

The Use of the Hypercircle Inequality in Deriving a Class of Numerical Approximation Rules for Analytic Functions

By Richard A. Valentin

1. Introduction. The relation of the hypercircle inequality to a large class of numerical approximation problems has been demonstrated by Golomb and Weinberger [1]. The applications considered in the present paper, while closely related to those of [1], concern a particular space of analytic functions and are similar in many respects to the results of Davis [2], [3] on errors of numerical approximation. Consideration of the functions analytic in $|z| < R$ and continuous on $|z| = R$ yields a Hilbert space having as its reproducing kernel an elementary form of the Szegő kernel function. For approximation in terms of a given set of point functionals, the simplicity of this kernel yields equations defining the various terms of the hypercircle approximation which may be solved in closed form.

In Section 2 we review the hypercircle inequality and the various equations relevant to its use in constructing approximation rules. These results are specialized in the following section to a particular space of analytic functions, while Section 4 contains examples to illustrate their use in specific cases. Presented as examples are an elementary interpolation rule and an integration formula for complex functions closely analogous to the 5-point rule of Birkhoff and Young.

2. The Hypercircle Inequality and Related Results. We begin by listing certain definitions and results relating to the theory of a complex Hilbert space \mathcal{H} . The theorems included are well known and will be stated without proof.*

DEFINITION 2.1. Let x_1, \dots, x_n be n independent elements of \mathcal{H} and c_1, \dots, c_n be n given constants. The set $y \in \mathcal{H}$ such that $(y, x_i) = c_i, i = 1, \dots, n$ is called a hyperplane of codimension n and shall be denoted by P_n .

DEFINITION 2.2. The portion of \mathcal{H} common to P_n and to the ball $\|y\| \leq r$ is called a hypercircle C_r .

THEOREM 2.1. P_n may be represented in terms of the orthonormal set $\{x_i^*\}$ derived by the Gram-Schmidt process from $\{x_i\}$ as

$$(y, x_i^*) = a_i, \quad i = 1, \dots, n,$$

where

$$(2.1) \quad \bar{a}_k = \bar{c}_1 [g(x_1)]^{-1/2} \begin{vmatrix} (x_1, x_1) & \cdots & (x_k, x_1) \\ \vdots & & \vdots \\ (x_1, x_{k-1}) & \cdots & (x_k, x_{k-1}) \\ \bar{c}_1 & \cdots & \bar{c}_k \end{vmatrix} \cdot [g(x_1, \dots, x_{k-1})g(x_1, \dots, x_k)]^{-1/2}$$

Received November 9, 1966.

* For relevant proofs see Davis [4], Chapter IX.

and $g(x_1, \dots, x_k)$ is the Gram matrix of order k .

THEOREM 2.2. *The element of P_n closest to the origin (i.e., of minimum norm) is $w = \sum_{i=1}^n a_i x_i^*$ where the constants a_i are as in (2.1).*

THEOREM 2.3. *If L is a bounded linear functional over \mathcal{H} , then there exists a unique element $h \in \mathcal{H}$ such that $Lx = (x, h)$, $x \in \mathcal{H}$. The element h is called the representer of L .*

THEOREM 2.4 (THE HYPERCIRCLE INEQUALITY). *Let w be the element of P_n which is nearest the origin. Then for any $x \in C_r$ and any bounded linear functional L having representer h*

$$(2.2) \quad |Lx - Lw|^2 \leq (r^2 - \|w\|^2) \left(\|h\|^2 - \sum_{k=1}^n |Lx_k^*|^2 \right).$$

In the class of approximations to be considered, it will be assumed that we are given a bounded linear functional L over \mathcal{H} , a bound r on the norm of some given $x \in \mathcal{H}$ and auxiliary conditions on x defining a hyperplane P_n . In particular, we shall be given data of the form $L_j x = c_j$, $j = 1, \dots, n$, where L_j are the point functionals associated with a fixed, discrete set of points in the complex plane. That is

$$(2.3) \quad L_j x = x(z_j).$$

We wish to express the approximation Lw in terms of a weighted sum of the values c_j and to obtain explicit expressions for this approximation and for all of the terms occurring in the error bound implied by (2.2).

We assume \mathcal{H} is such that it possesses a known reproducing kernel $K(z, \bar{w})$. That is

$$(2.4) \quad (x(z), K(z, \bar{w})) = x(w), \quad x \in \mathcal{H}.$$

If L is a bounded linear functional over \mathcal{H} , then the representer h of L is known to be given by operation on K as

$$h(w) = \overline{L_z K(z, \bar{w})}$$

and thus

$$(2.5) \quad Lx = (x, h), \quad x \in \mathcal{H}.$$

From (2.5) we may express the given point functional data in the form of a hyperplane P_n , where x_j , the representer of L_j , is

$$(2.6) \quad x_j(w) = \overline{K(z_j, \bar{w})}.$$

Assume that we have applied the Gram-Schmidt process to the set $\{x_i\}$ obtained as in (2.6) and that the elements of the resulting orthonormal set $\{x_i^*\}$ have been written in the form

$$x_i^* = \sum_{k=1}^i b_{ik} x_k, \quad i = 1, \dots, n,$$

where the b_{ik} are known constants. Considering P_n expressed in the form $(x, x_i^*) = a_i$, $i = 1, \dots, n$, we see that this is equivalent to

$$a_i = \left(x, \sum_{k=1}^i b_{ik} x_k \right) = \sum_{k=1}^i \overline{b_{ik}} c_k$$

and that we may write the estimate Lw as a weighted sum of the known hyperplane data c_i by noting

$$\begin{aligned} Lw &= \sum_{k=1}^n a_k Lx_k^* = \sum_{k=1}^n a_k(x_k^*, h) \\ &= \sum_{k=1}^n \left(\sum_{l=1}^k b_{kl} c_l \right) (x_k^*, h) \\ &= \sum_{k=1}^n \alpha_k c_k, \end{aligned}$$

where

$$(2.7) \quad \alpha_k = \sum_{l=k}^n \overline{b_{lk}} (x_l^*, h).$$

If we now consider the minimum problem

$$\min_{\beta_k} \left\| h - \sum_{k=1}^n \beta_k x_k \right\|$$

and its solution

$$(2.8) \quad \left\| h - \sum_{k=1}^n (h, x_k^*) x_k^* \right\|,$$

we note that

$$\sum_{k=1}^n (h, x_k^*) x_k^* = \sum_{k=1}^n \left(\sum_{l=k}^n b_{lk} (h, x_l^*) \right) x_k,$$

and hence that $\beta_k = \overline{\alpha}_k$. However, the Fourier solution (2.8) is equivalent to solving the normal equations

$$\sum_{i=1}^n \beta_i (x_i, x_j) = (h, x_j), \quad j = 1, \dots, n,$$

and hence we shall have determined the weight coefficients α_k by inverting the Gram matrix $[(x_i, x_j)]_{i,j=1}^n$. Denoting the required inverse matrix by $[\gamma_{ij}]_{i,j=1}^n$, we have

$$(2.9) \quad \alpha_k = \sum_{l=1}^n \overline{\gamma}_{lk} (x_l, h).$$

Having (2.9), the terms of the error bound implied by the hypercircle inequality are readily computed, since

$$\begin{aligned} \sum_{k=1}^n |Lx_k^*|^2 &= \sum_{k=1}^n (x_k^*, h) (h, x_k^*) \\ (2.10) \quad &= L \left[\sum_{k=1}^n (h, x_k^*) x_k^* \right] = L \left[\sum_{k=1}^n \overline{\alpha}_k x_k \right] \\ &= \sum_{k=1}^n \overline{\alpha}_k (x_k, h), \end{aligned}$$

while the minimum distance term $\|w\|$ is just given by $\|w\|^2 = \sum_{k=1}^n |a_k|^2$.

From the above results we see that the problem of obtaining the estimate and bound implied by the hypercircle inequality may be summarized by stating that one has obtained a complete solution if an inverse is known for the Gram matrix formed by the representers x_k and one can compute the constants a_k defined by (2.1).

3. The Hilbert Space $H^2(R)$. We shall consider the space of functions analytic in $|z| < R$ and continuous on $|z| = R$. It is known that each function of this class, which we shall denote by $H^2(R)$, possesses the representation**

$$(3.1) \quad f(z) = \frac{R}{2\pi} \int_{|w|=R} \frac{f(w)}{(R^2 - z\bar{w})} ds_w,$$

where ds_w denotes the element of arclength along $|w| = R$. Equivalently, $H^2(R)$ is a Hilbert space where the appropriate inner product and norm definitions are

$$(3.2) \quad \begin{aligned} (f, g) &= \int_{|w|=R} f\bar{g} ds_w, \quad f, g \in H^2(R), \\ \|f\| &= (f, f)^{1/2} < \infty, \quad f \in H^2(R). \end{aligned}$$

From (3.2) we see that the representation (3.1) is equivalent to stating that $H^2(R)$ has as its reproducing kernel

$$(3.3) \quad K(z, \bar{w}) = (2\pi)^{-1}R(R^2 - z\bar{w})^{-1}.$$

If one now considers a given set of point functionals

$$L_j f = f(z_j), \quad j = 1, \dots, n, \quad f \in H^2(R),$$

then from (3.3) and (2.6), the corresponding set of representers are

$$(3.4) \quad x_j(w) = (2\pi)^{-1}R(R^2 - \bar{z}_j w)^{-1}$$

and, hence,

$$(3.5) \quad (x_i, x_j) = (2\pi)^{-1}R(R^2 - \bar{z}_i z_j)^{-1}.$$

From the form of the inner product (3.5) it is evident that an explicit solution of the normal equations and evaluation of the constants a_i of (2.1) depends on the ability to compute determinants having the form of Cauchy's double alternant

$$D_n = \det [(p_i + q_j)^{-1}]_{i,j=1}^n.$$

However, it is known [4, p. 268] that

$$(3.6) \quad D_n = \prod_{i>j}^n (p_i - p_j)(q_i - q_j) \left[\prod_{i,j=1}^n (p_i + q_j) \right]^{-1},$$

and thus the required evaluations are a matter of applying (3.6) and performing some elementary algebra. Omitting the details, we note the following results:

For the coefficients defined in (2.1)

** Our definition of $H^2(R)$ is quite restrictive in the sense that all of the results stated in this section hold for a much wider class of functions; however, for the intended applications, the above will be sufficient.

$$\begin{aligned}
 \bar{a}_1 &= (2\pi R^{-1})^{1/2} (R^2 - z_1 \bar{z}_1)^{1/2} \bar{c}_1 \\
 \bar{a}_k &= (2\pi R^{-1})^{1/2} R^{-k+1} (R^2 - z_k \bar{z}_k)^{1/2} e^{i\phi_k} \left[\prod_{l=1}^{k-1} (\bar{z}_k - \bar{z}_l) \right]^{-1} \\
 &\quad \sum_{j=1}^k (-1)^{k+j} \bar{c}_j \left[\prod_{l=1}^{k-1} (R^2 - \bar{z}_j z_l) \right] \quad (k = 2, 3, \dots, n),
 \end{aligned}
 \tag{3.7}$$

where

$$e^{i\phi_k} \equiv \prod_{l=1}^{k-1} \left\{ \frac{(\bar{z}_k - \bar{z}_l)}{|z_k - z_l|} \left[\frac{(R^2 - \bar{z}_k z_l)}{|R^2 - \bar{z}_k z_l|} \right]^{-1} \right\}.
 \tag{3.8}$$

The factor $e^{i\phi_k}$ may be neglected, since in the applications we are only interested in $|a_i|$.

For the elements of the inverse matrix needed in (2.9) we obtain

$$\begin{aligned}
 \gamma_{ij} &= (2\pi R^{-1}) R^{-(2n-2)} \left[\prod_{l=1}^n (R^2 - \bar{z}_j z_l) \prod_{k=1; (k \neq j)}^n (R^2 - \bar{z}_k z_i) \right] \\
 &\quad \cdot \left[\prod_{l=1; (l \neq i)}^n (z_i - z_l) \prod_{k=1; (k \neq j)}^n (\bar{z}_j - \bar{z}_k) \right]^{-1}.
 \end{aligned}
 \tag{3.9}$$

One may quite easily derive expressions for the orthonormal set $\{x_i^*\}$ and the coefficients b_{ij} relating $\{x_i^*\}$ to $\{x_i\}$; however, they are not needed in the applications and hence will not be included.

4. Applications. Consider an elementary interpolation problem where we are given $f(z_1) = c_1$, $f(z_2) = c_2$ and we wish to approximate $f(z_0)$. In this case $Lf = f(z_0)$ and the representer h of L is simply

$$h(w) = (2\pi)^{-1} R (R^2 - \bar{z}_0 w)^{-1}$$

while

$$\|h\|^2 = (2\pi)^{-1} R (R^2 - \bar{z}_0 z_0)^{-1}.$$

Also, one has

$$\overline{(h, x_k)} = (x_k, h) = (2\pi)^{-1} R (R^2 - \bar{z}_k z_0)^{-1}, \quad k = 1, 2.$$

Applying (3.9) and (2.9) yields

$$Lw = \alpha_1 c_1 + \alpha_2 c_2,$$

where

$$\begin{aligned}
 \alpha_1 &= \frac{(R^2 - \bar{z}_1 z_1)(R^2 - \bar{z}_2 z_1)}{(R^2 - \bar{z}_1 z_0)(R^2 - \bar{z}_2 z_0)} \frac{(z_0 - z_2)}{(z_1 - z_2)}, \\
 \alpha_2 &= \frac{(R^2 - \bar{z}_1 z_2)(R^2 - \bar{z}_2 z_2)}{(R^2 - \bar{z}_1 z_0)(R^2 - \bar{z}_2 z_0)} \frac{(z_1 - z_0)}{(z_1 - z_2)}.
 \end{aligned}
 \tag{4.1}$$

From (2.10) and the above values of α_1 , α_2 , and $\|h\|^2$ one obtains

$$\|h\|^2 - \sum_{k=1}^2 |Lx_k^*|^2 = \frac{R}{2\pi} \frac{R^4 |z_0 - z_1|^2 |z_0 - z_2|^2}{|R^2 - \bar{z}_1 z_0|^2 |R^2 - \bar{z}_2 z_0|^2 (R^2 - z_0 \bar{z}_0)},
 \tag{4.2}$$

while for the computation of $\|w\|^2 = |a_1|^2 + |a_2|^2$, Eq. (3.7) yields

$$(4.3) \quad \begin{aligned} a_1 &= (2\pi R^{-1})^{1/2} (R^2 - z_1 \bar{z}_1)^{1/2} c_1 \\ a_2 &= (2\pi R^{-1})^{1/2} (R^2 - z_2 \bar{z}_2)^{1/2} R^{-1} (z_2 - z_1)^{-1} \\ &\quad \times [-C_1 (R^2 - z_1 \bar{z}_1) + C_2 (R^2 - z_2 \bar{z}_1)]. \end{aligned}$$

The weights (4.1) are similar in form to the Lagrange interpolation coefficients, and in the limit as $R \rightarrow \infty$ they reduce to these coefficients. In considering n -point interpolation formulas, one may show that the weights are given by the obvious generalization of (4.1) and, similarly, that (4.2), the portion of the error bound independent of the particular function being considered, generalizes in the obvious way. The expression for $\|w\|^2$, however, is not conveniently expressible in the general case.

As a second example we shall consider an approximation analogous to a special case of the 5-point integration rule

$$(4.4) \quad \int_{z_3}^{z_1} f(z) dz \cong \frac{h}{15} [24f_0 + 4(f_1 + f_3) - (f_2 + f_4)]$$

developed by Birkhoff and Young [5]. In (4.4), $f_k = f(z_k)$ and $z_k = z_0 + (i)^{k-1}h$, $k = 1, \dots, 4$, where f is analytic in $|z - z_0| \leq h_1$, $h_1 > |h|$, with h a known complex constant. If one takes $z_0 = 0$, $h = i$, then (4.4) becomes

$$(4.5) \quad \int_{-i}^i f(z) dz \cong \frac{i}{15} \{24f(0) + 4[f(i) + f(-i)] - [f(1) + f(-1)]\}.$$

We shall consider the hypercircle result equivalent to (4.5) by taking

$$Lf = \int_{-i}^i f(z) dz, \quad f \in H^2(R), \quad R > 1,$$

with $L_k f = f(z_k) = c_k$, $k = 1, \dots, 5$, where

$$z_1 = -z_5 = \lambda, \quad z_2 = -z_4 = i\lambda, \quad z_3 = 0,$$

and λ is an as yet unspecified real constant. This yields the coefficient matrix

$$\frac{R}{2\pi} \begin{bmatrix} G_1 & G_2 & G_3 & \bar{G}_2 & G_4 \\ \bar{G}_2 & G_1 & G_3 & G_4 & G_2 \\ G_3 & G_3 & G_3 & G_3 & G_3 \\ G_2 & G_4 & G_3 & G_1 & \bar{G}_2 \\ G_4 & \bar{G}_2 & G_3 & G_2 & G_1 \end{bmatrix},$$

where $G_1 = (R^2 - \lambda^2)^{-1}$, $G_2 = (R^2 + i\lambda^2)^{-1}$, $G_3 = R^{-2}$, $G_4 = (R^2 + \lambda^2)^{-1}$. Application of the results of the previous section yields the inverse

$$\left(\frac{2\pi}{R}\right) \frac{(R^8 - \lambda^8)}{16\lambda^8 R^4} \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 & \bar{\gamma}_2 & \gamma_4 \\ \bar{\gamma}_2 & \gamma_1 & \gamma_3 & \gamma_4 & \gamma_2 \\ \gamma_3 & \gamma_3 & \gamma_5 & \gamma_3 & \gamma_3 \\ \gamma_2 & \gamma_4 & \gamma_3 & \gamma_1 & \bar{\gamma}_2 \\ \gamma_4 & \bar{\gamma}_2 & \gamma_3 & \gamma_2 & \gamma_1 \end{bmatrix},$$

where

$$\begin{aligned} \gamma_1 &= (R^4 + \lambda^4)(R^2 + \lambda^2), & \gamma_2 &= (R^4 - \lambda^4)(R^2 + i\lambda^2), & \gamma_3 &= -4R^6, \\ \gamma_4 &= (R^4 + \lambda^4)(R^2 - \lambda^2), & \gamma_5 &= 16R^{14}(R^8 - \lambda^8)^{-1}, \end{aligned}$$

while for the remaining inner products we have

$$\begin{aligned} (h, x_1) &= (h, x_5) = -(2\pi)^{-1}iR \tan^{-1}(\lambda R^{-2}) \\ (h, x_2) &= (h, x_4) = -(2\pi)^{-1}iR \tanh^{-1}(\lambda R^{-2}) \\ (h, x_3) &= -(2\pi)^{-1}iR(\lambda R^{-2}). \end{aligned}$$

Thus one finds

$$Lw = \alpha_1[f(\lambda) + f(-\lambda)] + \alpha_2[f(i\lambda) + f(-i\lambda)] + \alpha_3f(0),$$

where

$$\begin{aligned} \alpha_1 &= i(R^8 - \lambda^8)(4\lambda^9 R^2)^{-1}[(R^4 + \lambda^4) \tan^{-1}(\lambda R^{-2}) \\ &\quad + (R^4 - \lambda^4) \tanh^{-1}(\lambda R^{-2}) - 2\lambda R^2], \\ (4.6) \quad \alpha_2 &= i(R^8 - \lambda^8)(4\lambda^9 R^2)^{-1}[(R^4 - \lambda^4) \tan^{-1}(\lambda R^{-2}) \\ &\quad + (R^4 + \lambda^4) \tanh^{-1}(\lambda R^{-2}) - 2\lambda R^2], \\ \alpha_3 &= -i(R^8 - \lambda^8)R^2\lambda^{-9}[\tan^{-1}(\lambda R^{-2}) + \tanh^{-1}(\lambda R^{-2}) \\ &\quad - 2\lambda R^6(R^8 - \lambda^8)^{-1}]. \end{aligned}$$

The constants determining $\|w\|$ are

$$\begin{aligned} a_1 &= (2\pi R^{-1})^{1/2}(R^2 - \lambda^2)^{1/2}c_1, \\ a_2 &= (2\pi R^{-1})^{1/2}(R^2 - \lambda^2)^{1/2}[-\lambda R(1 - i)]^{-1}[-(R^2 - \lambda^2)c_1 + (R^2 - i\lambda^2)c_2], \\ a_3 &= (2\pi R^{-1})^{1/2}R[i\lambda^2 R^2]^{-1} \\ &\quad \times [(R^2 - \lambda^2)(R^2 + i\lambda^2)c_1 - (R^2 - i\lambda^2)(R^2 - \lambda^2)c_2 + R^4 c_3], \\ a_4 &= (2\pi R^{-1})^{1/2}(R^2 - \lambda^2)^{1/2}[2\lambda^3 R^8(1 + i)]^{-1} \\ &\quad \times R^2[-(R^2 - \lambda^2)(R^2 + i\lambda^2)c_1 + (R^2 - i\lambda^2)(R^2 - \lambda^2)c_2 \\ &\quad - R^4 c_3 + (R^2 + i\lambda^2)(R^2 + \lambda^2)c_4], \\ a_5 &= (2\pi R^{-1})^{1/2}(R^2 - \lambda^2)^{1/2}[4\lambda^4 R^4]^{-1} \\ &\quad \times R^2[(R^2 - \lambda^2)(R^4 + \lambda^4)c_1 - (R^4 - \lambda^4)(R^2 - i\lambda^2)c_2 + R^6 c_3 \\ &\quad - (R^4 - \lambda^4)(R^2 + i\lambda^2)c_4 + (R^4 + \lambda^4)(R^2 + \lambda^2)c_5], \end{aligned}$$

while that portion of the error bound independent of the particular function operated upon is

$$\begin{aligned} \sigma^2 &\equiv \|h\|^2 - \sum_{k=1}^5 |Lx_k^*|^2 = 2(\pi R)^{-1} \sum_{k=0}^{\infty} (2k + 1)^{-2} R^{-4k} \\ &\quad - \{(2\pi R\lambda^{10})^{-1}(R^8 - \lambda^8)[R^4(\tan^{-1}(\lambda R^{-2}) + \tanh^{-1}(\lambda R^{-2}))^2 \\ &\quad \quad + \lambda^4(\tan^{-1}(\lambda R^{-2}) - \tanh^{-1}(\lambda R^{-2}))^2] \\ &\quad \quad + 2R^7(\pi\lambda^8)^{-1} - 2R(\pi\lambda^9)^{-1}(R^8 - \lambda^8)[\tan^{-1}(\lambda R^{-2}) + \tanh^{-1}(\lambda R^{-2})]\}. \end{aligned}$$

In the limit as $R \rightarrow \infty$, the weights (4.6) reduce to

$$\alpha_1 \rightarrow i(30\lambda^4)^{-1}(3 - 5\lambda^2),$$

$$\alpha_2 \rightarrow i(30\lambda^4)^{-1}(3 + 5\lambda^2),$$

$$\alpha_3 \rightarrow -2i(5\lambda^4)^{-1}(1 - 5\lambda^2),$$

which for the case $\lambda = 1$ are identical with those of the Birkhoff-Young rule (4.5). The constant λ has been included to indicate the possibility of producing an optimum choice of points z_k . For example, one may form $\partial\sigma/\partial\lambda = 0$ and solve the resulting transcendental equation to obtain $\lambda = \lambda(R)$ such that σ is a minimum for each given R . We shall not carry out the details of such a computation, but merely note that expanding the resulting transcendental equation in powers of R^{-1} yields as the leading term $(4/63)\lambda^{12}R^{-20}(7\lambda^4 - 3)$ and, hence, for the integration of functions having a large radius of analyticity, a near optimal choice of λ would be $\lambda = (3/7)^{1/4}$.

Reactor Engineering Division
Argonne National Laboratory
Argonne, Illinois

1. M. GOLOMB & H. F. WEINBERGER, "Optimal approximation and error bounds," *On Numerical Approximation*, Proc. Sympos., Madison, April 21-23, 1958, Univ. of Wisconsin Press, Madison, Wis., 1959, pp. 117-196. MR 22 #12697.
2. P. DAVIS, "Errors of numerical approximation for analytic functions," *J. Rational Mech. Anal.*, v. 2, 1953, pp. 303-313. MR 14, 907.
3. P. DAVIS, "Errors of numerical approximation for analytic functions," *Surveys of Numerical Analysis*, McGraw-Hill, New York, 1962, pp. 468-484. MR 24 #B1766.
4. P. DAVIS, *Interpolation and Approximation*, Blaisdell, New York, 1963. MR 28 #393.
5. G. BIRKHOFF & D. YOUNG, "Numerical quadrature of analytic and harmonic functions," *J. Math. Phys.*, v. 29, 1950, pp. 217-221. MR 12, 445.