

# Construction of Gauss-Christoffel Quadrature Formulas

By Walter Gautschi\*

**1. Introduction.** Let  $w(x)$  be a given function ("weight function") defined on a finite or infinite interval  $(a, b)$ . Consider a sequence of quadrature rules

$$(1.1) \quad \int_a^b f(x)w(x)dx \doteq \sum_{r=1}^n \lambda_r^{(n)} f(\xi_r^{(n)}), \quad n = 1, 2, 3, \dots$$

Each of these rules will be called a *Gauss-Christoffel quadrature formula* if it has maximum degree of exactness, i.e. if (1.1) is an exact equality whenever  $f$  is a polynomial of degree  $2n - 1$ . It is a well-known fact, due to Christoffel [3], that such quadrature formulas exist uniquely, provided the weight function  $w(x)$  is nonnegative, integrable with  $\int_a^b w(x)dx > 0$ , and such that all its moments

$$(1.2) \quad \mu_k = \int_a^b x^k w(x)dx, \quad k = 0, 1, 2, \dots,$$

exist. Then, moreover,  $\xi_r^{(n)} \in (a, b)$ , and  $\lambda_r^{(n)} > 0$ . If  $w(x)$  is not of constant sign, Gauss-Christoffel formulas still exist if certain Hankel determinants in the moments are different from zero [21]. In this case, however, some of the abscissas  $\xi_r^{(n)}$  may fall outside the interval  $(a, b)$ ; in particular, they may become complex. We shall call  $\xi_r^{(n)}$  the *Christoffel abscissas*, and  $\lambda_r^{(n)}$  the *Christoffel weights* associated with the weight function  $w(x)$ .

Gauss [7] originally considered the case  $w(x) \equiv 1$  on  $[-1, 1]$ . Other classical cases are associated with the names of Jacobi, Laguerre, and Hermite. In more recent times, the subject has experienced a considerable resurgence, as is evidenced by the appearance of numerous numerical tables [15], [21], both relative to classical and nonclassical weight functions. The emergence of powerful high-speed computers, undoubtedly, has been a major force in this development. Curiously enough, the constructive (algorithmic) aspect of the subject, until very recently, has remained at the state of development in which it was left by Christoffel, and Stieltjes [20]. The generally recommended procedure still consists [1] in constructing the system  $\{\pi_r\}$  of orthogonal polynomials associated with the weight function  $w(x)$ , and to obtain  $\xi_r^{(n)}$  as the zeros of  $\pi_n$ , and  $\lambda_r^{(n)}$ , in a number of possible ways, in terms of these orthogonal polynomials. An alternative procedure, suggested by Rutishauser [19], makes use of the quotient-difference algorithm, while Golub and Welsch [11] use Francis' *QR*-transformations to compute  $\xi_r^{(n)}$  as eigenvalues of a Jacobi matrix and  $\lambda_r^{(n)}$  as the first component of the corresponding eigenvectors. These methods, as interesting as they are, appear to be computationally feasible, for large  $n$ , only if the orthogonal polynomials  $\pi_r$ , or the associated Stieltjes continued fraction, are explicitly known. Otherwise, they are subject to severe numerical instability,

---

Received September 11, 1967.

\* This work was performed in part at the Argonne National Laboratory, Argonne, Illinois, under the auspices of the U. S. Atomic Energy Commission.

making it virtually impossible to obtain meaningful answers, unless one resorts to multiple-precision work.

The reason for this is the ill-conditioned character of the problem which these methods attempt to solve. The problem, basically, is the purely algebraic one of deriving  $\xi_r^{(n)}$ ,  $\lambda_r^{(n)}$  from the first  $2n$  moments of  $w(x)$ , i.e. of solving the algebraic system of equations

$$(1.3) \quad \sum_{r=1}^n \lambda_r^{(n)} [\xi_r^{(n)}]^k = \mu_k \quad (k = 0, 1, 2, \dots, 2n - 1).$$

It will be shown (Section 2) that for a finite interval  $(0, 1)$  the (asymptotic, relative) condition number  $\kappa_n$  for this problem can be estimated from below by

$$(1.4) \quad \kappa_n > \min \left( \mu_0, \frac{1}{\mu_0} \right) \max_{1 \leq r \leq n} \left\{ (1 + \xi_r^{(n)}) \prod_{k=1; k \neq r}^n \left( \frac{1 + \xi_k^{(n)}}{\xi_r^{(n)} - \xi_k^{(n)}} \right)^2 \right\}.$$

Considering that the abscissas  $\xi_r^{(n)}$ , for large  $n$ , tend to cluster near the endpoints of the interval  $(0, 1)$ , many of the differences  $\xi_r^{(n)} - \xi_k^{(n)}$  will be quite small in absolute value. Consequently, some of the products in (1.4), and thus the lower bound for  $\kappa_n$ , are likely to be very large when  $n$  is large.

To give a more concrete idea of just how large  $\kappa_n$  may become, we note [22, p. 309] that for a wide class of weight functions the abscissas  $\xi_r^{(n)}$  ultimately (as  $n \rightarrow \infty$ ) assume an arc cos-distribution, i.e.

$$(1.5) \quad \xi_r^{(n)} \doteq \frac{1}{2}(1 + \cos \theta_r^{(n)}), \quad \theta_r^{(n)} = (2r - 1)\pi/2n.$$

Replacing the  $\xi_r^{(n)}$  in (1.4) by their approximate values in (1.5), one finds that

$$(1.6) \quad \kappa_n > \min \left( \mu_0, \frac{1}{\mu_0} \right) \frac{(17 + 6\sqrt{8})^n}{64n^2} > \min \left( \mu_0, \frac{1}{\mu_0} \right) \frac{(33.97)^n}{64n^2}.$$

Numerical values of the lower bound in (1.6), for  $\mu_0 = 1$  and a few selected values of  $n$ , are shown in Table 1.

TABLE 1  
Lower bound for condition number  $\kappa_n$

$n$	$(33.97)^n/64n^2$
5	$2.8 \times 10^4$
10	$3.2 \times 10^{11}$
15	$6.4 \times 10^{18}$
20	$1.6 \times 10^{26}$

It is thus seen that in the presence of rounding errors the above-mentioned methods, if they rely on the moments, must be expected to suffer a loss of at least 11 decimal digits, if  $n = 10$ , and a loss of 26 digits, when  $n = 20$ . This is well above the attrition level one is normally willing to accept!

The lesson to be learned from this analysis is evident: the moments are not suitable, as data, for constructing Gauss-Christoffel quadrature formulas of large order  $n$ . Apart from the fact that they are not always easy to compute, small changes in the moments (due to rounding, for example) may result in very large changes in the Christoffel numbers.

In Section 3 we propose an alternative procedure for generating Gauss-Christoffel formulas, which is based on a suitable discretization of the inner product  $(f, g) = \int_a^b f(x)g(x)w(x)dx$ , and thus bypasses the moments altogether. As the discretization is made infinitely fine, the process converges to the desired Christoffel numbers provided the singularities of  $w(x)$ , if any, are located at the endpoints of the interval and are monotonic. Extensive tests have shown that the method is reasonably accurate, relatively "inexpensive," and requiring only single-precision arithmetic. A computer algorithm (in ALGOL) is to appear in [10].

Cases may arise in which our method converges very slowly. While approximate Christoffel numbers are still obtained, it may be desirable to further improve their accuracy. This can be done by applying Newton's method to a system of equations, equivalent to (1.3), using as initial approximations the approximate Christoffel numbers already obtained. An appropriate procedure for this will be described in Section 4. Unfortunately, this iterative refinement calls for the moments of the weight function, and therefore is of limited practical value, unless one is prepared to use higher-precision work in some preliminary parts of the computation.

The ability to generate Gauss-Christoffel quadrature formulas, as needed, is of considerable practical interest, not only for integrating singular functions, but also for the numerical solution of integral equations and boundary value problems. We also remark that this new capability may well be useful in future systems of "automated numerical analysis," such as the NAPSS system currently under development at Purdue University [18].

In the appendix are collected a few general properties, more or less known, of orthogonal polynomials which are relevant to our discussion in Sections 3, 4.

Extensions of our work seem possible to quadrature formulas of maximum degree of exactness, where some of the abscissas are prescribed, or the quadrature sum involves derivative values as well as function values. Such generalizations, however, will not be considered here.

**2. Condition of the Classical Approach.** In this section we discuss the condition of the problem of solving the system of algebraic equations (1.3). In particular we derive the estimates (1.4) and (1.6) for the asymptotic condition number, and compare them with the condition of inverting Hilbert matrices.

It will be useful, first, to consider the condition of a mapping  $M$ , say, from one normed space  $X$  into another,  $Y$ :

$$M : X \rightarrow Y .$$

Following Rice [17], we define the (relative)  $\delta$ -condition number  $\kappa(\delta)$  of  $M$  at  $x_0 \in X$  by

$$(2.1) \quad \kappa(\delta) \stackrel{\text{def}}{=} \max_{\|h\|=\delta} \frac{\|M(x_0 + h) - Mx_0\|}{\|Mx_0\|} \bigg/ \frac{\delta}{\|x_0\|} .$$

Thus,  $\kappa(\delta)$  represents the maximum amount by which a (relative) perturbation in the space  $X$ , as given by  $\delta/\|x_0\|$ , is magnified under the mapping  $M$ . Since the perturbations to be considered are small (rounding errors!) it is natural to consider the (relative) asymptotic condition number  $\kappa$  of  $M$  at  $x_0$ , as defined by

$$(2.2) \quad \kappa = \lim_{\delta \rightarrow 0} \kappa(\delta),$$

where the existence of the limit, of course, is assumed.

In solving the system of equations (1.3) we are dealing with the mapping  $M: X \rightarrow Y$  of a  $2n$ -dimensional Euclidean space into itself, if we identify  $X$  with the "moment space," and  $Y$  with the space of Christoffel numbers. This mapping is one-to-one in the neighborhood of the exact solution of (1.3). We may write (1.3) in the compact form

$$(2.3) \quad F(y) = x,$$

where  $x^T = (\mu_0, \mu_1, \dots, \mu_{2n-1})$ ,  $y^T = (\lambda_1, \dots, \lambda_n, \xi_1, \dots, \xi_n)$ ,  $F^T = (F_1, F_2, \dots, F_{2n})$ , and

$$(2.4) \quad F_k(y) = \sum_{r=1}^n \lambda_r \xi_r^{k-1} \quad (k = 1, 2, \dots, 2n).$$

The (relative) asymptotic condition number  $\kappa = \kappa_n$  for solving the nonlinear system of equations (2.3) at  $x_0$  is well known to be (cf. [17])

$$(2.5) \quad \kappa_n = \frac{\|x_0\|}{\|y_0\|} \| [F_y(y_0)]^{-1} \|,$$

where  $y_0$  is the solution of  $F(y) = x_0$ , and  $F_y(y)$  denotes the Jacobian matrix of  $F$ . The matrix norm in (2.5) is assumed to be subordinate to the vector norm chosen in  $X$  and  $Y$ . From (2.4) we obtain by a simple computation that

$$(2.6) \quad F_y(y_0) = \Xi \Lambda,$$

where

$$(2.7) \quad \Xi = \begin{bmatrix} 1 & \dots & 1 & 0 & \dots & 0 \\ \xi_1 & \dots & \xi_n & 1 & \dots & 1 \\ \xi_1^2 & \dots & \xi_n^2 & 2\xi_1 & \dots & 2\xi_n \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ \xi_1^{2n-1} & \dots & \xi_n^{2n-1} & (2n-1)\xi_1^{2n-2} & \dots & (2n-1)\xi_n^{2n-2} \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & \lambda_1 & \\ & & & & & \lambda_2 & \\ & & & & & & \ddots \\ & & & & & & & \lambda_n \end{bmatrix}.$$

(For simplicity, we have written  $\xi_r$  for  $\xi_r^{(n)}$ , and  $\lambda_r$  for  $\lambda_r^{(n)}$ .) Hence, by (2.5),

$$(2.8) \quad \kappa_n = \frac{\|x_0\|}{\|y_0\|} \|\Lambda^{-1}\Xi^{-1}\|.$$

We now choose our norms. We take as vector norm  $\|x\| = \max_k |x_k|$ , and correspondingly as matrix norm  $\|A\| = \max_k \sum_r |a_{kr}|$ . We further assume the basic interval to be  $(0, 1)$ , and  $w(x) \geq 0$ . Clearly,  $\|x_0\| \geq \mu_0$ . Since  $\lambda_r > 0$ , and  $\sum_{r=1}^n \lambda_r = \mu_0$ , we have  $\lambda_r < \mu_0$ . Also,  $0 < \xi_r < 1$ . Therefore,

$$\|y_0\| = \max_r (\lambda_r, \xi_r) < \max(1, \mu_0).$$

Moreover, with the matrix norm as defined,

$$\|\Lambda^{-1}\Xi^{-1}\| \geq \min(1, 1/\mu_0) \|\Xi^{-1}\|.$$

It thus follows from (2.8), that

$$\kappa_n > \frac{\mu_0 \min(1, 1/\mu_0)}{\max(1, \mu_0)} \|\Xi^{-1}\|,$$

or, equivalently,

$$(2.9) \quad \kappa_n > \min(\mu_0, 1/\mu_0) \|\Xi^{-1}\|.$$

Further discussion now hinges on obtaining a lower bound for  $\|\Xi^{-1}\|$ , where  $\Xi$  is the matrix in (2.7), a confluent Vandermonde matrix [8].

**THEOREM 2.1.** *Let  $\xi_1, \xi_2, \dots, \xi_n$  be mutually distinct positive numbers, and  $\Xi$  the matrix defined in (2.7). Then*

$$(2.10) \quad u_1 \leq \|\Xi^{-1}\| \leq \max(u_1, u_2),$$

where  $\|\cdot\|$  denotes the maximum row sum norm, and

$$(2.11) \quad u_i = \max_{1 \leq r \leq n} b_r^{(i)} \prod_{k=1; k \neq r}^n \left( \frac{1 + \xi_k}{\xi_r - \xi_k} \right)^2 \quad (i = 1, 2),$$

$$(2.12) \quad b_r^{(1)} = 1 + \xi_r, \quad b_r^{(2)} = \left| 1 + 2\xi_r \sum_{k=1; k \neq r}^n \frac{1}{\xi_r - \xi_k} \right| + 2 \left| \sum_{k=1; k \neq r}^n \frac{1}{\xi_r - \xi_k} \right|.$$

*Proof.* It was shown in [8] that

$$\Xi^{-1} = \begin{bmatrix} A \\ B \end{bmatrix},$$

where  $A = (a_{rs}), B = (b_{rs})$  are  $(n \times 2n)$ -matrices satisfying

$$(2.13) \quad \sum_{s=1}^{2n} |a_{rs}| \leq b_r^{(2)} \prod_{k \neq r} \left( \frac{1 + \xi_k}{\xi_r - \xi_k} \right)^2, \quad \sum_{s=1}^{2n} |b_{rs}| = b_r^{(1)} \prod_{k \neq r} \left( \frac{1 + \xi_k}{\xi_r - \xi_k} \right)^2.$$

Letting

$$\alpha = \max_{1 \leq r \leq n} \sum_{s=1}^{2n} |a_{rs}|, \quad \beta = \max_{1 \leq r \leq n} \sum_{s=1}^{2n} |b_{rs}|,$$

we have by (2.11) and (2.13),  $\alpha \leq u_2, \beta = u_1$ . Now, either  $\alpha \leq \beta$ , or  $\alpha > \beta$ . In the

first case,  $\|\Xi^{-1}\| = \beta = u_1$ , in the second case,  $u_1 < \|\Xi^{-1}\| = \alpha \leq u_2$ . Hence, (2.10) holds in either case, and Theorem 2.1 is proved.

We remark that in the case  $u_1 \geq u_2$  we have  $\|\Xi^{-1}\| = u_1$ .

Applying Theorem 2.1 to (2.9), we obtain

$$(2.14) \quad \kappa_n > \min\left(\mu_0, \frac{1}{\mu_0}\right) \max_{1 \leq r \leq n} \left\{ (1 + \xi_r) \prod_{k \neq r} \left( \frac{1 + \xi_k}{\xi_r - \xi_k} \right)^2 \right\},$$

the result already stated in (1.4).

Using the approximations (cf. (1.5))

$$\xi_r \doteq \frac{1}{2}(1 + x_r), \quad x_r = \cos \theta_r, \quad \theta_r = (2r - 1)\pi/2n,$$

where  $x_r$  are the zeros of the Chebyshev polynomial  $T_n(x)$ , we may estimate

$$(2.15) \quad \begin{aligned} (1 + \xi_r) \prod_{k \neq r} \left( \frac{1 + \xi_k}{\xi_r - \xi_k} \right)^2 &\doteq \frac{1}{2} (3 + x_r) \prod_{k \neq r} \left( \frac{3 + x_k}{x_r - x_k} \right)^2 \\ &= \frac{1}{2(3 + x_r)} \left[ \frac{T_n(3)}{T_n'(x_r)} \right]^2 > \frac{1}{8} \left[ \frac{T_n(3)}{T_n'(x_r)} \right]^2. \end{aligned}$$

We have

$$(2.16) \quad T_n'(x_r) = T_n'(\cos \theta_r) = n \frac{\sin(n\theta_r)}{\sin \theta_r} = (-1)^{r-1} \frac{n}{\sin \theta_r}.$$

Now the maximum in (2.14) is obviously larger than the respective expression evaluated for any fixed  $r = r_0$ . Choosing  $r_0 = [n/2] + 1$ , we obtain in view of (2.15), (2.16)

$$\kappa_n \gtrsim \frac{1}{8} \min\left(\mu_0, \frac{1}{\mu_0}\right) \left[ \frac{c_n T_n(3)}{n} \right]^2, \quad \begin{aligned} c_n &= 1 && (n \text{ odd}), \\ c_n &= \cos(\pi/2n) && (n \text{ even}). \end{aligned} *$$

Since  $\cos(\pi/2n) \geq 1/\sqrt{2}$  ( $n \geq 2$ ), it follows that  $c_n \geq 1/\sqrt{2}$ , and so

$$\kappa_n \gtrsim (1/16n^2) \min(\mu_0, 1/\mu_0) [T_n(3)]^2.$$

As is well known,  $z_n = T_n(3)$  satisfies

$$z_{n+1} - 6z_n + z_{n-1} = 0, \quad z_0 = 1, \quad z_1 = 3.$$

Hence, using standard results from the theory of linear difference equations,

$$z_n = T_n(3) = \frac{1}{2}(t_1^n + t_2^n), \quad t_1 = 3 + \sqrt{8}, \quad t_2 = 3 - \sqrt{8}.$$

It follows that  $T_n(3) > \frac{1}{2}t_1^n$ , and we finally obtain

$$(2.17) \quad \kappa_n \gtrsim \min\left(\mu_0, \frac{1}{\mu_0}\right) \frac{(17 + 6\sqrt{8})^n}{64n^2},$$

the result already stated in (1.6).

We note from (2.17) that  $\kappa_n$  grows at least at a rate essentially equal to  $\exp[n \ln(17 + 6\sqrt{8})] = \exp(3.5255 \cdots n)$ . Surprisingly, this coincides with the rate of growth of the (Turing) condition number for the  $n$ th order segment of

---

\* We use the symbol  $\gtrsim$  to remind the reader that we are now dealing with an *approximate* lower bound.

the Hilbert matrix, as estimated by Todd [23]. Computing Christoffel numbers on the interval (0, 1) from given moments is therefore about as ill-conditioned as the inversion of Hilbert matrices!

**3. Computation of Christoffel Numbers by Orthogonal Polynomials of a Discrete Variable.** We begin with the classical construction due to Christoffel. We introduce the inner product

$$(3.1) \quad (f, g) = \int_a^b f(x)g(x)w(x)dx,$$

and let  $\{\pi_r\}_{r=0}^\infty$  denote the associated orthonormal polynomials (cf. Example 1 of the appendix),

$$(3.2) \quad (\pi_r, \pi_s) = \delta_{rs}, \quad \text{degree}(\pi_r) = r.$$

Let  $\xi_r^{(n)}$  be the zeros of  $\pi_n(x)$  in (say) increasing order. Then  $\xi_r^{(n)}$  are precisely the Christoffel abscissas corresponding to the weight function  $w(x)$ . The Christoffel weights can be found, e.g., from

$$(3.3) \quad \lambda_r^{(n)} = \frac{1}{\sum_{k=0}^{n-1} [\pi_k(\xi_r^{(n)})]^2}.$$

This representation is particularly suitable for computation since it involves the summation of positive terms.

It seems appropriate, at this point, to distinguish two cases:

(a) The polynomials  $\{\pi_r\}$  are known explicitly, i.e. either the coefficients of  $\pi_r(x)$ , or the coefficients in the three-term recurrence relation [cf. (A.7)], \* are known in closed form. We may refer to this as the *classical case*, and call the corresponding weight functions “classical.” In this case the approach just outlined is entirely satisfactory for computational purposes.

(b) The polynomials  $\{\pi_r\}$  are not explicitly known. We refer to this as the *non-classical case*, and call the corresponding weight functions “nonclassical.” In this case it is necessary to progressively generate either the coefficients of  $\pi_r(x)$ , or the coefficients in the three-term recurrence relation for the  $\pi_r$ . This amounts to an orthogonalization of the successive powers, and hence requires knowledge of the moments of the given weight function. We are therefore in essence solving the ill-conditioned problem discussed in Section 2, and must thus be prepared to encounter severe numerical instability.

The following approach is specifically designed to handle the case of nonclassical weight functions.

Let

$$Q_N(\phi) \stackrel{\text{def}}{=} \sum_{k=1}^N w_k^{(N)} \phi(x_k^{(N)}), \quad w_k^{(N)} > 0, \quad N > n,$$

denote a sequence of auxiliary quadrature formulas with positive weights,

$$(3.4) \quad Q_N(\phi) \doteq \int_a^b \phi(x)dx.$$

---

\* (A.7) refers to formula (7) of the appendix.

We assume first  $(a, b)$  a finite interval, say  $(-1, 1)$  for definiteness. We define a new inner product,

$$(3.5) \quad [f, g]_N \stackrel{\text{def}}{=} Q_N(fgw),$$

that is, more explicitly,

$$(3.5') \quad [f, g]_N = \sum_{k=1}^N W_k^{(N)} f(x_k^{(N)})g(x_k^{(N)}), \quad W_k^{(N)} = w_k^{(N)}w(x_k^{(N)}).$$

Since  $W_k^{(N)} > 0$  (we assume here that  $w(x_k^{(N)}) \neq 0$  for  $k = 1, 2, \dots, N$ ), the inner product (3.5') gives rise to a set  $\{\pi_{r,N}\}_{r=0}^{N-1}$  of orthonormal polynomials of a discrete variable (cf. Example 2 of the appendix),

$$(3.6) \quad [\pi_{r,N}, \pi_{s,N}]_N = \delta_{rs}, \quad r, s = 0, 1, 2, \dots, N - 1.$$

These polynomials may be generated as described in the appendix. The process requires the computation of inner products of the form (3.5'), which in turn requires only a finite summation and the evaluation of  $w(x)$  at the points  $x_k^{(N)}$  (no moments!).

In analogy with the classical approach we now define  $\xi_{r,N}^{(n)}$  to be the zeros of  $\pi_{n,N}(x)$  (known to be real), and let

$$(3.7) \quad \lambda_{r,N}^{(n)} = \frac{1}{\sum_{k=0}^{n-1} [\pi_{k,N}(\xi_{r,N}^{(n)})]^2}.$$

The  $\xi_{r,N}^{(n)}, \lambda_{r,N}^{(n)}$ , suitably ordered, are taken to approximate  $\xi_r^{(n)}, \lambda_r^{(n)}$ , respectively. These approximations depend on the parameter  $N$ , and hopefully converge to the desired Christoffel numbers as  $N \rightarrow \infty$ .

We may now rephrase Theorem 4 of the appendix, and its Corollary, as follows:

**THEOREM 3.1.** *Suppose that  $\lim_{N \rightarrow \infty} [f, g]_N = (f, g)$ , whenever  $f$  and  $g$  are polynomials. Then*

$$(3.8) \quad \lim_{N \rightarrow \infty} \pi_{r,N}(x) = \pi_r(x),$$

and

$$(3.9) \quad \lim_{N \rightarrow \infty} \xi_{r,N}^{(n)} = \xi_r^{(n)}, \quad \lim_{N \rightarrow \infty} \lambda_{r,N}^{(n)} = \lambda_r^{(n)}.$$

Under the assumption of Theorem 3.1, our construction thus yields a convergent process. The stated assumption, in essence, requires that the quadrature rule  $Q_N$  in (3.4) be convergent for integrands of the form  $\phi(x) = p(x)w(x)$ , where  $p(x)$  is a polynomial, and  $w(x)$  is the given weight function. Since  $w(x)$  might be singular, we require, in other words, *convergence of the quadrature rule in the presence of singularities*. Fortunately, most of the common quadrature formulas do converge, even in the presence of singularities, particularly if the singularities occur at the endpoints of the interval and are monotonic [5], [16].

From the computational point of view, convergence alone, while desirable, is far from sufficient. Practical considerations lead us to impose the following additional requirements on the quadrature rules  $Q_N$ :

- (i) Convergence should be reasonably fast, even in the presence of singularities;
- (ii) The quadrature rule  $Q_N$  should be easy to generate for arbitrary, and especially large, values of  $N$ ;



(iii) The interval  $[x_1^{(N)}, x_2^{(N)}, \dots, x_N^{(N)}]$  spanned by  $x_1^{(N)}, \dots, x_N^{(N)}$  should contain the desired Christoffel abscissas  $\xi_1^{(n)}, \dots, \xi_n^{(n)}$ .

The first requirement assures that the value of  $N$ , necessary for given accuracy, is not excessively large. This is important, since the work involved in generating the discrete polynomials  $\pi_{r,N}$  is proportional to  $N$ . The second requirement provides flexibility of the process, and also eliminates the need for storing a large number of high-order quadrature formulas in the computer memory. The third requirement is necessary because of the known fact that the zeros of  $\pi_{n,N}(x)$  are all located in the interval  $[x_1^{(N)}, \dots, x_N^{(N)}]$ . Since these zeros are supposed to approximate the abscissas  $\xi_r^{(n)}$ , these latter had better be contained in that interval!

These, of course, are hard criteria to accommodate. In view of the tendency of the  $\xi_r^{(n)}$  to crowd near the endpoints of  $(-1, 1)$ , requirements (i) and (iii) suggest that we choose the abscissas  $x_k^{(N)}$  to have the same property. This rules out the most common quadrature rules, such as trapezoidal, midpoint, and Simpson rules. The classical Gaussian quadrature formula, on the other hand, is in conflict with requirement (ii). A quadrature rule which comes close to satisfying all the requirements is the Newton-Cotes formula for the abscissas

$$(3.10) \quad x_k^{(N)} = \cos \theta_k^{(N)}, \quad \theta_k^{(N)} = (2k - 1)\pi/2N,$$

the zeros of the Chebyshev polynomial  $T_N(x)$ . The corresponding weight factors  $w_k^{(N)}$  can be written down explicitly, as was already pointed out by Fejér [6]. In fact,

$$(3.11) \quad w_k^{(N)} = \frac{2}{N} \left\{ 1 - 2 \sum_{m=1}^{[N/2]} \frac{\cos(2m\theta_k^{(N)})}{4m^2 - 1} \right\}.$$

This takes care of the requirements (ii) and (iii), although it may be argued that (3.10), (3.11) require the evaluation of a large number of cosines. Actually, only *one* value of the cosine, viz.  $\cos(\pi/2N)$ , is needed, since all the others, both in (3.10) and (3.11), can be generated by well-known recurrence formulas! For best accuracy, however, it is recommended that only the cosines in (3.11) be computed recursively, especially if  $N$  is very large (say, exceeding 200).

As to requirement (i) we have recently shown [9] that the Fejér quadrature formula does indeed converge, not only for continuous functions, but also for singular functions, provided the singularities occur at the endpoints and are monotonic. The exact nature of the singularity is otherwise irrelevant. The rate of convergence, of course, depends on the type of singularity, though in a manner which is not well understood at the present time. Numerical experience indicates that convergence can be rather fast for some singularities (e.g. logarithmic singularities), but discouragingly slow for others (e.g. square-root singularities).

Another quadrature formula, which might be suitable, is the Gauss-Chebyshev formula

$$\int_{-1}^1 \frac{\Phi(x)}{(1-x^2)^{1/2}} dx \doteq \frac{\pi}{N} \sum_{k=1}^N \Phi(x_k^{(N)}),$$

if it is rewritten in the form

$$(3.12) \quad \int_{-1}^1 \phi(x) dx \doteq \frac{\pi}{N} \sum_{k=1}^N (\sin \theta_k^{(N)}) \phi(x_k^{(N)}).$$

Here we have exact equality if  $\phi(x) = p_{2N-1}(x)(1 - x^2)^{-1/2}$ , where  $p_{2N-1}(x)$  is a polynomial of degree  $2N - 1$ . The formula (3.12) is therefore particularly suitable in cases where the weight function  $w(x)$  has square-root singularities at the endpoints  $\pm 1$ , which is one of the cases where the Fejér formula converges very slowly.

It is interesting to point out the close kinship between the Fejér formula (3.10), (3.11) and the Gauss-Chebyshev formula (3.12), noting that the right-hand side in (3.11) is nothing else but the truncated Fourier expansion of  $(\pi/N) \sin \theta_k^{(N)}$ , the weight factor in (3.12)!

We may also remark, at this point, that in the process of generating the polynomials  $\pi_{r,N}$  ( $r = 0, 1, \dots, n$ ), one needs to evaluate inner products  $[f, g]_N$  only for polynomials  $f, g$  of degree  $\leq n$ . Using the Fejér quadrature formula, which is of interpolatory type, it thus follows from (3.1), (3.5) that for such  $f$  and  $g$ ,  $[f, g]_N = (f, g)$  whenever  $w(x)$  is a polynomial of degree  $m$ , and  $N > 2n + m$ . As a result, *our process of constructing Christoffel numbers, based on the Fejér formula (3.10), (3.11), is exact if  $w(x)$  is a polynomial of degree  $m$  and  $N > 2n + m$ . The process, in this case, converges trivially. Similarly, our process of constructing Christoffel numbers, based on the Gauss-Chebyshev formula (3.12), is exact if  $w(x) = (1 - x^2)^{-1/2}$  and  $N > n$ .*

Our development so far assumed  $[-1, 1]$  as the basic interval. This is no restriction of generality. In fact, the case of an arbitrary finite interval  $[a, b]$  is readily reduced to the case considered by a linear transformation of the independent variable. In the case of a half-infinite interval, say  $(0, \infty)$ , let  $\phi(t)$  be any continuously differentiable monotonically increasing function mapping the interval  $(-1, 1)$  onto  $(0, \infty)$ . Then

$$(f, g) = \int_0^\infty f(x)g(x)w(x)dx = \int_{-1}^1 f(\phi(t))g(\phi(t))w(\phi(t))\phi'(t)dt,$$

and we can proceed as before if we define

$$[f, g]_N = \sum_{k=1}^N W_k^{(N)} f(\phi(x_k^{(N)}))g(\phi(x_k^{(N)})),$$

where now

$$W_k^{(N)} = w_k^{(N)}w(\phi(x_k^{(N)}))\phi'(x_k^{(N)}).$$

An analogous device applies for a doubly infinite interval  $(-\infty, \infty)$ , in which case  $\phi(t)$  is to map  $(-1, 1)$  onto  $(-\infty, \infty)$ . Simple transformation functions, which proved satisfactory, are  $\phi(t) = (1 + t)/(1 - t)$  for  $(0, \infty)$ , and  $\phi(t) = t/(1 - t^2)$  for  $(-\infty, \infty)$ .

We conclude this section with a few comments on the computation of the zeros  $\xi_{r,N}^{(n)}$  of  $\pi_{r,N}(x)$ . We assume that the coefficients  $a_r, b_{r+1}$  in the recurrence relation (cf. (A.7\*))

$$\pi_{r+1,N}(x) = ((x - a_r)\pi_{r,N}(x) - b_r\pi_{r-1,N}(x))/b_{r+1} \tag{3.13} \qquad (r = 0, 1, \dots, n - 1),$$

$$\pi_{0,N}(x) = [1, 1]_N^{-1/2}, \quad \pi_{-1,N}(x) = 0,$$

have already been obtained by the methods described in the appendix. We propose two different procedures to find the Christoffel abscissas, depending on whether the

$\xi_r^{(k)}$  are desired for all  $k = 1, 2, \dots, n, r = 1, 2, \dots, k$ , or  $\xi_r^{(n)}, r = 1, 2, \dots, n$ , are desired for only one, or a few selected values of  $n$ .

In the first case we apply Newton's method to each of the equations  $\pi_{k,N}(x) = 0$  ( $k = 2, 3, \dots, n$ ), using  $(\xi_{r-1}^{(k-1)} + \xi_r^{(k-1)})/2$  as initial approximation for  $\xi_r^{(k)}$ . (Here,  $\xi_0^{(k-1)}$  is equal to  $a$ , if  $a$  is finite, or a lower bound for  $\xi_1^{(n)}$ , if  $a = -\infty$ . Similarly  $\xi_k^{(k-1)}$  is equal to  $b$ , if  $b$  is finite, or an upper bound for  $\xi_n^{(n)}$ , if  $b = \infty$ .) The choice of the initial approximation is motivated by the interlacing property of the zeros of  $\pi_{r,N}$  and is normally sufficiently accurate to assure rapid convergence of Newton's iteration. Occasionally, however, because of the highly oscillatory character of the polynomials  $\pi_{r,N}$ , it may happen that some of the Newton iterates fall astray. For this reason it is recommended that each Newton approximation be checked upon whether or not it satisfies the interlacing property. If not, the appropriate subinterval should be examined more carefully for possible zeros, and Newton's iteration repeated with a suitably revised initial approximation.

In the second case, the zeros  $\xi_r^{(n)}$  may be computed in their natural order, using Newton's method in combination with successive deflation. Thus suppose  $\xi_1 = \xi_1^{(n)}$  is already obtained. We then construct the deflated polynomial (we drop the second subscripts  $N$  for notational simplicity)

$$(3.14) \quad \pi_n^{[1]}(x) = (\pi_n(x) - \pi_n(\xi_1))/(x - \xi_1)$$

and compute its smallest zero by Newton's method, using  $\xi_1$  as initial approximation. Thereafter, we deflate again, and compute the smallest zero of the twice deflated polynomial. The process is repeated until all zeros are obtained. We note, that  $\pi_n^{[1]}$  can be obtained by a recurrence relation very similar to (3.13), namely

$$(3.13^1) \quad \begin{aligned} \pi_{r+1}^{[1]}(x) &= (\pi_r(\xi_1) + (x - a_r)\pi_r^{[1]}(x) - b_r\pi_{r-1}^{[1]}(x))/b_{r+1} \\ &\quad (r = 1, 2, \dots, n - 1), \\ \pi_1^{[1]}(x) &= \pi_0/b_1, \quad \pi_0^{[1]} = 0. \end{aligned}$$

This follows readily from (3.13), and the definition (3.14), where  $n$  is to be replaced by  $r$ . (This technique of deflation, in the context of matrices, was already described by Wilkinson [24, p. 468ff.]. He also analyzes its numerical stability.) Similarly, the  $m$ -times deflated polynomial  $\pi_n^{[m]}(x)$  can be generated from

$$(3.13^m) \quad \begin{aligned} \pi_{r+1}^{[m]}(x) &= (\pi_r^{[m-1]}(\xi_m) + (x - a_r)\pi_r^{[m]}(x) - b_r\pi_{r-1}^{[m]}(x))/b_{r+1} \\ &\quad (r = m, m + 1, \dots, n - 1), \\ \pi_m^{[m]}(x) &= \pi_{m-1}^{[m-1]}/b_m, \quad \pi_{m-1}^{[m]}(x) = 0. \end{aligned}$$

To avoid undesirable accumulation of error, it is recommended that each deflation (except the first) be preceded by a "refinement" of the respective zero using Newton's iteration applied to the original (undeflated) polynomial  $\pi_n(x)$ .

It should be noted that the initial approximations to the zeros, if successive deflations are used, are not as accurate as those used in the first procedure (without deflation).

**4. Iterative Refinement of Christoffel Numbers.** We assume now that we have certain approximations  $\xi_r^0, \lambda_r^0$  to the desired Christoffel numbers  $\xi_r^{(n)}, \lambda_r^{(n)}$ , which

are sufficiently accurate to attempt solving the basic system of algebraic equations by Newton's method. The approximations  $\xi_r^0, \lambda_r^0$ , for example, may have been obtained by the procedure discussed in Section 3.

Let  $\{p_r\}_{r=0}^{2n-1}$  be a system of  $2n$  linearly independent polynomials, and define the "modified moments" by

$$(4.1) \quad m_k = \int_a^b p_k(x)w(x)dx .$$

The basic system of equations (1.3) is obviously equivalent to

$$(4.2) \quad \sum_{r=1}^n \lambda_r^{(n)} p_k(\xi_r^{(n)}) = m_k \quad (k = 0, 1, 2, \dots, 2n - 1) .$$

We wish to choose the polynomials  $p_r$  in such a way that the system (4.2), unlike (1.3), is well-conditioned. Ideally, this would be achieved if the Jacobian matrix  $J(\lambda_1, \dots, \lambda_n; \xi_1, \dots, \xi_n)$  of (4.2) evaluated at the exact solution  $\lambda_r = \lambda_r^{(n)}, \xi_r = \xi_r^{(n)}$ , is orthogonal. We shall settle for the next best, which is orthogonality of  $J(\lambda_1^0, \dots, \lambda_n^0; \xi_1^0, \dots, \xi_n^0)$ . Since

$$(4.3) \quad J(\lambda_r; \xi_r) = \begin{bmatrix} p_0(\xi_1) & \dots & p_0(\xi_n) & \lambda_1 p_0'(\xi_1) & \dots & \lambda_n p_0'(\xi_n) \\ p_1(\xi_1) & \dots & p_1(\xi_n) & \lambda_1 p_1'(\xi_1) & \dots & \lambda_n p_1'(\xi_n) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_{2n-1}(\xi_1) & \dots & p_{2n-1}(\xi_n) & \lambda_1 p_{2n-1}'(\xi_1) & \dots & \lambda_n p_{2n-1}'(\xi_n) \end{bmatrix} ,$$

the required orthogonality means that the rows in the matrix (4.3) be mutually orthonormal. In terms of the inner product

$$(4.4) \quad \{f, g\} = \sum_{r=1}^n [f(\xi_r^0)g(\xi_r^0) + (\lambda_r^0)^2 f'(\xi_r^0)g'(\xi_r^0)] ,$$

this in turn implies that

$$(4.5) \quad \{p_r, p_s\} = \delta_{rs} , \quad r, s = 0, 1, \dots, 2n - 1 .$$

We are led to the discrete analogue of Gröbner polynomials, considered in Example 3 of the appendix.

In choosing the polynomials  $p_r$  as described, we not only are achieving a well-conditioned system of algebraic equations, (4.2), but also assure that the linear systems of equations which need to be solved in Newton's method are all well-conditioned. This is so because the first of these is exactly orthogonal, while the remaining ones are nearly orthogonal.

Unfortunately, the modified moments (4.1) are not known in advance, and must be generated, along with the polynomials  $p_r$ . As is shown in the first section of the appendix, we have for  $\{p_r\}$  the recurrence relation

$$(4.6) \quad \begin{aligned} p_{r+1}(x) &= ((x - a_{r,r})p_r(x) - a_{r,r-1}p_{r-1}(x) - \dots - a_{r,0}p_0(x))/b_{r+1} \\ & \qquad \qquad \qquad (r = 0, 1, \dots, 2n - 2) , \\ p_0(x) &= \{1, 1\}^{-1/2} , \end{aligned}$$

where the coefficients  $a_{rs}$  and  $b_{r+1}$  can be computed as described in the appendix. Let us define, then,

$$(4.7) \quad m_{rk} = \int_a^b x^k p_r(x) w(x) dx .$$

We have, in particular,

$$(4.8) \quad m_{0k} = p_0 \mu_k , \quad m_{r0} = m_r ,$$

where  $\mu_k$  are the moments (1.2) of  $w(x)$ . From (4.6) and (4.7) we obtain

$$(4.9) \quad m_{r+1,k} = \left( m_{r,k+1} - \sum_{s=0}^r a_{rs} m_{sk} \right) / b_{r+1} .$$

We may consider  $m_{r,k}$  as entries at grid points of the triangular region  $r \geq 0, k \geq 0, r + k \leq 2n - 1$  in the first quadrant of the  $(r, k)$ -plane. The entries along the vertical boundary of the triangle, by (4.8), are  $p_0 \mu_k$ , which we assume to be known. The relation (4.9) then permits to progressively fill in the triangle, proceeding from left to right. When completed, the entries along the horizontal boundary will be found, which by (4.8) are precisely the modified moments  $m_r$ .

Our process of iterative refinement thus consists of two parts. First, the generation of the orthonormal polynomials  $p_r$  and, along with this, the generation of the modified moments  $m_r$ . Second, the solution of the system of equations (4.2) by Newton's method. Since the whole process (starting, as it does, with the moments  $\mu_k$ ) is unstable, and the second part is stable, we conclude that the first part must be unstable. In practice, therefore, unless  $n$  is small, this part should be carried out with high precision.

**5. Examples.** We select at random some of the possible applications of our procedure to numerical integration, and also point out some of its limitations. The examples, of course, could easily be multiplied. For additional numerical examples we refer to [10].

(a) In the theory of radiative equilibrium of stellar atmospheres one encounters integrals of the form

$$J(\tau) = \frac{1}{2} \int_0^\infty f(t) E_1(|t - \tau|) dt ,$$

$$F(\tau) = 2 \int_\tau^\infty f(t) E_2(t - \tau) dt - 2 \int_0^\tau f(t) E_2(\tau - t) dt ,$$

to evaluate mean intensities and fluxes. Here,  $f(t)$  is a known function, and  $E_m(x) = \int_1^\infty e^{-xt} t^{-m} dt$ , the exponential integral. After a suitable change of variables, one is thus faced with integrals of the form

$$\int_0^\infty f(x) E_m(x) dx , \quad \int_0^\tau f(x) E_m(\tau - x) dx .$$

Since  $E_m(x)$  has a logarithmic singularity at  $x = 0$ , and an essential singularity at  $x = \infty$ , it is natural to treat  $E_m(x)$  and  $E_m(\tau - x)$  as weight functions, and to apply the corresponding Gauss-Christoffel quadrature formulas [2, p. 65ff]. These may be constructed by our procedure of Section 3, both singularities being monotonic. A 20-point formula for  $w(x) = E_1(x), 0 < x < \infty$ , so obtained, may be found in [10].

(b) For the evaluation of Fourier coefficients it may be useful to compute

$$\frac{1}{2\pi} \int_0^{2\pi} f(x) \left(1 - \frac{\cos \alpha x}{\sin \alpha x}\right) dx$$

by Gaussian quadrature treating the trigonometric factor as a weight function [25].

(c) Fourier integrals, such as  $\int_0^\infty f(x) \cos \alpha x \, dx$ , may be treated by Gaussian quadrature, in a manner described in [14]. This calls for  $n$ -point Gauss-Christoffel formulas with weight function  $w(x) = (1 + \cos x)/(1 + x)^{2n+s}$  on  $(0, \infty)$ , where  $s > 0$  is a suitable number, depending on the behavior of  $f(x)$  at  $x = \infty$ .

We have here a case of a nonmonotonic singularity (at  $x = \infty$ ) and thus no theoretical justification for the process of Section 3. The process, accordingly, seems to converge very slowly, if at all. To illustrate, we display below the minimum and maximum relative errors in the abscissas  $\xi_r^{(n)}$  and weights  $\lambda_r^{(n)}$  for the case  $s = 1$ ,  $n = 5$ , and values of  $N$  as shown.

$N$	min. err. $\xi_r^{(5)}$	max. err. $\xi_r^{(5)}$	min. err. $\lambda_r^{(5)}$	max. err. $\lambda_r^{(5)}$
20	.00245	.07907	.00250	.30546
40	.00354	.05818	.00372	.18464
80	.00584	.18406	.00611	.39220
160	.00025	.01658	.00021	.04318
320	.00132	.04617	.00132	.09630

(d) In an attempt to integrate numerically the remainder term in the Euler-Maclaurin sum formula [25], one might use Gauss-Christoffel formulas with weight function  $w(x) = 1/x - [1/x]$  on  $(0, 1)$ . This function has an infinite number of discontinuities, accumulating at  $x = 0$ , and is all but monotonic there. Not surprisingly, our procedure of Section 3 does not seem to converge, not even for  $n$  as small as 5, as may be seen from the following results.

$N$	$\xi_{1,N}^{(5)}$	$\xi_{3,N}^{(5)}$	$\xi_{5,N}^{(5)}$
100	.04756	.47518	.89997
200	.04392	.47103	.89932
400	.04308	.47499	.89983
800	.04510	.47361	.89968

APPENDIX. ORTHOGONAL POLYNOMIALS

We collect here, for easy reference, some elementary properties, computational aspects, and examples of orthogonal polynomials which are useful in the context of Sections 3 and 4.

Consider a (real) linear function space  $S$  containing the powers  $x^r$ ,  $r = 0, 1, 2, \dots, N$ , where  $N$  may be finite or infinite. Designate by  $(\ , \ )$  an inner product in  $S$ . The set of orthogonal polynomials, relative to this inner product, will be denoted by  $\{p_r\}_{r=0}^N$ . Thus,

$$(1) \quad (p_r, p_s) = 0 \text{ for } r \neq s, \quad \text{degree}(p_r) = r.$$

These polynomials are uniquely determined if we require that each  $p_r$  has leading coefficient one. The orthonormal polynomials will be denoted by  $p_r^*$ . We have

$$(2) \quad p_r^*(x) = c_r p_r(x), \quad c_r = (p_r, p_r)^{-1/2}.$$

**1. Recurrence relations.**

**THEOREM 1.** *The orthogonal polynomials in (1), having leading coefficients one, satisfy the recurrence relation*

$$(3) \quad p_{r+1}(x) = (x - a_{r,r})p_r(x) - a_{r,r-1}p_{r-1}(x) - \dots - a_{r,0}p_0(x) \\ (r = 0, 1, 2, \dots, N - 1),$$

where

$$(4) \quad a_{r,s} = (xp_r, p_s)/(p_s, p_s) \quad (s = 0, 1, 2, \dots, r).$$

*Proof.* It is clear that the polynomials defined by (3), and  $p_0(x) = 1$ , have leading coefficients one and correct degrees. A simple computation shows that orthogonality of  $p_0, p_1, \dots, p_r$  implies orthogonality of  $p_0, p_1, \dots, p_{r+1}$ . Since  $p_0$  and  $p_1$  are orthogonal, Theorem 1 follows by induction.

A recurrence relation for the orthonormal polynomials  $p_r^*$  could be obtained in the obvious manner by substituting (2) into (3). Computationally, it is slightly more convenient to introduce

$$(5) \quad \tilde{p}_r(x) = c_{r-1}p_r(x) = c_{r-1}p_r^*(x)/c_r,$$

and to transform (3), (4) into

$$(3^*) \quad \tilde{p}_{r+1}(x) = (x - a_{r,r}^*)\tilde{p}_r(x) - a_{r,r-1}^*\tilde{p}_{r-1}(x) - \dots - a_{r,0}^*\tilde{p}_0(x), \\ p_{r+1}^*(x) = \tilde{p}_{r+1}(x)/(\tilde{p}_{r+1}, \tilde{p}_{r+1})^{1/2},$$

where

$$a_{r,s}^* = (xp_r^*, p_s^*) \quad (s = 0, 1, \dots, r).$$

**THEOREM 2.** *If the inner product satisfies*

$$(6) \quad (xf, g) = (f, xg),$$

then (3) is a three-term recurrence relation, i.e.

$$(7) \quad p_{r+1}(x) = (x - a_r)p_r(x) - b_r p_{r-1}(x) \quad (r = 0, 1, \dots, N - 1),$$

where

$$(8) \quad a_r = (xp_r, p_r)/(p_r, p_r) \quad (r = 0, 1, \dots, N - 1),$$

$$(9) \quad b_r = (xp_r, p_{r-1})/(p_{r-1}, p_{r-1}) = (p_r, p_r)/(p_{r-1}, p_{r-1}) \quad (r = 1, 2, \dots, N - 1).$$

(We adopt the convention, in (7), that  $p_{-1}(x) \equiv 0$ .)

*Proof.* By (6) we have  $(xp_r, p_s) = (p_r, xp_s) = 0$  if  $s < r - 1$ , since  $xp_s$  is a polynomial of degree  $\leq r - 1$ , and  $p_r$  is orthogonal to every polynomial of degree  $< r$ . Consequently, by (4),  $a_{r,s} = 0$  if  $s < r - 1$ , and Theorem 2 is a corollary of Theorem 1. The second expression for  $b_r$  is obtained by noting that  $(xp_r, p_{r-1}) = (p_r, xp_{r-1}) = (p_r, p_r)$ , since  $xp_{r-1}$  differs from  $p_r$  by a polynomial of degree  $< r$ .

We may interpret  $p_r(x)$  of Theorem 2 as the characteristic polynomial  $\det(xI_r - J_r)$  of the symmetric tridiagonal matrix

$$J_r = \begin{bmatrix} a_0 & \sqrt{b_1} & & & & \\ \sqrt{b_1} & a_1 & \sqrt{b_2} & & & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & \cdot & \\ & & & & & \sqrt{b_{r-1}} & a_{r-1} \end{bmatrix}.$$

Since, by the second relation in (9),  $b_r > 0$ , we have that  $J_r$  is a Jacobi matrix. Consequently, as is well known, the polynomials  $\{p_r(x)\}_{r=0}^{N-1}$  have the Sturm sequence property (cf. [24, p. 300]). In particular, the zeros of  $p_r$  separate those of  $p_{r+1}$ .

Using (5), we obtain for the orthonormal polynomials  $p_r^*$  of Theorem 2 the recursion

$$(7^*) \quad \tilde{p}_{r+1}(x) = (x - a_r^*)p_r^*(x) - b_r^*p_{r-1}^*(x), \quad p_{r+1}^*(x) = \tilde{p}_{r+1}(x)/b_{r+1}^*$$

where

$$a_r^* = (xp_r^*, p_r^*), \quad b_r^* = (\tilde{p}_r, \tilde{p}_r)^{1/2}.$$

This form of the recurrence relation is particularly convenient for computation [4, p. 234].

Noting that  $a_r^* = a_r$ ,  $b_r^* = \sqrt{b_r}$ , the Gershgorin circle theorem applied to the Jacobi matrix  $J_n$  permits one to find upper and lower bounds for the zeros of  $p_n(x)$  in terms of the coefficients  $a_r^*$  and  $b_r^*$ .

**2. Examples.**

*Example 1.* Let  $S = C[-1, 1]$ , the class of continuous functions on  $[-1, 1]$  (hence  $N = \infty$ ), and let the inner product be defined by

$$(10) \quad (f, g) = \int_{-1}^1 f(x)g(x)w(x)dx.$$

Here,  $w(x)$  is a weight function assumed to be positive for  $-1 < x < 1$ , and such that all its moments  $\int_{-1}^1 x^r w(x)dx$ ,  $r = 0, 1, 2, \dots$ , exist. The inner product (10) clearly satisfies (6).

The recursion (7) can be used, in principle, to generate the orthogonal polynomials  $p_r(x)$  successively for  $r = 1, 2, 3, \dots$ , starting with  $p_{-1}(x) \equiv 0$ ,  $p_0(x) = 1$ . In practice, this requires the computation of the inner products in (8), (9), which in view of (10) may be problematic, especially if  $w(x)$  is a singular function not of the standard type  $w(x) = (1 - x)^\alpha(1 + x)^\beta$ ,  $\alpha > -1$ ,  $\beta > -1$ . In the latter case,  $p_r$  are the Jacobi polynomials, and the coefficients  $a_r, b_r$  in (7) are known explicitly [22].

*Example 2.* Let  $N = n - 1$  be a fixed positive integer, and  $S$  the set of polynomials of degree  $\leq N$ . Define



$$(11) \quad (f, g) = \sum_{r=1}^n w_r f(x_r)g(x_r) ,$$

where  $w_r, x_r$  are fixed real numbers with  $w_r > 0, x_r \neq x_s$  for  $r \neq s$ . We note that  $S$  is an inner product space, since  $(f, f) = 0$  implies  $f(x_r) = 0$  ( $r = 1, 2, \dots, n$ ), which in turn implies  $f \equiv 0, f$  being a polynomial of degree  $< n$ .

In contrast to Example 1, we now have a finite set of orthogonal polynomials depending on a parameter,  $n$ . To different values of  $n$  correspond different sets of orthogonal polynomials. As (6) is satisfied, these polynomials again obey the relations in (7)–(9). The successive computation of the coefficients  $a_r, b_r$  is now straightforward, since the inner product (11) requires only the evaluation of a finite sum.

*Example 3.* Let  $N = 2n - 1$  be fixed, and  $S$  the set of polynomials of degree  $\leq N$ . Define

$$(12) \quad (f, g) = \sum_{r=1}^n [u_r f(x_r)g(x_r) + v_r f'(x_r)g'(x_r)] ,$$

where  $u_r, v_r, x_r$  are fixed real numbers, with  $u_r > 0, v_r > 0$ . As in Example 2 one shows that  $S$  is an inner-product space. Unlike the previous example, however, the inner product now fails to satisfy (6). As a result, the associated orthogonal polynomials  $p_r$  obey the “long” recurrence relation (3). The coefficients  $a_{r,s}$  appearing in this relation are different from zero, in general, although in special circumstances some of them may vanish (cf. Theorem 3 below).

While it is true that the recurrence relation is now more complicated, it can still be used, as in Example 2, to successively build up the coefficients  $a_{r,s}$ . The inner products required in (4) are readily computed by the finite summation in (12), using for the derivatives the recursion

$$(13) \quad p'_{r+1}(x) = p_r(x) + (x - a_{r,r})p'_r(x) - a_{r,r-1}p'_{r-1}(x) - \dots - a_{r,1}p'_1(x) .$$

We remark that the continuous analogues of the polynomials considered in Example 3 were recently studied by Gröbner [12].

**3. Symmetry Properties.** If  $w(x)$  is an *even* function on  $(-a, a)$ , where  $0 < a \leq \infty$ , then the associated orthogonal polynomials satisfy

$$p_r(x) = (-1)^r p_r(-x) .$$

In particular, the zeros of  $p_r$  are located symmetrically with respect to the origin, and  $x = 0$  is a zero of  $p_r$  if  $r$  is odd.

This property may be used to essentially cut in half the amount of work required to construct the Christoffel numbers for an even weight function. Indeed, the polynomials  $p_{n,e}(x) = p_{2n}(\sqrt{x})$  form a set of orthogonal polynomials relative to the inner product

$$(f, g)_e = \int_0^{a^2} f(x)g(x) \frac{w(\sqrt{x})}{\sqrt{x}} dx .$$

It follows that the Christoffel numbers  $\xi_{r,e}^{(n)}, \lambda_{r,e}^{(n)}$  of  $p_{n,e}$  are related to those of  $p_{2n}$  by

$$\xi_{r,e}^{(n)} = [\xi_r^{(2n)}]^2, \quad \lambda_{r,e}^{(n)} = 2\lambda_r^{(2n)} \quad (r = 1, 2, \dots, n) ,$$

where  $\xi_r^{(2n)}$  are the *positive* zeros of  $p_{2n}$  and  $\lambda_r^{(2n)}$  the corresponding weight factors.

Similarly, the polynomials  $p_{n,0}(x) = (1/\sqrt{x})p_{2n+1}(\sqrt{x})$  are orthogonal with respect to the inner product

$$(f, g)_0 = \int_0^{a^2} f(x)g(x)\sqrt{x}w(\sqrt{x})dx,$$

and their zeros and weight factors are given by

$$\xi_{r,0}^{(n)} = [\xi_r^{(2n+1)}]^2, \quad \lambda_{r,0}^{(n)} = 2\xi_{r,0}^{(n)}\lambda_r^{(2n+1)} \quad (r = 1, 2, \dots, n).$$

Here again  $\xi_r^{(2n+1)}$  denotes the *positive* zeros of  $p_{2n+1}$  and  $\lambda_r^{(2n+1)}$  the corresponding weight factors. Moreover,

$$\int_{-a}^a w(x)dx - \sum_{r=1}^n \lambda_{r,0}^{(n)}/\xi_{r,0}^{(n)} = \lambda_0^{(2n+1)}$$

is the weight factor corresponding to the zero  $\xi_0^{(2n+1)} = 0$  of  $p_{2n+1}$ .

The inner product (12) may be called *equilibrated* if

$$(14) \quad \begin{aligned} x_{n+1-r} &= x_1 + x_n - x_r & (r = 1, 2, \dots, n) \\ u_{n+1-r} &= u_r, \quad v_{n+1-r} = v_r \end{aligned}$$

**THEOREM 3.** *If the inner product (12) is equilibrated, in the sense of (14), then the associated orthogonal polynomials  $p_r$  satisfy*

$$(15) \quad p_r(x_1 + x_n - x) = (-1)^r p_r(x).$$

Moreover, every other coefficient in the recursion (3) is zero, i.e.

$$(16) \quad a_{r,r-2s} = 0 \quad (s = 1, 2, 3, \dots).$$

The proof of Theorem 3 is elementary, and is omitted here.

**4. Discrete vs. Continuous Orthogonal Polynomials.** The orthogonal polynomials of Example 2 may be considered discrete analogues of those in Example 1. It is reasonable to expect that the former approach the latter, as  $n \rightarrow \infty$ , if the inner product in (11) converges to the inner product in (10).

**THEOREM 4.** *Let  $(f, g)$  denote the inner product in (10), and let*

$$(17) \quad [f, g]_n = \sum_{r=1}^n w_r^{(n)} f(x_r^{(n)})g(x_r^{(n)}),$$

where  $w_r^{(n)}$  are positive numbers and  $x_r^{(n)}$ , for each  $n$ , are  $n$  distinct numbers in  $[-1, 1]$ . Let  $\{p_r\}_{r=0}^\infty$  denote the set of orthogonal polynomials associated with (10), and  $\{p_{r,n}\}_{r=0}^{n-1}$  the set of orthogonal polynomials associated with (17). Suppose that

$$(18) \quad \lim_{n \rightarrow \infty} [f, g]_n = (f, g),$$

whenever  $f$  and  $g$  are polynomials. Then for each  $r = 0, 1, 2, \dots$  we have the limit relation

$$(19) \quad \lim_{n \rightarrow \infty} p_{r,n}(x) = p_r(x)$$

for any fixed  $x$ , and thus uniformly for  $x$  in any finite interval.

*Proof.* We begin with the observation that

$$|[f, g]_n| \leq \sum_{r=1}^n w_r^{(n)} \max_{-1 \leq x \leq 1} |f(x)| \cdot \max_{-1 \leq x \leq 1} |g(x)|$$

for any continuous functions  $f, g$ , and therefore

$$(20) \quad |[f, g]_n| \leq \|f\| \|g\| [1, 1]_n.$$

The polynomials  $p_r$ , by Theorem 2, satisfy (7)–(9), while the polynomials  $p_{r,n}$ , by the same theorem, satisfy

$$(21) \quad p_{r+1,n}(x) = (x - a_{r,n})p_{r,n}(x) - b_{r,n}p_{r-1,n}(x),$$

with

$$(22) \quad a_{r,n} = \frac{[xp_{r,n}, p_{r,n}]_n}{[p_{r,n}, p_{r,n}]_n}, \quad b_{r,n} = \frac{[xp_{r,n}, p_{r-1,n}]_n}{[p_{r-1,n}, p_{r-1,n}]_n}.$$

Suppose now that (19) is true for  $r = s$  and  $r = s - 1$ . We want to show that (19) holds for  $r = s + 1$ . For this it suffices to show that

$$(23) \quad a_{s,n} \rightarrow a_s, \quad b_{s,n} \rightarrow b_s \quad (n \rightarrow \infty),$$

since by (21), this implies  $p_{s+1,n}(x) \rightarrow (x - a_s)p_s(x) - b_s p_{s-1}(x) = p_{s+1}(x)$ .

We have

$$(24) \quad [p_{s,n}, p_{s,n}]_n = [p_s + (p_{s,n} - p_s), p_s + (p_{s,n} - p_s)]_n \\ = [p_s, p_s]_n + 2[p_s, p_{s,n} - p_s]_n + [p_{s,n} - p_s, p_{s,n} - p_s]_n.$$

The first term on the right, by (18), has the limit  $(p_s, p_s)$  as  $n \rightarrow \infty$ . To the second term we apply (20), with the result that

$$|[p_s, p_{s,n} - p_s]_n| \leq \|p_s\| \|p_{s,n} - p_s\| [1, 1]_n.$$

Since  $[1, 1]_n \rightarrow (1, 1)$ , and  $p_{s,n} \rightarrow p_s$  (by assumption), we see that the bound on the right tends to zero as  $n \rightarrow \infty$ . By the same reasoning, one shows that the last term in (24) also tends to zero. Consequently,

$$\lim_{n \rightarrow \infty} [p_{s,n}, p_{s,n}]_n = (p_s, p_s).$$

In the same manner, analogous limit relations can be established for all the other inner products appearing in (22), thus proving (23).

Since, trivially,  $p_{0,n} \rightarrow p_0$ ,  $p_{-1,n} \rightarrow p_{-1}$ , the assertion (19) now follows by induction.

Theorem 4 may also be obtained from a general theorem of B. R. Kripke [13] on best approximation with respect to nearby norms, if one observes that  $x^r - p_{r,n}(x)$  and  $x^r - p_r(x)$  are the best approximations to  $x^r$ , from polynomials of degree  $r - 1$ , in the norms of (17) and (10), respectively. The author is indebted to Professor J. R. Rice for this remark.

**COROLLARY.** *Let the zeros of  $p_r(x)$ , in increasing order, be denoted by  $x_1^{(r)}, x_2^{(r)}, \dots, x_r^{(r)}$ , and the zeros of  $p_{r,n}(x)$ , in the same order, by  $x_{1,n}^{(r)}, x_{2,n}^{(r)}, \dots, x_{r,n}^{(r)}$ . Under the assumptions of Theorem 4, we have*

$$(25) \quad \lim_{n \rightarrow \infty} x_{s,n}^{(r)} = x_s^{(r)}, \quad \lim_{n \rightarrow \infty} p_{t,n}(x_{s,n}^{(r)}) = p_t(x_s^{(r)}) \quad (s = 1, 2, \dots, r; t < r).$$

*Proof.* The first relation in (25) follows from the continuity of the zeros of an algebraic equation. The second relation follows from

$$p_{t,n}(x_{s,n}^{(r)}) - p_t(x_s^{(r)}) = [p_{t,n}(x_{s,n}^{(r)}) - p_t(x_{s,n}^{(r)})] + [p_t(x_{s,n}^{(r)}) - p_t(x_s^{(r)})]$$

by observing that  $|p_{t,n}(x_{s,n}^{(r)}) - p_t(x_{s,n}^{(r)})| \leq \max_{-1 \leq x \leq 1} |p_{t,n}(x) - p_t(x)| \rightarrow 0$  ( $n \rightarrow \infty$ ), and  $p_t(x_{s,n}^{(r)}) \rightarrow p_t(x_s^{(r)})$  ( $n \rightarrow \infty$ ).

Computer Sciences Department  
Purdue University  
Lafayette, Indiana 47907

1. D. G. ANDERSON, "Gaussian quadrature formulae for  $\int_0^1 -\ln(x)f(x)dx$ ," *Math. Comp.* v. 19, 1965, pp. 477-481. MR 31 #2826.
2. S. CHANDRASEKHAR, *Radiative Transfer*, Oxford Univ. Press, 1950, Chapter II. MR 13, 136.
3. E. B. CHRISTOFFEL, "Sur une classe particulière de fonctions entières et de fractions continues," *Ann. Mat. Pura Appl.*, (2), v. 8, 1877, pp. 1-10.
4. P. J. DAVIS, *Interpolation and Approximation*, Blaisdell, New York, 1963. MR 28 #393.
5. P. J. DAVIS & P. RABINOWITZ, "Ignoring the singularity in approximate integration," *SIAM J. Numer. Anal.*, v. 2, 1965, pp. 367-383. MR 33 #3459.
6. L. FEJÉR, "Mechanische Quadraturen mit positiven Cotesschen Zahlen," *Math. Z.*, v. 37, 1933, pp. 287-309.
7. C. F. GAUSS, "Methodus nova integralium valores per approximationem inveniendi," *Comment. Soc. Regiae Scientiarum Recentiores*, v. 3, 1816; *Werke*, Vol. 3, pp. 163-196.
8. W. GAUTSCHI, "On inverses of Vandermonde and confluent Vandermonde matrices. II," *Numer. Math.*, v. 5, 1963, pp. 425-430. MR 29 #1734.
9. W. GAUTSCHI, "Numerical quadrature in the presence of a singularity," *SIAM J. Numer. Anal.*, v. 4, 1967, pp. 357-362.
10. W. GAUTSCHI, "Algorithm, Gaussian quadrature formulas," *Comm. ACM.* (To appear.)
11. G. H. GOLUB & J. H. WELSCH, *Calculation of Gauss Quadrature Rules*, Comput. Sci. Dept. Tech. Rep. No. CS 81, Stanford University, Calif., 1967.
12. W. GRÖBNER, "Orthogonale Polynomsysteme die gleichzeitig mit  $f(x)$  auch deren Ableitung  $f'(x)$  approximieren," *Funktionalanalysis, Approximationstheorie, Numerische Mathematik*, edited by L. Collatz, G. Meinardus, and H. Unger, Birkhäuser, Basel, 1967, pp. 24-32.
13. B. R. KRIPKE, "Best approximation with respect to nearby norms," *Numer. Math.*, v. 6, 1964, pp. 103-105. MR 29 #1483.
14. L. G. KRUGLIKOVA & V. I. KRYLOV, "Numerical Fourier transform," *Dokl. Akad. Nauk BSSR*, v. 5, 1961, pp. 279-283. (Russian) MR 26 #886.
15. V. I. KRYLOV & L. T. SUL'GINA, *Handbook on Numerical Integration*, "Nauka," Moscow, 1966. (Russian)
16. P. RABINOWITZ, "Gaussian integration in the presence of a singularity," *SIAM J. Numer. Anal.*, v. 4, 1967, pp. 191-201.
17. J. R. RICE, "A theory of condition," *SIAM J. Numer. Anal.*, v. 3, 1966, pp. 287-310.
18. J. R. RICE & S. ROSEN, "NAPSS—a numerical analysis problem solving system," Proc. ACM 21st Natl. Conf., Los Angeles, Calif. (August 1966), Thompson, Washington, D. C., 1966, pp. 51-56.
19. H. RUTISHAUSER, "On a modification of the QD-algorithm with Graeffe-type convergence," Proc. IFIP Congress 62, pp. 93-96, North-Holland, Amsterdam, 1963.
20. T. J. STIELTJES, "Quelques recherches sur la théorie des quadratures dites mécaniques," *Ann. Sci. École Norm. Sup.*, (3), v. 1, 1884, pp. 409-426; *Oeuvres Complètes*, Vol. I, pp. 377-394.
21. A. H. STROUD & DON SECREST, *Gaussian Quadrature Formulas*, Prentice-Hall, Englewood Cliffs, N. J., 1966. MR 34 #2185.
22. G. SZEGÖ, *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publ., Vol. 23, Amer. Math. Soc., Providence, R. I., 1959. MR 21 #5029.
23. J. TODD, "The condition of the finite segments of the Hilbert matrix," Nat. Bur. Standards Appl. Math. Ser., No. 39, U. S. Government Printing Office, Washington, D. C., 1954, pp. 109-116. MR 16, 861.
24. J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965. MR 32 #1894.
25. I. ZAMFIRESCU, "An extension of Gauss' method for the calculation of improper integrals," *Acad. R. P. Romine Stud. Cerc. Mat.*, v. 14, 1963, pp. 615-631. (Romanian) MR 32 #1906.