissae, within  $D$ , exist.

COROLLARY 3. If the appropriate derivatives and distinct abscissae, within  $D$ , exist, the approximation of the form (1), which is optimal with respect to both weights and abscissae, treats the  $2n$  functions  $\{K(y, \bar{x}_k), (\partial/\partial \bar{x}_k)K(y, \bar{x}_k); k = 1, 2, \dots, n\}$  exactly.

COROLLARY 4. If the appropriate derivatives and abscissae, within  $D$ , exist  $dr(y)/dy$  vanishes at the distinct optimal abscissae  $\{x_k; k = 1, 2, \dots, n\}$ .

This follows by inspection of Eqs. (9) and (10).

Existence of the appropriate derivatives is not a trivial limitation, as the following example illustrates. Consider the class of continuous, real-valued, piecewise differentiable functions of a real variable  $x \in [0, 1]$ , which vanish where  $x = 0$ . This becomes a Hilbert space if an inner product is defined by

$$(f, g) = \int_0^1 f' \cdot g' \cdot dx,$$

and the corresponding reproducing kernel function in the completed space is given by

$$K(x, y) = \min(x, y).$$

Optimality of the approximation rule can still be defined in terms of minimising  $S$  with respect to the  $\{x_j\}$  but can no longer be expressed in terms of Theorem 2.

The reason for this can be interpreted in several equivalent ways; for example, we have

$$\begin{aligned} \frac{\partial}{\partial y} K(x, y) &= 0; & x < y, \\ &\text{undefined;} & x = y, \\ &= 1; & x > y, \end{aligned}$$

which would be the representer of  $f'(x)$  if this were a bounded linear functional on the space. However, the discontinuity at  $x = y$  means that  $(\partial/\partial y)K(x, y) \notin \mathcal{H}$ , corresponding to the fact that  $f'(x)$  is not a bounded functional, and also makes Eq. (9) meaningless, since  $(\partial/\partial \bar{x}_k)K(x_j, \bar{x}_k)$  is not defined for  $j = k$ .

If we now define the matrix  $G$  and vector  $\mathbf{h}$  by means of the relations

$$\begin{aligned} G_{jk} &= K(x_k, \bar{x}_j); & k, j &= 1, 2, \dots, n, \\ h_j &= g(x_j) = \overline{L^v K(y, \bar{x}_j)}; & j &= 1, 2, \dots, n, \end{aligned}$$

the usual result for the minimal value of  $S$  may be expressed as

$$(11) \quad S_{\min} = \|g\|^2 - \mathbf{h}'G^{-1}\mathbf{h},$$

where the prime denotes "complex conjugate transpose". Hence,

$$(12) \quad \frac{|Rf|^2}{\|f\|^2} \leq \|g\|^2 - \mathbf{h}'G^{-1}\mathbf{h} = \|g\|^2 - \mathbf{h}'\mathbf{w},$$

where, of course,  $\mathbf{w}$  is the vector of optimal weight values.

Introducing the matrix  $F$  and vector  $\mathbf{e}$  defined by the relations

$$(13) \quad \begin{aligned} F_{jk} &= \frac{\partial}{\partial \bar{x}_j} K(x_k, \bar{x}_j); & j, k &= 1, 2, \dots, n, \\ e_j &= \overline{L^v \frac{\partial}{\partial \bar{x}_j} K(y, \bar{x}_j)} = \frac{d}{dx_j} g(x_j); & j &= 1, 2, \dots, n, \end{aligned}$$

we see that Eqs. (7) and (9) may be expressed as

$$(14) \quad G\mathbf{w} = \bar{\mathbf{h}}, \quad F\mathbf{w} = \bar{\mathbf{e}}.$$

Eliminating  $\mathbf{w}$  we obtain

$$(15) \quad FG^{-1}\bar{\mathbf{h}} = \bar{\mathbf{e}}$$

as a set of  $n$  simultaneous equations which the optimal values of the abscissae must satisfy. In many cases it may be sufficient to determine these optimal values numerically, either by solving Eq. (15) or by minimising the form

$$S = ||g||^2 - \mathbf{h}'\mathbf{G}^{-1}\mathbf{h}$$

with respect to  $\{x_j; j = 1, 2, \dots, n\}$ . However, for purposes of comparison with known results, we shall proceed by studying certain special cases analytically.

Notice that there is no a priori guarantee that Eq. (15) can be satisfied by distinct values of  $\{x_j; j = 1, 2, \dots, n\}$  within  $D$ . Indeed, if  $Lf$  is chosen to be  $f(x_0), f'(x_0)$ , or any higher derivative, for some fixed  $x_0 \in D$ , the only values satisfying (15) may (not unreasonably) be

$$(16) \quad x_j = x_0; \quad j = 1, 2, \dots, n.$$

However, the above analysis is interesting particularly in connection with approximate quadrature—an application which we explore in the next section.

Richter and Rabinowitz [10] have studied Bergman-Hilbert spaces, with inner product integration regions which are symmetric about the origin in the complex plane. For the case in which  $Lf$  is an integral along a finite segment  $(-d, d)$  of the real line, they have shown that the optimal abscissae are contained in  $[-d, d]$  and the optimal weights are all positive.

**3. Three Examples.** As the first example we take  $\mathcal{H}$  to be the Hilbert space of functions of a complex variable  $x$  which are analytic within the region  $|x| < r$  and continuous where  $x = r$ ,  $r$  being a positive, real number. The inner product will be defined as an integral around the circle, i.e.

$$(17) \quad (f, g) = \int_{|x|=r} fg \cdot |dx|,$$

whence the reproducing kernel function is the Szegő kernel

$$(18) \quad K(y, \bar{x}) = \frac{r}{2\pi} (r^2 - y\bar{x})^{-1}.$$

Further details may be found in Meschkowski [5].

From Corollary 1, we know that any optimal rule of the form (1) for estimating the value of a bounded, linear functional must be exact for the  $2n$  functions:

$$(19) \quad \frac{r}{2\pi} (r^2 - y\bar{x}_k)^{-1}, \quad \frac{ry}{2\pi} (r^2 - y\bar{x}_k)^{-2}; \quad k = 1, 2, \dots, n,$$

and hence, for any linear combination of these.

In general, it will be possible to express the  $2n$  functions

$$(20) \quad q_j(y) = y^j \prod_{k=1}^n \left(1 - \frac{y\bar{x}_k}{r^2}\right)^{-2}; \quad j = 0, 1, 2, \dots, 2n - 1,$$

as linear combinations of the above functions, so that  $\{q_j(y)\}$  must also be treated exactly by the optimal approximation rule. However, as  $r \rightarrow \infty$ ,

$$(21) \quad q_j(y) \rightarrow y^j; \quad j = 0, 1, 2, \dots, 2n - 1,$$

and exact treatment of these limiting functions is the characteristic of quadrature rules of Gaussian type. Thus, for example, the classic Gaussian quadrature rules appear as the limit, when  $r \rightarrow \infty$ , of a sequence of *optimal* quadrature rules.

Although, in principle, the asymptotic form of  $S_{\min}$  may be determined for any given  $L$ , this information does not enable us to make use of (12) for the purpose of bounding the error of approximation, since  $\|f\|$  increases without limit as  $r \rightarrow \infty$ . This observation agrees with the fact that extra information, often in terms of derivatives of  $f$ , is necessary in order to obtain bounds on the error of Gaussian quadrature rules. It is of course, still possible to bound the error in terms of values of  $\|f\|$  and  $S$  computed in a space for which a Gaussian approximation rule is not optimal; for example, one could choose a fixed, finite value of  $r$  for this purpose.

(ii) For the second example we consider the class of real valued functions of a real variable  $x$  which have square integrable second derivatives over the interval  $[0, 1]$ . With an inner product defined by

$$(22) \quad (f, g) = \alpha f(0)g(0) + \beta f'(0)g'(0) + \int_0^1 f''g'' dx,$$

where  $\alpha$  and  $\beta$  are positive, real constants, this class becomes a Hilbert space possessing the reproducing kernel function

$$(23) \quad \begin{aligned} K(x, y) = K(y, x) &= -y^3/6 + y^2x/2 + yx/\beta + 1/\alpha; & y \leq x, \\ &= -x^3/6 + x^2y/2 + xy/\beta + 1/\alpha; & y \geq x, \end{aligned}$$

From the previous theorems we know that an optimal linear approximation rule of the form (1) must be exact for any linear combination of the  $2n$  functions  $\{K(y, x_j), J(y, x_j); j = 1, 2, \dots, n\}$ , where

$$(24) \quad \begin{aligned} K(y, x_j) &= -y^3/6 + y^2x_j/2 + yx_j/\beta + 1/\alpha; & y \leq x_j, \\ &= -x_j^3/6 + x_j^2y/2 + x_jy/\beta + 1/\alpha; & y \geq x_j, \end{aligned}$$

and

$$(25) \quad \begin{aligned} J(y, x_j) &= y^2/2 + y/\beta; & y \leq x_j, \\ &= -x_j^2/2 + x_jy + y/\beta; & y \geq x_j. \end{aligned}$$

In particular, if we construct the  $2(n-1)$  functions

$$(26) \quad \begin{aligned} M_i(y) &= K(y, x_{i+1}) - K(y, x_i) - \frac{(x_{i+1} - x_i)}{2} [J(y, x_{i+1}) + J(y, x_i)], \\ N_i(y) &= \frac{J(y, x_{i+1}) - J(y, x_i)}{x_{i+1} - x_i}; \quad j = 1, 2, \dots, n-1, \end{aligned}$$

independent of  $\alpha$  and  $\beta$ , so that

$$\begin{aligned} M_i(y) &= 0; & y \leq x_i, \\ &= -y^3/6 + y^2(x_i + x_{i+1})/4 - x_ix_{i+1}y/2 + x_i^2(3x_{i+1} - x_i)/12; \\ & & x_i \leq y \leq x_{i+1}, \\ &= (x_{i+1} - x_i)^3/12; & y \geq x_{i+1}, \end{aligned}$$

and

$$\begin{aligned}
 N_j(y) &= 0; & y &\leq x_j, \\
 &= \frac{1}{2}(y - x_j)^2/(x_{j+1} - x_j); & x_j &\leq y \leq x_{j+1}, \\
 &= y - (x_j + x_{j+1})/2; & y &\geq x_{j+1},
 \end{aligned}$$

these too must be treated exactly by the optimal approximation rule.

Now notice that as  $\alpha$  and  $\beta$  become very small the linear parts of  $K(y, x_j)$  and  $J(y, x_j)$  become dominant. Thus, in the limit where  $\alpha$  and  $\beta$  separately approach zero, the optimal approximation must be exact for the functions  $f = 1$  and  $f = y$ ; indeed, this is the reason for being interested in the limiting case. These two conditions, together with the  $2(n - 1)$  similar conditions on  $\{M_j(y), N_j(y); j = 1, 2, \dots, n - 1\}$  from which  $\alpha$  and  $\beta$  have been eliminated, will, in general, provide enough information to enable us to determine optimal values of the parameters  $\{w_j, x_j; j = 1, 2, \dots, n\}$  in the limiting case.

Specialising further, we consider the case where

$$(27) \quad Lf = \int_0^1 f(y) dy.$$

Application of the above exactness conditions then leads to the equations

$$\begin{aligned}
 1: \quad & \sum_{k=1}^n w_k = 1, \\
 y: \quad & \sum_{k=1}^n w_k x_k = \frac{1}{2}, \\
 (28) \quad M_j(y): \quad & \sum_{k=j}^n w_k (x_{j+1} - x_j)^3 / 12 = \frac{(x_{j+1} - x_j)^3}{12} \cdot \left(1 - \frac{x_j + x_{j+1}}{2}\right), \\
 & \qquad \qquad \qquad j = 1, 2, \dots, n - 1, \\
 N_j(y): \quad & \sum_{k=j}^n w_k \left(x_k - \frac{x_j + x_{j+1}}{2}\right) = \frac{1}{2} - \frac{x_j + x_{j+1}}{2} + \frac{x_j^2 + x_j x_{j+1} + x_{j+1}^2}{6}, \\
 & \qquad \qquad \qquad j = 1, 2, \dots, n - 1,
 \end{aligned}$$

for the required weights and abscissae.

After some algebraic manipulation, it may be verified that Eqs. (28) are satisfied by the following values for  $\{w_j, x_j; j = 1, 2, \dots, n\}$ :

$$\begin{aligned}
 (29) \quad w_1 &= \frac{1}{2} - (n/2 - 1)h = w_n; & w_j &= h; & j &= 2, 3, \dots, n - 1, \\
 x_1 &= \frac{1}{2} - (n - 1)h/2; & x_{j+1} &= x_j + h; & j &= 1, 2, \dots, n - 2; \\
 (30) \quad x_n &= \frac{1}{2} + (n - 1)h/2,
 \end{aligned}$$

where the "step-length" parameter  $h$  satisfies the equation

$$(31) \quad (3n^2 - 6n + 1)h^2 - 6(n - 1)h + 3 = 0,$$



whence

$$(32) \quad h = \frac{3(n-1) \pm \sqrt{6}}{3n^2 - 6n + 1}.$$

If the negative root is chosen in (32) all the abscissae lie within the support of the functions. Choice of the positive root in (32) leads to a situation wherein  $x_1$  and  $x_n$  lie just outside the interval  $[0, 1]$ , which is unacceptable for functions defined only within that interval. The above results agree with those obtained, using special techniques, by Krylov ([4, p. 140], etc.) and Stern [8], although in the English translation, there appears to be a misprint in the expression given by Krylov for the values of the weights.

(iii) For the third example we consider the Paley-Wiener space (Payley and Wiener, [6]) of entire functions  $f(z)$  whose restrictions to the real axis have finite bandwidth  $2a$ , with inner product defined by

$$(33) \quad (f, g) = \int_{-\infty}^{\infty} f(t) \cdot \bar{g}(t) \cdot dt < \infty.$$

The finite bandwidth property means that, for any function  $f(z)$  in the space,

$$F(w) = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} e^{iwt} f(t) \cdot dt$$

exists for all  $w$  and vanishes for real  $w$  outside the interval  $[-a, a]$ . An alternative characterisation of the space is that its members are entire functions of "exponential type at most  $a$ "; in other words, there exists  $a$  such that for all  $f(z)$  in the space,

$$(34) \quad \limsup_{|z| \rightarrow \infty} \left[ \frac{\log |f(z)|}{|z|} \right] < a < \infty$$

(de Branges, [3]).

The reproducing kernel function for this space is given by

$$(35) \quad K(z, \bar{x}) = \frac{\text{Sin } [a(z - \bar{x})]}{\pi(z - \bar{x})}.$$

Thus, if the abscissae  $\{x_j; j = 1, 2, \dots, n\}$  are prescribed, Theorem 1 tells us that an optimal linear estimation rule must be exact for functions

$$(36) \quad \frac{\text{Sin } [a(z - \bar{x}_j)]}{\pi(z - \bar{x}_j)}; \quad j = 1, 2, \dots, n.$$

In particular, if the real  $\{x_j\}$  are equispaced with an interval  $\pi/a$ , we find that the optimal weights are then given by

$$(37) \quad w_j = L \left( \frac{\text{Sin } [a(z - x_j)]}{a(z - x_j)} \right)$$

for any bounded linear functional  $L$ . For example, if we choose

$$Lf = f(x), \quad \text{for some fixed, real } x,$$

we find that the optimal interpolation rule, in this Hilbert space with these pre-

scribed abscissae, may be written in the form

$$(38) \quad \hat{f}(x) = \sum_{j=1}^n f_j \frac{\text{Sin} [a(x - x_j)]}{a(x - x_j)},$$

i.e., the optimal interpolating function is the well-known cardinal function! (Whittaker, [9]). Optimal values of bounded linear functionals can now be found by applying the functionals to  $\hat{f}(x)$  (cf. Corollary 1).

Returning again to quadrature rules, if we attempt to choose

$$Lf = \int_{-\infty}^{\infty} f(t) \cdot dt$$

we find that the representer of  $L$  is unity over  $(-\infty, \infty)$ ; i.e., it is not strictly a member of the Hilbert space under consideration since its norm is not finite. However, for suitable  $f(t)$  and real  $\epsilon$ , we can write

$$(39) \quad \int_{-\infty}^{\infty} f(t) \cdot dt = \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} \frac{\text{Sin} (\epsilon t)}{\epsilon t} \cdot f(t) \cdot dt = \lim_{\epsilon \rightarrow 0} L_{\epsilon} f, \quad \text{say.}$$

Here  $L_{\epsilon}$  is a bounded linear functional: since its representer  $(\text{Sin} (\epsilon t))/\epsilon t$  is a member of the space if  $\epsilon < a$ ; hence, we can apply the results of Section 2.

In this case the abscissae which are optimal for the purpose of estimating  $L_{\epsilon} f$  are found by maximising the quantity

$$(40) \quad T = \sum_{j=1}^n \sum_{k=1}^n \frac{\text{Sin} (\epsilon x_j)}{\epsilon x_j} \cdot \frac{\text{Sin} (\epsilon \bar{x}_k)}{\epsilon \bar{x}_k} \cdot [G^{-1}]_{jk},$$

where the matrix  $G$  is given by

$$(41) \quad G_{jk} = \frac{\text{Sin} [a(x_j - \bar{x}_k)]}{x_j - \bar{x}_k}; \quad j, k = 1, 2, \dots, n.$$

Now, although

$$(42) \quad \left\| \frac{\text{Sin} (\epsilon t)}{\epsilon t} \right\|^2 = \frac{\pi}{\epsilon} \rightarrow \infty \quad \text{as} \quad \epsilon \rightarrow 0,$$

limiting values of the optimal abscissae may be found by maximising the sum of the elements of the inverse of the Gram matrix  $G$ . However, it is necessary to observe that these limiting values will be arbitrary to the extent of an additive constant, in agreement with the fact that the integral of a function over the real line is unchanged by a finite shift of origin. Furthermore, it is clear by inspection of Eq. (40) that if the  $\{x_j; j = 1, 2, \dots, n\}$  maximise  $T$ , so do the values  $\{\bar{x}_j; j = 1, 2, \dots, n\}$  and hence, also the values  $\{-x_j; j = 1, 2, \dots, n\}$  and the values  $\{x_{n-j+1}; j = 1, 2, \dots, n\}$ . Thus, we can take the limiting optimal abscissae as real and symmetrically distributed about the origin.

A table of corresponding optimal values of weights and abscissae is given below, for various values of  $n$ . The maximisation of  $T$  was performed numerically, using a conjugate gradient method (Fletcher and Reeves, [11]).

It is conjectured that, as  $n \rightarrow \infty$ , the optimal weights and the spacing between optimal abscissae all approach the value  $2\pi$ .

Table of Optimal Quadrature Constants in Paley-Wiener-Hilbert Space

$n$	$ax_i$	$w_i$	$n$	$ax_i$	$w_i$
2	$\pm 2.2467047$	4.01344867	10	$\pm 2.9188177$	5.77116011
3	0.0	4.85312962		$\pm 8.7426567$	5.73687292
	$\pm 4.9049245$	4.27649995		$\pm 14.5180889$	5.64933802
4	$\pm 2.6463646$	5.10581247		$\pm 20.1773847$	5.43372596
	$\pm 7.7109632$	4.40320621		$\pm 25.4895296$	4.62744777
5	0.0	5.35165706	11	0.0	5.81627482
	$\pm 5.4422714$	5.22869607		$\pm 5.8708587$	5.80398102
	$\pm 10.5936134$	4.47811596		$\pm 11.7198148$	5.76213088
6	$\pm 2.7932979$	5.47077261		$\pm 17.5152246$	5.66962713
	$\pm 8.3169775$	5.30201789		$\pm 23.1910859$	5.45055408
	$\pm 13.5233788$	4.52780036		$\pm 28.5174116$	4.64118298
7	0.0	5.58754472	12	$\pm 2.9517948$	5.84866593
	$\pm 5.6649413$	5.54180744		$\pm 8.8472671$	5.82879692
	$\pm 11.2404772$	5.35102863		$\pm 14.7156617$	5.78195825
	$\pm 16.4851663$	4.56327339		$\pm 20.5270356$	5.68600816
8	$\pm 2.8706889$	5.65704249		$\pm 26.2164566$	5.46443334
	$\pm 8.5854996$	5.58931584		$\pm 31.5545963$	4.65268310
	$\pm 14.1972269$	5.38624824	13	0.0	5.88072356
	$\pm 19.4701975$	4.58993178		$\pm 5.9278222$	5.87314378
9	0.0	5.72548305		$\pm 11.8423367$	5.84827293
	$\pm 5.7899561$	5.70348109		$\pm 17.7262660$	5.79797010
	$\pm 11.5395763$	5.62348848		$\pm 23.5507171$	5.69953235
	$\pm 17.1781264$	5.41286087		$\pm 29.2514594$	5.47609002
	$\pm 22.4729048$	4.61073592		$\pm 34.5996047$	4.66246074

Computing Centre  
Queen's University  
Kingston, Ontario, Canada

1. N. ARONSZAJN, "Theory of reproducing kernels," *Trans. Amer. Math. Soc.*, v. 68, 1950, pp. 337-404. MR 14, 479.
2. M. GOLOMB & H. F. WEINBERGER, *Optimal Approximation and Error Bounds*, Proc. Sympos. (Madison, Wis., 1958) Univ. of Wisconsin Press, Madison, Wis., 1959, pp. 117-190. MR 22 #12697.
3. L. DE BRANGES, *Hilbert Spaces of Entire Functions*, Prentice-Hall, Englewood Cliffs, N. J., 1968. MR 37 #4590.
4. V. I. KRYLOV, *Approximate Calculation of Integrals*, Fizmatgiz, Moscow, 1959; English transl., Macmillan, New York, 1962. MR 22 #2002; MR 26 #2008.
5. H. MESCHKOWSKI, *Hilbertsche Räume mit Kernfunktion*, Die Grundlehren der math. Wissenschaften, Band 113, Springer-Verlag, Berlin, 1962. MR 25 #4326.
6. R. E. A. C. PALEY & N. WIENER, *Fourier Transforms in the Complex Domain*, Amer. Math. Soc. Colloq. Publ., vol. 19, Amer. Math. Soc., Providence, R. I., 1934; reprint, 1967.
7. A. SARD, "Best approximate integration formulas," *Amer. J. Math.*, v. 71, 1949, pp. 80-91. MR 10, 576.
8. M. D. STERN, "Optimal quadrature formulae," *Comput. J.*, v. 9, 1967, pp. 396-403. MR 35 #3885.
9. E. T. WHITTAKER, "On the functions which are represented by the expansions of the interpolation theory," *Proc. Edinburg Math. Soc.*, v. 35, 1915, pp. 181-194.
10. N. RICHTER & P. RABINOWITZ, Weizmann Institute of Science, Rehovot, Israel. (Private communication.)
11. R. FLETCHER & C. M. REEVES, "Function minimization by conjugate gradients," *Comput. J.*, v. 7, 1964, pp. 149-154. MR 32 #4827.