

SIMPLIFIED TECHNIQUE FOR CONSTRUCTING ORTHONORMAL FUNCTIONS

M. O. PEACH

1. **Introduction.** An orthonormalization process starts with a set of linearly independent functions

$$f_1, f_2, \dots,$$

and the set of complex conjugate functions

$$\bar{f}_1, \bar{f}_2, \dots,$$

all defined over a given region R . From these are constructed a set of functions

$$g_1, g_2, \dots,$$

and the set of complex conjugate functions

$$\bar{g}_1, \bar{g}_2, \dots,$$

defined over R and such that

$$\int_R g_m \bar{g}_n dR = \begin{cases} 0, & \text{if } m \neq n, \\ 1, & \text{if } m = n. \end{cases}$$

The standard method¹ of constructing orthonormal functions, while completely satisfying logically, has certain practical disadvantages. For example, if the integrations must be done numerically (as would be necessary if either the f_i or the boundary of R were complicated functions, or if the f_i were tabular functions) then the mere tabulation of the intermediate functions which appear becomes burdensome. One would prefer to perform the necessary integrations on the original functions f_i and then proceed by a purely algebraic or numerical process to obtain the g_i . This can be done. If we let N_i be the numerator and D_i the denominator of the orthonormal function g_i , and if we put $F_{ij} = \int_R f_i \bar{f}_j dR$, then the standard orthonormalization process can be shown, by simple algebra, to result in the following:

$$\begin{aligned} N_1 &= f_1, & D_1^2 &= F_{11}, \\ N_2 &= \begin{vmatrix} F_{11} & f_1 \\ F_{21} & f_2 \end{vmatrix}, & D_2^2 &= F_{11} \cdot \begin{vmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{vmatrix}, \end{aligned}$$

Presented to the Society, September 13, 1943; received by the editors August 18, 1943, and, in revised form, December 8, 1943.

¹ Courant and Hilbert, *Methoden der mathematischen Physik*, vol. 1, p. 41.

$$N_3 = \left\| \begin{array}{cc|cc} F_{11} & F_{12} & F_{11} & f_1 \\ F_{21} & F_{22} & F_{21} & f_2 \\ \hline F_{11} & F_{13} & F_{11} & f_1 \\ F_{21} & F_{23} & F_{31} & f_3 \end{array} \right\|,$$

$$D_3^2 = F_{11} \cdot \left\| \begin{array}{cc|cc} F_{11} & F_{12} & F_{11} & F_{13} \\ F_{21} & F_{22} & F_{21} & F_{23} \\ \hline F_{11} & F_{13} & F_{11} & F_{13} \\ F_{21} & F_{23} & F_{31} & F_{33} \end{array} \right\|.$$

The next quantities N_4 and D_4 would each occupy eight lines; N_5 and D_5 sixteen lines, and so on. The quantity N_4 would be a second order determinant whose elements would also be second order determinants, which in turn would have as elements second order determinants whose elements would be the F_{ij} and f_i . We observe that the construction of N_i and D_i reduces to the successive evaluation of second order determinants. All determinants are obtained from a matrix of the original quantities F_{ij} and f_i by a uniform and simple procedure which is described later as the "process P ."

The present paper generalizes the above observation into a technique which (a) provides a definite arrangement of work which minimizes repetition of symbols and lessens the chance of computational errors, (b) can be used by a person whose mathematical training extends no further than a knowledge of the algebraic rules of sign, (c) is adapted to the use of modern computing machines.

To carry through the proof of the method, it is necessary to introduce a more compact notation than that used above. Therefore the proof is somewhat abstract, so we illustrate the extreme simplicity of the final result by an example, namely, the evaluation of the first four Legendre polynomials. The advantages of the method become more pronounced as the number of orthonormal functions desired increases; but space limitations forbid consideration of more than four.

2. Example. We shall construct the first four orthonormal functions corresponding to the linearly independent set of functions $1, x, x^2, \dots$, and the region $-1 \leq x \leq +1$.

Using the standard notation $(f, g) = \int_R f(x)g(x)dR$ we obtain by integration $(1, 1) = 2.000$, $(1, x) = 0$, $(1, x^2) = 0.667$, $(1, x^3) = 0$, $(x, x) = 0.667$, $(x, x^2) = 0$, $(x, x^3) = 0.400$, $(x^2, x^2) = 0.400$, $(x^2, x^3) = 0$, $(x^3, x^3) = 0.286$. With these values as elements we form a symmetric determinant of order four, and adjoin to it the unit matrix of order four, thus obtaining the following 4×8 Matrix No. 1:

$$\left\| \begin{array}{cccc|cccc} 2.000 & 0 & 0.667 & 0 & 1.000 & 0 & 0 & 0 \\ 0 & 0.667 & 0 & 0.400 & 0 & 1.000 & 0 & 0 \\ 0.667 & 0 & 0.400 & 0 & 0 & 0 & 1.000 & 0 \\ 0 & 0.400 & 0 & 0.286 & 0 & 0 & 0 & 1.000 \end{array} \right\|.$$

Now for every element not in either the first row or first column we can define a determinant of order two by blotting out all rows and all columns except those containing either the element under consideration or the leading element of the matrix. For example, associated with the element 0.400 appearing in the third row and third column is the determinant

$$\left| \begin{array}{cc} 2.000 & 0.667 \\ 0.667 & 0.400 \end{array} \right| = +0.356.$$

All such determinants are evaluated and the quantities so obtained are placed in the same position as their associated elements of Matrix No. 1, thus forming a 3×7 Matrix No. 2:

$$\left\| \begin{array}{ccc|cccc} +1.334 & 0 & +0.800 & 0 & +2.000 & 0 & 0 \\ 0 & +0.356 & 0 & -0.667 & 0 & +2.000 & 0 \\ +0.800 & 0 & +0.571 & 0 & 0 & 0 & +2.000 \end{array} \right\|.$$

This is treated in the same way as Matrix No. 1, thus forming the 2×6 Matrix No. 3:

$$\left\| \begin{array}{cc|cccc} +0.474 & 0 & -0.889 & 0 & +2.667 & 0 \\ 0 & +0.122 & 0 & -1.600 & 0 & +2.667 \end{array} \right\|.$$

In similar fashion we form the 1×5 Matrix No. 4:

$$\left\| +0.058 \mid 0 \quad -0.759 \quad 0 \quad +1.264 \right\|.$$

The r th orthonormal function can now be written down as a fraction whose denominator is the square root of the product of the leading elements of all the matrices up to and including the r th matrix, and whose numerator is a polynomial in the original functions $1, x, x^2, \dots$, with coefficients the last n elements in the top row of the r th matrix. The first four are:

$$g_1(x) = \frac{1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3}{(2.000)^{1/2}} = 0.707,$$

$$g_2(x) = \frac{0 \cdot 1 + 2.000x + 0 \cdot x^2 + 0 \cdot x^3}{(2.000 \times 1.334)^{1/2}} = 1.222x,$$

$$g_3(x) = \frac{(-0.889) \cdot 1 + 0 \cdot x + 2.667x^2 + 0 \cdot x^3}{(2.000 \times 1.334 \times 0.474)^{1/2}} = -0.790 + 2.37x^2,$$

$$g_4(x) = \frac{0 \cdot 1 + (-0.759)x + 0 \cdot x^2 + 1.264x^3}{(2.000 \times 1.334 \times 0.474 \times 0.058)^{1/2}} = -0.280x + 4.67x^3.$$

In actual computation the original matrix is laid out on a large sheet with sufficient space to enter the new quantities as they are determined, so that when the computation is finished all n matrices appear superimposed upon the same sheet. Two L -shaped guides prevent errors in picking out the elements of each determinant. It is seen that the unit operation is the evaluation of a second order determinant, which operation requires but slight mathematical knowledge. This operation can be performed as a single step on most modern computing machines without writing down the intermediate products.

3. **The process P .** Let us be given an n th order determinant

$$D^1 = |d_{p,q}^1|, \quad p, q = 1, \dots, n,$$

where the element $d_{p,q}^1$ lies in the p th row and q th column. We define the process P by the functional equation

$$P(d_{p,q}^1) = d_{p,q}^2, \quad p, q = 2, \dots, n,$$

where

$$d_{p,q}^2 = \begin{vmatrix} d_{1,1}^1 & d_{1,q}^1 \\ d_{p,1}^1 & d_{p,q}^1 \end{vmatrix}.$$

Notice that $P(d_{p,1}^1)$ and $P(d_{1,q}^1)$ are not defined. Suppressing these undefined quantities we form a determinant of order $n-1$

$$D^2 = |d_{p,q}^2|, \quad p, q = 2, \dots, n.$$

Ordinarily, we would regard this determinant as possessing rows and columns numbered $1, 2, \dots, (n-1)$. It is more convenient for our present purpose to regard the first row and first column as being absent, so we agree instead that D^2 contains rows and columns numbered $2, 3, \dots, n$. With this convention, we complete the definition of D^2 by saying that the element $d_{p,q}^2$ lies in the p th row (row numbered p) and q th column (column numbered q) of the determinant D^2 .

In a similar way we define

$$P^2(d_{p,q}^1) = P(d_{p,q}^2) = \begin{vmatrix} d_{2,2}^2 & d_{2,q}^2 \\ d_{p,2}^2 & d_{p,q}^2 \end{vmatrix} = d_{p,q}^3, \quad p, q = 3, \dots, n,$$

and, in general:

$$P^{r-1}(d_{p,q}^1) = P(d_{p,q}^{r-1}) = \begin{vmatrix} d_{r-1,r-1}^{r-1} & d_{r-1,q}^{r-1} \\ d_{p,r-1}^{r-1} & d_{p,q}^{r-1} \end{vmatrix} = d_{p,q}^r, \quad p, q = r, \dots, n.$$

Also:

$$P^{r-1}(D^1) = D^r = |d_{p,q}^r|, \quad p, q = r, \dots, n.$$

Notice that $P(d_{p,s}^r)$ and $P(d_{s,q}^r)$ are not defined, when $s < r$. We agree that D^r is of order $n-r+1$ and contains rows and columns numbered $r, r+1, \dots, n$, the element $d_{p,q}^r$ lying in the p th row (row numbered p) and the q th column (column numbered q). This convention as to the suppression of undefined quantities and the numbering of elements will be extended to all determinants used in this paper.

Suppose now that the elements in the n th column of D^1 are integrable functions defined over a given region R , while all other elements are constants. Denote these functions by

$$d_{p,n}^1 = f_p, \quad 1 \leq p \leq n.$$

Let g be an arbitrary integrable function defined over R . Then, by the simple properties of determinants, it follows easily that

- (a) The elements in the n th column of D^r are linear combinations of the functions f_i ,
- (b) $\int_{RG} \cdot D^r dR$ differs from D^r only in that every f_i appearing in D^r is replaced by the quantity $\int_{RG} \cdot f_i dR$,
- (c) $\int_{RG} \cdot P^r(D^1) dR = P^r(\int_{RG} \cdot D^1 dR)$.

4. Proof of the method. Let us be given a linearly independent set of complex functions of real variables defined in a given region R , and integrable over R :

$$f = f_1, f_2, \dots, f_n, \dots$$

Form the set of conjugate complex functions:

$$\bar{f} = \bar{f}_1, \bar{f}_2, \dots, \bar{f}_n, \dots$$

Define the three determinants of order n :

$$D^1 = |d_{p,q}^1|, \quad d_{p,q}^1 = \int_R f_p \bar{f}_q dR, \quad p, q = 1, \dots, n,$$

$$F^1 = |f_{p,q}^1|, \quad f_{p,q}^1 = d_{p,q}^1, \quad q \neq n, \quad p, q = 1, \dots, n,$$

$$f_{p,n}^1 = f_p,$$

$$\bar{F}^1 = |\bar{f}_{p,q}^1|, \quad \bar{f}_{p,q}^1 = d_{p,q}^1, \quad p \neq n, \quad p, q = 1, \dots, n,$$

$$\bar{f}_{n,q}^1 = \bar{f}_q.$$

Define $F^r = P^{r-1}(F^1)$ and $\bar{F}^r = P^{r-1}(\bar{F}^1)$. Define $f_{p,q}^r$ as the element lying in the p th row and q th column of F^r , and $\bar{f}_{p,q}^r$ as the element lying in the p th row and q th column of \bar{F}^r .

We now prove in succession a set of results to culminate in our main theorem, Theorem J, and its corollary.

THEOREM A. $\int_R f_{p,n}^r \cdot \bar{f}_r dR = d_{p,r}^r$, if $r \leq p \leq n$.

PROOF. $\int_R f_{p,n}^r \cdot \bar{f}_r dR = \int_R P^{r-1}(f_{p,n}^1) \cdot \bar{f}_r dR = P^{r-1}(\int_R F_p \cdot \bar{f}_r dR) = P^{r-1}D_{p,r}^1 = d_{p,r}^r$, by (c) of §3 and definitions.

THEOREM B. $\int_R f_{p,n}^{r+1} \cdot \bar{f}_r dR = 0$, if $r+1 \leq p \leq n$.

PROOF:

$$\int_R f_{p,n}^{r+1} \cdot \bar{f}_r dR = \int_R \begin{vmatrix} d_{r,r}^r & f_{r,n}^r \\ d_{p,r}^r & f_{p,n}^r \end{vmatrix} \cdot \bar{f}_r dR,$$

$$= \begin{vmatrix} d_{r,r}^r \int_R f_{r,n}^r \cdot \bar{f}_r dR \\ d_{p,r}^r \int_R f_{p,n}^r \cdot \bar{f}_r dR \end{vmatrix} = 0,$$

since, by Theorem A, the last column is identical to the first.

THEOREM C. $\int_R f_{p,n}^r \cdot \bar{f}_s dR = 0$, if $s < r \leq p \leq n$.

PROOF. Assume the theorem true for $r=k$. Then

$$\int_R f_{p,n}^{k+1} \cdot \bar{f}_s dR = \begin{vmatrix} f_{k,k}^k \int_R f_{k,n}^k \cdot \bar{f}_s dR \\ f_{p,k}^k \int_R f_{p,n}^k \cdot \bar{f}_s dR \end{vmatrix}.$$

By hypothesis the elements in the last column are zeros, and thus the theorem holds for $k+1$. By Theorem B it holds for $r=s+1$, hence it holds for all r satisfying the conditions of the theorem.

THEOREM D. $\int_R f_{n,n}^n \cdot \bar{f}_s dR = 0$ if $s < n$ and equals $d_{n,n}^n$ if $s=n$.

This follows from Theorems A and C by placing $r=p=n$.

THEOREM E. *An analogue to each of the previous theorems is obtained by taking conjugates of each side of the respective equations.*

THEOREM F. $\int_R f_{n,n}^n \cdot \bar{f}_{n,n}^n dR = \prod_{s=1}^n d_{s,s}^s.$

PROOF. Write

$$\bar{f}_{n,n}^n = \sum_{s=1}^n A_s \bar{f}_s, \quad A_s = \text{constants.}$$

Then

$$\int_R f_{n,n}^n \sum_{s=1}^n A_s \bar{f}_s dR = \sum_{s=1}^n A_s \int_R f_{n,n}^n \cdot \bar{f}_s dR = \sum_{s=1}^n A_s Q_s,$$

where

$$Q_s = \begin{cases} 0, & \text{if } s \neq n, \\ d_{n,n}^n, & \text{if } s = n. \end{cases}$$

We see that if we replace \bar{f}_s by Q_s we change $\bar{f}_{n,n}^n$ into $\int_R f_{n,n}^n \cdot \bar{f}_{n,n}^n dR$. Suppose we make this replacement in the determinant \bar{F}^1 forming a new determinant \bar{Q}^1 . Then since $P^{n-1}(\bar{F}^1) = \bar{F}^n$ it follows that $P^{n-1}(\bar{Q}^1) = \bar{Q}^n$ where the element $\bar{q}_{n,n}^n = \int_R f_{n,n}^n \cdot \bar{f}_{n,n}^n dR$ of \bar{Q}^n corresponds to the element $\bar{f}_{n,n}^n$ of \bar{F}^n . Consider \bar{Q}^1 . It is identical to \bar{D}^1 except that all the elements in the bottom row, except the one in the last column, are replaced by zeros. Therefore the process P applied to $\bar{q}_{n,n}^1$ of this determinant reduces to multiplication. We have

$$\begin{aligned} P(\bar{q}_{n,n}^1) &= (d_{n,n}^n)(d_{1,1}^1), \\ P^2(\bar{q}_{n,n}^1) &= (d_{n,n}^n)(d_{1,1}^1)(d_{2,2}^2), \\ &\dots \dots \dots \dots \dots \dots \dots, \\ P^{n-1}(\bar{q}_{n,n}^1) &= \prod_{s=1}^n d_{s,s}^s. \end{aligned}$$

THEOREM G. *Suppose that in addition to the determinants already mentioned we have a determinant \bar{F}^{o1} defined exactly the same as \bar{F}^1 except that it is of order m where $m \neq n$. Then*

$$\int_R f_{n,n}^n \cdot \bar{f}_{m,m}^{om} dR = 0, \quad m \neq n.$$

PROOF. Suppose $m < n$. Then

$$\begin{aligned} \bar{f}_{m,m}^{om} &= \sum_{s=1}^m A_s^o \bar{f}_s, \quad A_s^o = \text{constants,} \\ \int_R f_{n,n}^n \sum_{s=1}^m A_s^o \bar{f}_s dR &= \sum_{s=1}^m A_s^o \int_R f_{n,n}^n \cdot \bar{f}_s dR = 0, \end{aligned}$$

by Theorem D.

If $m > n$ we expand $f_{n,n}^n$ as a linear combination of the functions $f_s, s = 1, 2, \dots, n$, and use the conjugate theorem of Theorem D.

THEOREM H. *The function*

$$g_n = f_{n,n}^n / \left(\prod_{s=1}^n d_{s,s}^s \right)^{1/2}$$

is the n th orthonormal function for the given set f and the given domain R .

This follows immediately from Theorems F and G.

THEOREM J. *Let us adjoin the unit matrix of order n to the determinant D^1 and apply the process P^{n-1} to the combined matrix. Form a matrix whose first row is the top row of the unit matrix, whose second row is the top row of the first transform of the unit matrix, and, in general, whose r th row is the top row of the $(r-1)$ st transform of the unit matrix. Let this matrix be denoted by*

$$\begin{vmatrix} B_{1,1} & B_{1,2} & \cdots & B_{1,n} \\ B_{2,1} & B_{2,2} & \cdots & B_{2,n} \\ \cdot & \cdot & \cdots & \cdot \\ B_{n,1} & B_{n,2} & \cdots & B_{n,n} \end{vmatrix}.$$

Then the n th orthonormal function is given by:

$$g_n = \frac{B_{n,1}f_1 + B_{n,2}f_2 + \cdots + B_{n,n}f_n}{(d_{1,1}^1)(d_{2,2}^2) \cdots (d_{n,n}^n)}.$$

PROOF. We know $f_{n,n}^n = k_{n,1}f_1 + k_{n,2}f_2 + \cdots + k_{n,n}f_n$, where the $k_{n,s}$ are constants. To evaluate $k_{n,s}$ we put $f_s = 1$ and all other $f = 0$ and see what happens to $f_{n,n}^n$. Suppose we make these replacements in F^1 forming a new determinant K^1 , that is, we replace the n th column of F^1 by the s th column of the unit matrix. Then $P^n(F^1)$ is identical to $P^n(K^1)$ except that its n th column is replaced by the s th column of the transformed unit matrix. In particular the element $f_{n,n}^n$ is replaced by the quantity $B_{n,s}$. Hence $k_{n,s} = B_{n,s}$ and the theorem is proved.

COROLLARY. *In Theorem J it is also true that*

$$g_r = \frac{B_{r,1}f_1 + B_{r,2}f_2 + \cdots + B_{r,n}f_n}{(d_{1,1}^1)(d_{2,2}^2) \cdots (d_{r,r}^r)},$$

where $r < n$.

PROOF. Apply Theorem J for the case $n=r$. We obtain for g_r an expression which differs from the one just written only in the fact that the terms $B_{r,r+1}f_{r+1} + B_{r,r+2}f_{r+2} + \dots + B_{r,n}f_n$ are missing from its numerator. But the coefficients $B_{r,s}=0$ when $r < s$. Hence the two expressions are equal.

REMARK. The generalization of the method to orthonormalization with respect to a general norming or weight function p is obvious. One applies the process described to the functions $(p)^{1/2}f_i$ and obtains functions g_i (\equiv linear combinations of the $(p)^{1/2}f_i$) which are orthonormal with respect to the weight function unity. Dividing through by the common factor $(p)^{1/2}$ one forms functions $g_i(p)^{-1/2}$ which are orthonormal with respect to p .

CARNEGIE INSTITUTE OF TECHNOLOGY

VERTICES OF PLANE CURVES

S. B. JACKSON

1. Introduction. The Four-Vertex Theorem, proved first by Mukhopadhyaya [1],¹ states that on any oval, not a circle, there are at least four vertices, that is, extrema of the curvature. This result was extended by Fog [2] and Graustein [3] to any simple closed curve with continuous curvature. The discussion by Graustein makes it clear that the Four-Vertex Theorem is valid also for a very large number of non-simple curves. Indeed the class of curves having only two vertices is relatively quite small. The main object of the present paper is to characterize geometrically, as far as possible, the curves with just two vertices. It is thus a proof of the Four-Vertex Theorem by exclusion.

Since a curve with just two vertices consists of two arcs of monotone curvature, a study is made of such arcs (§2). The most useful fact is that this monotone character of an arc is invariant under direct circular transformations. The property that a point be a vertex of a curve is similarly invariant. This makes it possible to simplify many of the discussions by suitably chosen transformations. Monotone arcs are found to be essentially simple and possess a spiral character.

The existence of vertices on certain types of arcs is established (§4)

Presented to the Society, April 29, 1944; received by the editors February 15, 1944.

¹ Numbers in brackets refer to the Bibliography at the end of the paper.