Estimates in Quadratic Formulas

Golub, G. H.
1993

Estimates in Quadratic Formulas

Gene H. Golub* Zdeněk Strakoš †

August 9, 1994

Revised version

Abstract

Let A be a real symmetric positive definite matrix. We consider three particular questions, namely estimates for the error in linear systems $Ax = b$, minimizing quadratic functional $\min_x(x^TAx - 2b^Tx)$ subject to the constraint $\| x \| = \alpha$, $\alpha < \| A^{-1}b \|$, and estimates for the entries of the matrix inverse A^{-1}. All of these questions can be formulated as a problem of finding an estimate or an upper and lower bound on $u^TF(A)u$, where $F(A) = A^{-1}$ resp. $F(A) = A^{-2}$, u is a real vector. This problem can be considered in terms of estimates in the Gauß-type quadrature formulas which can be effectively computed exploiting the underlying Lanczos process. Using this approach, we first recall the exact arithmetic solution of the questions formulated above and then analyze the effect of rounding errors in the quadrature calculations. It is proved that the basic relation between the accuracy of Gauß quadrature for $f(\lambda) = \lambda^{-1}$ and the rate of convergence of the corresponding conjugate gradient process holds true even for finite

*Department of Computer Science, Stanford University, Stanford, CA 94305, U.S.A., e-mail: na.golub@na-net.ornl.gov.
The work of this author was supported in part by the NSF under grant NSF CCR-882/078.

†Institute of Computer Science, Academy of Science of the Czech Republic, Pod vodárenskou věží 2, 182 07 Praha 8, Czech Republic, e-mail: na.strakos@na-net.ornl.gov.
A part of this work was performed while visiting the Institute of Mathematics and Its Applications, University of Minnesota, and the Department of Computer Science, Stanford University.

precision computation. This allows us to explain experimental results observed in quadrature calculations and in physical chemistry and solid state physics computations which are based on continued fraction recurrences.

1 Introduction

Throughout the paper, A denotes a real N by N symmetric positive definite (SPD) matrix, $\lambda_1, \lambda_2, \ldots, \lambda_N$ its eigenvalues and $u_1, u_2, \ldots, u_N$ its corresponding orthonormalized eigenvectors, $\Lambda = diag(\lambda_i)$, U is the orthonormal matrix with columns $(u_1, u_2, \ldots, u_N)$. To simplify the notation we assume, without any loss of generality, that the eigenvalues of A are distinct, i.e., $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_N$.

We consider the following problem. Let u be a real vector with N components. We want to estimate

$$u^T F(A) u, \tag{1}$$

where $F(A)$ is a given function of the matrix A. Often one wishes to determine upper and lower bounds on (1), namely.

$$\underline{\kappa} \leq u^T F(A) u \leq \overline{\kappa}. \tag{2}$$

We note that the more general case

$$v^T F(A) w,$$

where v and w are real vectors with N components, $v \neq w$, can be easily converted into the symmetric case (1) using the identity

$$v^T F(A) w = 1/2[v^T F(A) v + w^T F(A) w - (v - w)^T F(A)(v - w)]. \tag{3}$$

In this paper we are particularly interested in the case $F(A) = A^{-1}$ and $F(A) = A^{-2}$. We will give three specific examples below.

Problem (I)

Given a system of linear equations,

$$Ax = b$$

where $b \in R^N$ is an arbitrary right hand side. Let x^0 be an initial approximation to the solution x. We wish to determine an estimate, upper and lower bound for the Euclidean resp. A-norm of the error

$$\| x - x^0 \| \quad \text{resp.} \quad \| x - x^0 \|_A \ .$$

Considering the initial residual $r^0 = b - Ax^0$,

$$\| x - x^0 \|^2 = (r^0)^T A^{-2} r^0, \tag{4}$$

$$\| x - x^0 \|_A^2 = (r^0)^T A^{-1} r^0. \tag{5}$$

Upper and lower bounds on (4) and (5) were developed in [DEG-72], [DGN-78]. Moreover, the approach proposed in [DGN-78] led to determining upper and lower bounds for the A-norm in the k-th step of the conjugate gradient algorithm. This problem is closely related to the convergence of continued fractions.

Problem (II)

Consider the problem of minimizing the quadratic functional

$$\min_x (x^T A x - 2b^T x) \tag{6}$$

subject to the constraint

$$\| x \| = \sigma, \tag{7}$$

where $b \in R^N$ is a given nonzero vector and σ is choosen so that

$$\sigma < \| A^{-1} b \| \ . \tag{8}$$

We use the method of Lagrange multipliers. Defining

$$h(x,\gamma) = x^T A x - 2b^T x + \gamma(x^T x - \sigma^2)$$

we see that the equation $grad_x h(x,\gamma) = 0$ leads to the linear system

$$(A + \gamma I)x = b \tag{9}$$

where the multiplier γ is determined from the condition $x^T x - \sigma^2 = 0$ and this leads to

$$b^T (A + \gamma I)^{-2} b = \sigma^2, \tag{10}$$

i.e.

$$(U^T b)^T (\Lambda + \gamma I)^{-2} (U^T b) = \sigma^2$$

which can be rewritten as

$$\sum_{i=1}^{N} \frac{(u_i, b)^2}{(\lambda_i + \gamma)^2} = \sigma^2. \tag{11}$$

In order to find the Lagrange multiplier γ we need to solve this equation (the condition (8) quarantees that $\gamma > 0$). To construct (11), we must know all the eigenvalues and eigenvectors of A. For A large and sparse, this can be a difficult problem which we prefer to avoid.

In most cases, we do not need to solve (10) precisely. An approximate solution of (10) can be found using an estimate for $b^T (A + \gamma I)^{-2} b$. Problem (II) is considered in detail in [GM-91] where the comprehensive description of solving constrained least squares and of minimizing quadratic forms subject to the constraint $\|x\| = \sigma$ is given. For some details see also [GS-93]. We are not going to repeat these considerations here.

Problem (III)

Finally, in some situations we want to estimate the individual entries of the matrix inverse in terms of the entries of the original matrix. This problem was studied in [RW-92]. The diagonal entries can be written as

$$(A^{-1})_{ii} = e_i^T A^{-1} e_i, \tag{12}$$

where e_i is the i-th unit vector. This is again, a particular example of (1). For the other entries $(A^{-1})_{ij}$, $i \neq j$, one may use the identity (3) with $v = e_i$, $w = e_j$.

All these problems can be formulated as a problem of estimates in the Gauß-type quadrature which can be solved in an elegant way using the connection to the underlying Lanczos process.

This paper is organized as follows. In Section 2 and 3 we summarize very briefly some classical results about the Gauß and Gauß-Radau quadratures and their close relations to the Lanczos algorithm and Jacobi matrices. In Section 4 we recall the known solutions to Problems (I) and (III) assuming exact arithmetic and present some simple alternative proofs. For the more detail exposition we refer to the report [GS-93] and to the papers [Ga-81], [Sz-39], [DR-84], [Pa-87], [HS-52], [S-55], [Go-73], [DEG-72], [DGN-78], [GM-91], [RW-92], [GM-93], [FG-92]. In Section 5, we analyze the effect of rounding errors. Based on the backward error results developed in [Gr-89], it is proved that the basic relation between the accuracy of the Gauß quadrature for $f(\lambda) = \lambda^{-1}$ and the rate of convergence of the corresponding conjugate gradient (CG) process holds true in finite precision arithmetic, even though the computed quantities may differ significantly from their exact precision counterparts. As a consequences, we show how this approach can be used for "reconstructing" the A-norm of the error in the finite precision CG run which is illustrated by numerical experiments. We note that the particular case of (1)

$$u^T(\lambda I - A)^{-1}u, \tag{13}$$

where λ is outside the A's spectrum, is of great importance in physical chemistry and solid state physics computation. The Lanczos method and method of continued fractions are succesfully used in this context for years, see e.g. [MF-86], [MP-89], [R-79], [HSa-84], [HSa-85], [AB-88], [H-87], [H-89]. The results presented in this paper offer a comprehensive insight into the riddle of this computation. Section 5 explains some well known practical observations concerning the effect of rounding errors.A summary is given in Section 6.

Throughout this paper matrices, vectors and functions are real.

2 Gauß and Gauß—Radau Quadratures

In this section we recall some basic results from the theory of Gauß and Gauß-Radau quadratures.

Let $\zeta \le \lambda_1, \lambda_N \le \xi$, $\sigma_j \neq 0$, $j = 1, 2, \ldots, N$. We consider a distribution function $\omega(\lambda)$ with the N points of increase $\lambda_1, \lambda_2, \ldots, \lambda_N$, defined by

$$
\begin{aligned}
\omega(\lambda) &= 0 \qquad\qquad \text{for } \zeta \le \lambda < \lambda_1 \\
\omega(\lambda) &= \sum_{j=1}^{l} \sigma_j^2 \quad \text{for } \lambda_l \le \lambda < \lambda_{l+1},\ l = 1, 2, \ldots, N-1, \qquad (14) \\
\omega(\lambda) &= \sum_{j=1}^{N} \sigma_j^2 \quad \text{for } \lambda_N \le \lambda \le \xi.
\end{aligned}
$$

Suppose that $f(x)$ is a function which is continuous at each λ_j, $j = 1, 2, \ldots, N$. Then the Rieman-Stieltjes integral of the function $f(\lambda)$ over the interval $\langle \zeta, \xi \rangle$ with the distribution function $\omega(\lambda)$ reduces to the finite sum

$$
\int_{\zeta}^{\xi} f(\lambda) d\omega(\lambda) = \sum_{j=1}^{N} \sigma_j^2 f(\lambda_j). \qquad (15)
$$

Nevertheless, we will keep the "integral" notation in some places for its convenience.

Let $q_0(\lambda), q_1(\lambda), \ldots, q_N(\lambda)$, where $q_i(\lambda)$ is of exact degree i, $i = 0, 1, \ldots, N$, be a sequence of monic orthogonal polynomials with respect to the inner product induced by (15), i.e.

$$
\int_{\zeta}^{\xi} q_i(\lambda) q_j(\lambda) d\omega(\lambda) = 0\ , \quad i \neq j,\ 0 \le i,\ j \le N.
$$

$$
q_N(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_N).
$$

Let

$$
\begin{aligned}
&\kappa_i = \Big(\int_{\zeta}^{\xi} q_i^2(\lambda) d\omega(\lambda) \Big)^{-1/2}, \quad i = 0, 1, \ldots, N-1, \\
&\kappa_N = \kappa_{N-1}.
\end{aligned}
$$

Then $p_0(\lambda), p_1(\lambda), \ldots, p_N(\lambda)$ defined by

$$p_i(\lambda) = \kappa_i q_i(\lambda), \quad i = 0, 1, \ldots, N,$$

is the corresponding sequence of orthonormal polynomials.

Theorem 2.1 (Gauß quadrature) . *Let $1 \leq n < N$. There exist n points μ_j, $j = 1, 2, \ldots, n$, $\zeta \leq \mu_1 < \mu_2 < \ldots < \mu_n \leq \xi$, and n positive weight coefficients $\omega_1, \ldots, \omega_n$ such that for any $f \in C^{2n}[\zeta, \xi]$*

$$\begin{aligned} (16) \quad \int_\zeta^\xi f(\lambda) d\omega(\lambda) &= \sum_{j=1}^n \omega_j f(\mu_j) + R_n(f), \\ R_n(f) &= \int_\zeta^\xi f(\lambda, \mu_1, \mu_1, \mu_2, \mu_2 \ldots, \mu_n, \mu_n) q_n^2(\lambda) d\omega(\lambda) = \\ &= \frac{f^{(2n)}(\nu)}{(2n)!} \int_\zeta^\xi q_n^2(\lambda) d\omega(\lambda), \quad \nu \in (\zeta, \xi), \end{aligned}$$

$f(\lambda, \mu_1, \mu_1, \mu_2, \mu_2 \ldots, \mu_n, \mu_n)$ is the $2n$-th divided difference of a function f with respect to the abscissas $\lambda, \mu_1, \mu_1 \ldots, \mu_n, \mu_n$ (see e.g. [DR-84], Section 4.5, pp. 297-298, or [Ga-81], p. 83).

The abscissas are equal to the roots of the polynomial $q_n(\lambda)$. The weights ω_j can be determined using the following particular choice of f:

$$\int_\zeta^\xi q_{n-1}(\lambda) \frac{q_n(\lambda)}{\lambda - \mu_j} d\omega(\lambda) = \omega_j q_{n-1}(\mu_j) q_n'(\mu_j),$$

and using the orthogonality

$$\int_\zeta^\xi q_{n-1}(\lambda) \frac{q_n(\lambda)}{\lambda - \mu_j} d\omega(\lambda) = \int_\zeta^\xi q_{n-1}^2(\lambda) d\omega(\lambda) = \kappa_{n-1}^{-2}.$$

Consequently,

$$(17) \qquad \omega_j = \frac{\kappa_{n-1}^{-2}}{q_{n-1}(\mu_j) q_n'(\mu_j)}, \quad j = 1, \ldots, n.$$

For the theory of the Gauß quadrature and related topics we refer to the excellent classical monography [Sz-39], chapters 1-3 and 15, and to the

deep survey paper [Ga-81] from the book devoted to the E.B. Christoffel (it is worth to note that this book contains many other remarkable contributions). They include many historical remarks and references to original papers. A comprehensive survey of the Gauß quadrature may be found in [DR-84], including the error analysis. [DR-84] deals with the weighted Riemann integral. Because the basic statements concerning integration rules of the Gauß type (chapter 2.7) rely on the theory of interpolation and the first mean value theorem (cf. [DR-84, p. 15 and Section 4.5]), they can be easily reformulated for the more general case of the Riemann – Stieltjes integral.

Orthonormal polynomials $p_0(\lambda), p_1(\lambda), \ldots, p_N(\lambda)$ and the corresponding monic orthogonal polynomials $q_0(\lambda), q_1(\lambda), \ldots, q_N(\lambda)$ satisfy three term recurrences

$$
\begin{aligned}
& p_0(\lambda) = 1, \quad q_0(\lambda) = 1 \\
& \beta_{i+1} p_i(\lambda) = \lambda p_{i-1}(\lambda) - \alpha_i p_{i-1}(\lambda) - \beta_i p_{i-2}(\lambda), \quad i = 1, \ldots, N-1, \\
& p_N(\lambda) = \lambda p_{N-1}(\lambda) - \alpha_N p_{N-1}(\lambda) - \beta_N p_{N-2}(\lambda), \\
& q_i(\lambda) = \lambda q_{i-1}(\lambda) - \alpha_i q_{i-1}(\lambda) - \beta_i^2 q_{i-2}(\lambda), \quad i = 1, \ldots, N,
\end{aligned} \tag{18}
$$

where

$p_{-1}(\lambda) = q_{-1}(\lambda) = 0,$
$\alpha_i = \int_\zeta^\xi \lambda p_{i-1}^2(\lambda) d\omega(\lambda), \quad i = 1, \ldots, N,$
$\beta_1 = 0,$
$\beta_{i+1} = (\int_\zeta^\xi (\lambda p_{i-1}(\lambda) - \alpha_i p_{i-1}(\lambda) - \beta_i p_{i-2}(\lambda))^2 d\omega(\lambda))^{1/2}, \ i = 1, 2, \ldots, N-1.$
Then

$$
\begin{aligned}
& \kappa_0 = 1 \\
& \kappa_i = (\prod_{j=1}^{i} \beta_{j+1})^{-1} \quad i = 1, 2, \ldots, N-1, \\
& \kappa_N = \kappa_{N-1}.
\end{aligned} \tag{19}
$$

Using quadrature formulas, we wish to determine computable upper and lower bounds for the integral (15). We consider an integration formula of Gauß type with one preassigned abscissa $\mu_0 \in [\zeta, \xi]$,

$$\int_\zeta^\xi f(\lambda)d\omega(\lambda) = \omega_0 f(\mu_0) + \sum_{j=1}^{n} \omega_j f(\mu_j) + R_n(f) \tag{20}$$

where the n abscissas $\mu_1, \mu_2, \ldots, \mu_n$ and $n+1$ weights $\omega_0, \omega_1, \ldots, \omega_n$ are to be determined so that the rule is exact for polynomials of the highest possible degree (that is $2n$). As described below, we are interested in the special cases $\mu_0 = \zeta$ or $\mu_0 = \xi$.

Theorem 2.2 (Gauß-Radau quadrature) . *Let $1 \le n \le N-1$. Let $\mu_0 \in [\zeta, \xi]$ be a fixed preassigned abscissa. Then there exists a quadratic rule so that for any $f \in C^{2n+1}[\zeta, \xi]$*

$$\int_\zeta^\xi f(\lambda)d\omega(\lambda) = \sum_{j=0}^{n} \omega_j f(\mu_j) + R_n(f) \tag{21}$$

$$R_n(f) = \int_\zeta^\xi f(\lambda, \mu_0, \mu_1,\ \mu_1, \mu_2, \mu_2, \ldots, \mu_n, \mu_n)(\lambda - \mu_0) \prod_{i=1}^{n} (\lambda - \mu_i)^2 d\omega(\lambda),$$

where $f(\lambda, \mu_0, \mu_1, \mu_1, \ldots, \mu_n, \mu_n)$ is $(2n+1)$th divided difference with respect to the abscissas $\lambda, \mu_0, \mu_1, \mu_1, \ldots, \mu_n, \mu_n$. If $\mu_0 = \zeta$ resp. $\mu_0 = \xi$, then

$$R_n(f) = \frac{f^{2n+1}(\nu)}{(2n+1)!} \int_\zeta^\xi (\lambda - \mu_0) \prod_{i=1}^{n} (\lambda - \mu_i)^2 d\omega(\lambda), \qquad \nu \in (\zeta, \xi). \tag{22}$$

It is known that the abscissas $\mu_1, \ldots, \mu_n$ are equal to the roots of the polynomial $\tilde{q}_{n+1}(\lambda)$ which is given by the relation

$$\tilde{q}_{n+1}(\lambda) = (\lambda - \tilde{\alpha}_{n+1}) q_n(\lambda) - \beta_{n+1}^2 q_{n-1}(\lambda), \tag{23}$$

where the unknown coefficient $\tilde{\alpha}_{n+1}$ is determined using the condition $\tilde{q}_{n+1}(\mu_0) = 0$,

$$\tilde{\alpha}_{n+1} = \mu_0 - \beta_{n+1}^2 \left[\frac{q_n(\mu_0)}{q_{n-1}(\mu_0)} \right]^{-1}, \tag{24}$$

and for $\tau_n = q_n(\mu_0)/q_{n-1}(\mu_0)$, we have the recurrence

$$\tau_1 = \mu_0 - \alpha_1, \qquad \tau_j = \mu_0 - \alpha_j - \beta_j^2/\tau_{j-1}, \qquad j = 2, \ldots, n.$$

The weights $\omega_0, \omega_1, \ldots, \omega_n$ are given by

$$\omega_j = \frac{\prod_{k=2}^{n+1} \beta_k^2}{q_n(\mu_j)\tilde{q}'_{n+1}(\mu_j)}, \quad j = 0, 1, \ldots, n. \tag{25}$$

For proofs see, e.g. [GS-93].

As it is shown in next sections, we are interested particularly in computing upper and lower bounds for the integrals

$$\int_\zeta^\xi \lambda^{-1} d\omega(\lambda) \quad \text{and} \quad \int_\zeta^\xi \lambda^{-2} d\omega(\lambda). \tag{26}$$

Let $\zeta > 0$. If $f(\lambda) = \lambda^{-1}$, resp $f(\lambda) = \lambda^{-2}$, then from (22) it is clear that for $\mu_0 = \zeta$ both $R_n(\lambda^{-1})$ and $R_n(\lambda^{-2})$ are negative and for $\mu_0 = \xi$ both $R_n(\lambda^{-1})$ and $R_n(\lambda^{-2})$ are positive. Consequently, the Gauß-Radau rule (21) gives for $\mu_0 = \zeta$ the desired upper bounds while for $\mu_0 = \xi$ the desired lower bounds. Moreover, from (16) it follows that for $f(\lambda) = \lambda^{-1}$, resp. $f(\lambda) = \lambda^{-2}$, the truncation error $R_n(f)$ of the Gauß quadrature is always positive. Consequently, the Gauß rule gives for both these functions always lower bound for (15).

In the next section, we recall an elegant way of computing weights and abscissas of Gauß and Gauß-Radau quadratures based on reformulating the problem in the context of the Lanczos algorithm.

3 Connection to Jacobi matrices

It is well known that the Gauß and Gauß-Radau quadratures imply the fundamental orthonormality property, i.e., the first m polynomials $p_0(\lambda), \ldots, p_{m-1}(\lambda)$ orthonormal with respect to the innerproduct induced by (15) are also orthonormal with respect to the discrete m-point quadratic form

$$(f, g)_m = \sum_{j=j_0}^{n} \omega_j f(\mu_j) g(\mu_j) \tag{27}$$

where ω_j, μ_j are the weights and abscissas of the quadrature formula, $j_0 = 1$ and $m = n$ for the Gauß , $j_0 = 0$ and $m = n + 1$ for the Gauß-Radau rule. Using the duality between orthonormal polynomials defined by (27) and the Euclidean geometry (which was demonstrated in a similar context in the

classical paper [HS-52]) we can determine the quadrature parameters simply as the characteristics of the following Jacobi matrix $\tilde{T}_m$,

$$\tilde{T}_m = \begin{pmatrix} \alpha_1\beta_2 & & & \\ \beta_2\alpha_2 & & & \\ & \ddots & \ddots & \ddots \\ & & & \alpha_{m-1}\beta_m \\ & & & \beta_m \quad \tilde{\alpha}_m \end{pmatrix}, \tag{28}$$

where for the Gauß rule $\tilde{\alpha}_m$ is identical to α_n given by (18), while for the Gauß-Radau rule $\tilde{\alpha}_m = \tilde{\alpha}_{n+1}$ which is given by (24).

Let S be the orthonormal matrix with the normalized eigenvectors of $\tilde{T}_m$ denoted by s_j, $j = j_0, \ldots, n$, as its columns. Then

$$\omega_j = (e_1^T s_j)^2 = s_{1j}^2, \quad j = j_0, \ldots, n, \tag{29}$$

i.e., the weights in the Gauß (resp. Gauß–Radau) rule are the square of the first elements of the normalized eigenvectors of the Jacobi matrix $\tilde{T}_m$ given by (28). The abscissas are identical to the eigenvalues of $\tilde{T}_m$.

This result can be found in [W-62], but it was known earlier, cf. [Ga-81]. In [GW-69] the three term recurrency was related to the matrix of moments and the QR algorithm was proposed for computing the eigenvalues and eigenvectors of $\tilde{T}_m$. In [Go-73] the problem of computing $\tilde{\alpha}_{n+1}$ in the Gauß–Radau quadrature was considered as the inverse eigenvalue problem for the matrix $\tilde{T}_m$. Clearly, all the entries of $\tilde{T}_m$ except $\tilde{\alpha}_m = \tilde{\alpha}_{n+1}$ are given, and $\tilde{\alpha}_{n+1}$ is to be determined from the condition that μ_0 is a given $\tilde{T}_m$'s eigenvalue. Then

$$\tilde{\alpha}_{n+1} = \mu_0 + \eta_n, \tag{30}$$

where η_n is the last entry of the vector $z = (\eta_1, \eta_2, \ldots, \eta_n)^T$ which is the solution to the linear system

$$(T_n - \mu_0 I)z = \beta_{n+1}^2 e_n, \tag{31}$$

where T_n is the n by n leading principal submatrix of $\tilde{T}$. For the other related results we refer to [Ga-68], [DEG-72], [Go-74], [DGN-78], [Ga-81], [KG-83], [GK-89], [K-89]. A comprehensive discussion of the problem of numerically

generating the recursion coefficients of orthonormal polynomials can be found in [Ga-82].

For computing the Gauß or Gauß–Radau approximations to the integrals (26) one does not need to determine the eigenvalues and the first entries of the $\tilde{T}_m$'s eigenvectors. Indeed, for $f(\lambda) = \lambda^{-i}$, $i = 1, 2$,

$$\int_\zeta^\xi \lambda^{-i} d\omega(\lambda) = \sum_{j=j_0}^{n} \omega_j \mu_j^{-i} + R_n(\lambda^{-i})$$

where

$$\sum_{j=j_0}^{n} \omega_j \mu_j^{-i} = (S^T e_1)^T diag(\mu_j^{-i})(S^T e_1) = e_1^T \tilde{T}_m^{-i} e_1 = (\tilde{T}_m^{-i})_{11},$$

$$\int_\zeta^\xi \lambda^{-i} d\omega(\lambda) = (\tilde{T}_m^{-i})_{11} + R_n(\lambda^{-i}). \tag{32}$$

Solving the m by m system

$$\tilde{T}_m g = e_1, \tag{33}$$

one can simply determine

$$(\tilde{T}_m^{-1})_{11} = e_1^T g, \tag{34}$$

$$(\tilde{T}_m^{-2})_{11} = \| g \|^2 . \tag{35}$$

The next section shows an interesting applications of the given formulas. Using the relation of the Lanczos algorithm to continued fractions, the Gauß quadrature approximation to $\int_\zeta^\xi \lambda^{-1} d\omega(\lambda)$ will be computed by a simple three term recurrence.

4 Estimates in quadratic formulas — exact arithmetic

In this section we describe the solution of Problems I and III, assuming exact arithmetic, and in the Section 5, the effect of rounding errors will be analyzed.

As it is shown in Section 1, Problem I reduces to finding estimates, upper and lower bounds on

$$(r^0)^T A^{-i} r^0$$

where r^0 is the initial residual, $r^0 = b - Ax^0$, $i = 1, 2$. To simplify the notation we assume, as stated in Section 1, that the eigenvalues of A are distinct, and that r^0 has a nonvanishing projection in the direction of every eigenvector u_j of A, i.e.

$$\sigma_j = u_j^T r^0 / \| r^0 \| \neq 0, \quad j = 1, 2, \ldots, N. \tag{36}$$

The extension of the results presented below to the general case is trivial.

Considering the eigendecomposition $A = U \Lambda U^T$, (4) and (5) become

$$\| x - x^0 \|^2 = \| r^0 \|^2 \sum_{j=1}^{N} \frac{\sigma_j^2}{\lambda_j^2} \tag{37}$$

$$\| x - x^0 \|_A^2 = \| r^0 \|^2 \sum_{j=1}^{N} \frac{\sigma_j^2}{\lambda_j}. \tag{38}$$

Using the distribution function $\omega(\lambda)$ defined by (14) for some ζ, ξ, $\zeta \leq \lambda_1$, $\lambda_N \leq \xi$, (37) and (38) can be written as

$$\| x - x^0 \|^2 = \| r^0 \|^2 \int_{\zeta}^{\xi} \lambda^{-2} d\omega(\lambda) \tag{39}$$

$$\| x - x^0 \|_A^2 = \| r^0 \|^2 \int_{\zeta}^{\xi} \lambda^{-1} d\omega(\lambda). \tag{40}$$

The desired estimates, upper and lower bounds can be found by computing the Gauß and Gauß–Radau quadratures (32) which can be easily done by using (33)-(35). The coefficients α_j, β_{j+1}, $j = 1, 2, \ldots, n$, defined by (18) are identical with those given by n steps of the Lanczos process for the original matrix A with the initial vector $h^1 = r^0 / \| r^0 \|$,

$$\begin{aligned} h^0 &= 0, \beta_1 = 0 \\ &\alpha_k = (Ah^k - \beta_k h^{k-1}, h^k) \\ \tilde{h}^{k+1} &= Ah^k - \alpha_k h^k - \beta_k h^{k-1} \end{aligned} \tag{41}$$

$$\begin{aligned} \beta_{k+1} &= \| \tilde{h}^{k+1} \| \\ h^{k+1} &= \tilde{h}^{k+1}/\beta_{k+1}, \quad k = 1, 2, \ldots, n, \end{aligned}$$

$n \leq N-1$. For $n = N$ $\beta_{N+1} = 0$, and the Gauß quadrature formula is identical with the original integral (15). Thus, for computing estimates, upper and lower bounds on (37) and (38) we run n steps of the Lanczos process (41), then set $m = n$ and determine the Gauß estimate using (33)-(35), or set $m = n+1$, $m \leq N-1$, compute $\tilde{\alpha}_{n+1}$ for a properly choosen $\mu_0 \leq \lambda_1$ (resp. $\mu_0 \geq \lambda_N$) from (30)-(31) and determine the Gauß-Radau upper and lower bounds from (33)-(35), cf. [DGN-78]. Note, that for approximating μ_0 we can use the information about the location of the extremal eigenvalues λ_1 and λ_N obtained from the Lanczos process (41).

The described approach leads to two questions which we give below.

First, let T_N be the tridiagonal matrix given in N steps of (41),

$$T_N = \begin{pmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & & & \\ & \ddots & \ddots & \ddots & \\ & & & & \beta_N \\ & & & \beta_N & \alpha_N \end{pmatrix}. \tag{42}$$

Then

$$\| x - x^0 \|_A^2 = (r^0)^T A^{-1} r^0 = \| r^0 \|^2 \, (e_1)^T T_N^{-1} e_1 = \| r^0 \|^2 \, (T_N^{-1})_{11}$$

and (38) gives

$$(T_N^{-1})_{11} = \sum_{j=1}^{N} \frac{\sigma_j^2}{\lambda_j}. \tag{43}$$

It is well known (cf. [Sz-39, HS-52]), that (43) can be expanded into a continued fraction:

$$\sum_{j=1}^{N} \frac{\sigma_j^2}{\lambda_j} = C_N = \frac{G_N}{E_N} = \cfrac{1}{\alpha_1 - \cfrac{\beta_2^2}{\alpha_2 - \cfrac{\beta_3^2}{\vdots \cfrac{}{\alpha_{N-1} - \cfrac{\beta_N^2}{\alpha_N}}}}}, \tag{44}$$

where the numerator E_N and denominator G_N are given by the three term reccurences

$$\begin{aligned} E_0 &= 1, E_1 = \alpha_1, E_j = -\alpha_j E_{j-1} - \beta_j^2 E_{j-2} \\ G_0 &= 0, G_1 = 1, G_j = -\alpha_j G_{j-1} - \beta_j^2 G_{j-2}, \quad j = 2, \ldots, N. \end{aligned} \tag{45}$$

The Gauß estimate for $\int_\zeta^\xi \lambda^{-1} d\omega(\lambda)$ is given after n steps of (41), $n \leq N-1$, by $(T_n^{-1})_{11}$, where T_n is the n-th leading principal submatrix of T_N. T_n can be considered as a result of the n-dimensional Lanczos proces applied to the matrix T_n with the initial vector $v^1 = e_1$. Therefore $(T_n^{-1})_{11}$ can be expanded into the continued fraction

$$(T_n^{-1})_{11} = \sum_{l=1}^{n} \frac{\omega_l}{\mu_l} = C_n, \tag{46}$$

where C_n is the n-th convergent of C_N. It can be determined simply by

$$C_n = \frac{G_n}{E_n}, \tag{47}$$

G_n, E_n are given by (45).

Summarizing, the Gauß estimate $(T_n^{-1})_{11}$ for (38) can be computed using the three term recurrences for continued fractions (45). This gives rise to the question:
Can we describe in a simple way the rate of convergence of C_n to C_N, or, equivalently, $(T_n^{-1})_{11}$ to $(T_N^{-1})_{11}$?

Note that we are not interested in the convergence of individual pairs μ_l, ω_l to some λ_j, σ_j^2 here. For the detailed (but still incomplete) discussion of the last problem in both the exact and finite precision arithmetic we refer

to [SG-92]. In a recent paper [FF-93] it is shown that, roughly speaking, a very good approximation (in some sense) to the eigenvalue and weight distribution $\{\lambda_j, \sigma_j^2\}_{j=1,\ldots,N}$ can be obtained after only few iterations of the (exact arithmetic) Lanczos algorithm. The approach by Freund and Fischer [FF-93] is also based on the connection of the Lanczos algorithm to the Gauß quadrature. In [FH-93] a general complex non-Hermitian matrices are considered; the Arnoldi process and the nonsymmetric Lanczos algorithm are associated with Gauß quadratures in the complex plane.

It should also be emphasized that the association of the Gauß quadrature with the continued fractions described above is known from the original work by Gauß at the beginning of the 19th century. His approach was based on the theory of continued fractions associated with hypergeometric series, see the survey paper [Ga-81].

Second, instead of using the Lanczos process (41), we can run the corresponding conjugate gradient (CG) process for A, r^0:

$$
\begin{aligned}
d^0 &= r^0, \\
&\quad \varphi_{k-1} = (r^{k-1}, r^{k-1})/(d^{k-1}, Ad^{k-1}) \\
x^k &= x^{k-1} + \varphi_{k-1} d^{k-1} \\
r^k &= r^{k-1} - \varphi_{k-1} A d^{k-1} \\
&\quad \psi_k = (r^k, r^k)/(r^{k-1}, r^{k-1}) \\
d^k &= r^k + \psi_k d^{k-1}, \qquad\qquad k = 1, 2, \ldots, n,
\end{aligned} \tag{48}
$$

and then compute Lanczos coefficients α_k, β_{k+1} from the well-known formulas

$$
\begin{aligned}
\alpha_k &= \frac{1}{\varphi_{k-1}} + \frac{\psi_{k-1}}{\varphi_{k-2}}, \\
\beta_{k+1} &= \frac{\psi_k^{1/2}}{\varphi_{k-1}}, \qquad k = 1, 2, \ldots, n,
\end{aligned} \tag{49}
$$

where $\psi_0 = 0, \varphi_{-1} = 1$. In this way, we can compute not only the estimates and bounds for the initial errors (37), (38), but we obtain also the n-th conjugate gradient approximation x^n, minimizing the A-norm of the error among all the polynomial acceleration (or Krylov subspace) methods (cf.,e.g., [HY-81], Chapter 7). This gives rise to the question:

Can we give a simple formula (based on the computed quantities) for the A-norm of the error $\| x - x^n \|_A$ *at the n-th step of the conjugate gradient process?*

These two questions are in fact identical. They were originally solved (using the theory of moments) in [DGN-78], by proving the identity

$$\| x - x^n \|_A^2 = \| r^0 \|^2 \, [(T_N^{-1})_{11} - (T_n^{-1})_{11}]. \tag{50}$$

The original proof is quite difficult, as is that one based on the Cayley-Hamilton theorem presented in [GS-93]. Formula (50) is, however, a straight consequence of Theorem 2.1 (Gauß quadrature) for the distribution function $\omega(\lambda)$ defined by (14) and $f(\lambda) = 1/\lambda$. Indeed,

$$\int_\zeta^\xi (1/\lambda) d\omega(\lambda) \equiv \sum_{i=1}^N \frac{\sigma_i^2}{\lambda_i} = \sum_{l=1}^n \frac{\omega_l}{\mu_l} + R_n(1/\lambda), \tag{51}$$

i.e.,

$$(T_N^{-1})_{11} = (T_n^{-1})_{11} + R_n(1/\lambda). \tag{52}$$

From (16),

$$R_n(1/\lambda) = \int_\zeta^\xi \frac{q_n^2(\lambda)}{\lambda(\prod_{l=1}^n \mu_l)^2} d\omega(\lambda) = \sum_{i=1}^N \frac{\sigma_i^2 q_n^2(\lambda_i)}{\lambda_i(\prod_{l=1}^n \mu_l)^2},$$

because for $f(\lambda) = \lambda^{-1}$

$$f(\lambda, \mu_1, \mu_1, \ldots, \mu_n, \mu_n) = [\lambda(\prod_{l=1}^n \mu_l)^2]^{-1},$$

which is easy to prove by induction. Considering

$$\frac{q_n^2(\lambda)}{(\prod_{l=1}^n \mu_l)^2} = \frac{\| r^n \|^2 \, p_n^2(\lambda)}{\| r^0 \|^2},$$

the error in the Gauß quadrature is written in the form

$$R_n(1/\lambda) \;=\; \frac{1}{\| r^0 \|^2} \sum_{i=1}^N \frac{\sigma_i^2 \, \| r^n \|^2 \, p_n^2(\lambda_i)}{\lambda_i} =$$

$$= \frac{1}{\| r^0 \|^2} \{ \| r^n \| p_n(A)h^1 \}^T A^{-1} \{ \| r^n \| p_n(A)h^1 \} =$$
$$= \frac{1}{\| r^0 \|^2} \| x - x^n \|_A^2 .$$

For details see [GS-93].

An alternative proof, which is very elegant, was proposed by the referee (this proof is based on the similar idea as the original one in [DGN-78], but uses different polynomial expansion). Consider the polynomial depending on n-parameters $\nu = (\nu_0, \ldots, \nu_{n-1})^T$

$$p(\lambda, \nu) = 1 - \lambda \sum_{j=0}^{n-1} \nu_j p_j(\lambda).$$

The conjugate gradient method minimizes the A-norm of the error, which may be viewed as a minimization problem in ν

$$\|x - x^n\|_A = \|r^n\|_{A^{-1}} = \min_{\nu}((r^0)^T p(A, \nu) A^{-1} p(A, \nu) r^0). \tag{53}$$

After a simple manipulation,

$$(r^0)^T p(A, \nu) A^{-1} p(A, \nu) r^0 = (r^0)^T A^{-1} r^0 - 2\|r^0\|^2 \nu_0 + \|r^0\|^2 \nu^T T_n \nu. \tag{54}$$

It is clear that the minimal solution $\tilde{\nu}$ is given by

$$T_n \tilde{\nu} = \|r^0\| e_1, \tag{55}$$

which, substituting to (54) and (53), gives (50).

Thus, the convergence of $(T_n^{-1})_{11}$ to $(T_N^{-1})_{11}$ is determined by the convergence of the A-norm of the error in the corresponding conjugate gradient process, and vice versa. At the step $n < N$ we do not know, of course, the value $(T_N^{-1})_{11}$, but we can easily compute its Gauß-Radau bounds. In this way, (50) can be used for computing bounds for $\| x - x^n \|_A^2$. A different approach to the energy error approximation is used in [D-93].

In practice, however, rounding errors may crucially affect the computation. Are the accuracy of the Gauß quadrature for $f(\lambda) = \lambda^{-1}$ and the convergence of the corresponding conjugate gradient process related in the way similar to (50) even in the finite precision case? The effect of rounding

errors is analyzed in the next section, and the positive answer to this question is given.

The solution of Problem III is in fact already given by the above considerations. The approach based on the Gauß-Radau bounds may also serve as a powerful tool for finding (in an easy way) analytical bounds on the entries of the matrix inverse expressed explicitly in terms of the entries of the original matrix. We demonstrate this on the next example.

We want to determine upper and lower bounds for $(A^{-1})_{jj}$. One step of the Lanzos process (41) with $h^1 = e_j$ gives $\alpha_1 = e_j^T A e_j = a_{jj}$, $\beta_2^2 = \sum_{l \neq j} a_{lj}^2 \neq 0$(we avoid the trivial case $\sum\limits_{l \neq j} a_{lj}^2 = 0$). From (30)-(31)

$$\tilde{\alpha}_2 = \mu_0 + \frac{\beta_2^2}{a_{jj} - \mu_0}$$

and (33) becomes the 2 by 2 system

$$\begin{pmatrix} a_{jj} & (\sum\limits_{l \neq j} a_{lj}^2)^{1/2} \\ (\sum\limits_{l \neq j} a_{lj}^2)^{1/2} & \tilde{\alpha}_2 \end{pmatrix} g = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

from which after some manipulations

$$(e_1^T g) = \frac{1}{\mu_0} - \frac{(\mu_0 - a_{jj})^2}{\mu_0((\sum\limits_{l=1} a_{lj})^2 - \mu_0 a_{jj})}.$$

Assuming $0 < \mu_0^L \leq \lambda_1$ and $\mu_0^U \geq \lambda_N$, where λ_1 is the smallest and λ_N the largest eigenvalue of A, we receive the bound

$$(56) \frac{1}{\mu_0^U} - \frac{(\mu_0^U - a_{jj})^2}{\mu_0^U((\sum\limits_{l=1}^{n} a_{lj})^2 - \mu_0^U a_{jj})} < (A^{-1})_{jj} < \frac{1}{\mu_0^L} - \frac{(\mu_0^L - a_{jj})^2}{\mu_0^L((\sum\limits_{l=1}^{n} a_{lj})^2 - \mu_0^L a_{jj})},$$

which was originally published in [RW-92].

For the off-diagonal entries the identity

$$(57) \qquad A_{ij}^{-1} = e_i^T A^{-1} e_j = 1/2[(A_{ii}^{-1} + A_{jj}^{-1} - (e_i - e_j)^T A^{-1}(e_i - e_j)]$$

can be used to develop bounds similar to (56) in a quite analogous way. As expected, it will result in rather complicated formulas.

5 Rounding error analysis

The effect of the rounding errors in the Lanczos process may crucially depend on the distribution of the eigenvalues of A and on the components of the initial vector in the directions of the eigenvectors of A (cf. [St-91], [GS-92], [SG-92], [N-93]). As a consequence, computed Lanczos coefficients (elements of the tridiagonal matrix) may differ substantially (even several orders of magnitude) from their theoretical counterparts.

It was observed in atomic physics and physical chemistry calculations that, despite this numerical difficulty, the method of continued fractions [HSa-84], [HSa-85], and other methods used for computing (13), see, e.g., [MP-89], frequently give very precise results. Similar behavior was observed in Lanczos process based Gauß quadrature calculations.

In the previous sections, we have shown that both of these observations are of the same origin because the problem of computing (13) can be formulated as the problem of computing the Gauß quadrature for $f(\lambda) = \lambda^{-1}$. In exact arithmetic, the accuracy of the Gauß quadrature for $f(\lambda) = \lambda^{-1}$ and for the distribution function $\omega(\lambda)$ with the finite points of increase, defined by (14), is determined by the energy norm of the error in the corresponding conjugate gradient process, see (50). In this section, we derive an analogy of (50) for the computed values in finite precision arithmetic. For the different approach to the rounding error analysis of the Gauß quadrature cf. [K-94].

We preserve the notation of previous sections for the theoretical exact arithmetic values, while denoting the actually computed quantities by the superscript $\hat{}$. The abscissas μ_j and weights ω_j of the n-point Gauß quadrature, $j = 1, 2, \ldots, n$, are in Sections 3,4 determined as the eigenvalues and the square of the first elements of the normalized eigenvectors of the Jacobi matrix T_n computed as a result of the n steps of the Lanczos or conjugate gradient algorithm. In exact arithmetic, the Lanczos and CG algorithms ((41) and (48)-(49)) produce the same tridiagonal matrices T_n. In the presence of rounding errors, the results computed by the Lanczos run may slightly differ from the corresponding results of the CG run. Though there are many theoretical arguments supporting the conjecture that these differences are small (see [Si-84a], [Si-84b], [Gr-89]), and it is in a good agreement with many experiments (see, e.g. [St-91]), the rigorous theoretical quantification is still missing. The detailed discussion of this question is out of the range of this paper. We are interested in the relation between the error in the finite

precision Gauß quadrature and the rate of convergence of the CG process. Therefore, henceforth we assume that all the actually computed quantities are determined by the finite precision CG run, i.e., using the formulas (48), (49).

As $\hat{T}_n$ differs from its theoretical counterpart T_n, the values $\hat{\mu}_j, \hat{\omega}_j$ may be very different from μ_j, ω_j, and the actual error of the computed Gauß quadrature can no longer be expressed as in (16).

We note here that we are primarily interested in the error in determining the abscissas and weights, caused by rounding errors in the CG process (48). We consider all the other rounding errors in computing the quadrature negligible (it concerns (49), the error in determining $\hat{\mu}_j, \hat{\omega}_j$ from $\hat{T}_n$, errors in the computation of function values and in forming sums; for more details we refer to [DR-84], Chapter 4) and we do not take them into account. This assumption will simplify our analysis remarkably and, unless the precision of the quadrature is close to the machine precision, it does not affect the correctness of the proved results.

Rounding errors in the Lanczos algorithm were thoroughly investigated by Paige in his Ph.D. Thesis [Pg-71] and in subsequent papers (see, e.g., [Pg-72], [Pg-76], [Pg-80]), and then by many other authors (for a partial survey, see, e.g., [Si-84a], [Si-84b], [GS-92]). Our analysis is based on the backward error results developed by Greenbaum [Gr-89]. According to that, for a given fixed n, the Lanczos matrix $\hat{T}_n$, produced in n steps of the finite precision CG process for the matrix A and the initial residual r^0, is **identical** to the Lanczos matrix $\overline{T}_n$, generated in n steps of the **exact** CG algorithm applied to a certain larger matrix $\overline{A}_n$, having possibly many more eigenvalues then A, but whose eigenvalues all lie within tiny intervals about the eigenvalues of A, with a certain initial residual $\overline{r}_n^0$. The matrix $\overline{A}_n$ (the exact distribution of $\overline{A}_n$'s eigenvalues) and the initial vector $\overline{r}_n^0$ of the exact CG process depend on the actual rounding errors in the steps 1 through n of the original finite precision run. The dimension of the "equivalent" exact CG proces is denoted by $\overline{N}$, $\overline{N} \geq N$. We emphasize, that $\overline{A}_n$, $\overline{r}_n^0$, and even the dimension $\overline{N}$, depend on n. For the Lanczos method the result can be formulated quite analogously.

We turn to the analysis of the effect of rounding errors in computing the Gauß quadrature. Let n be fixed, $n < N$. Based on the results by Greenbaum (cf. [Gr-89]), the problem of approximating the original sum

$$\sum_{i=1}^{N} \sigma_i^2 f(\lambda_i) \tag{58}$$

by the n-point Gauß quadrature approximation computed in the **finite precision arithmetic**,

$$\sum_{l=1}^{n} \hat{\omega}_l f(\hat{\mu}_l), \tag{59}$$

may now be considered as a problem of computing the **exact** n-point Gauß quadrature approximation (59) to the "original" sum

$$\sum_{j=1}^{\overline{N}} \overline{\sigma}_j^2 f(\overline{\lambda}_i), \tag{60}$$

defined by $\overline{A}_n$, $\overline{r}_n^0$.

If $f(\lambda)$ has a reasonably bounded first derivative in the neigborhoods of each λ_i, $\quad i = 1, 2, \ldots, N$, then (58) is close to (60). To show this, we need to quantify the relations between the eigenvalues and the squared first elements of the normalized eigenvectors of the matrices T_N and $\overline{T}_{\overline{N}}$. Resulting from the exact CG processes for A, r^0 and $\overline{A}_n, \overline{r}_n^0$, the Jacobi matrices T_N and $\overline{T}_{\overline{N}}$ have the same eigenvalues as A and $\overline{A}_n$, respectively. Using (29), the weights may be written as

$$\sigma_i^2 = s_{1i}^2, \quad \overline{\sigma}_j^2 = \overline{s}_{1j}^2, \quad i = 1, \ldots, N, \;\; j = 1, \ldots, \overline{N},$$

where s_i and $\overline{s}_j$ are the normalized eigenvectors of T_N and $\overline{T}_{\overline{N}}$, respectively.

Henceforth, $O(x)$ denotes the product of x with a constant independent of x. We recall two theorems, and then we prove Theorem 5.3, which relates (58) to (60). By $\|A\|$ we denote the spectral norm of A.

Theorem 5.1 (Gr-89) *For each eigenvalue $\overline{\lambda}_j$ of the Greenbaum matrix $\overline{A}_n$, constructed for the n steps of the finite precision CG process for A, r^0, there is some eigenvalue λ_i of A so that*

$$|\overline{\lambda}_j - \lambda_i| = \min_l |\overline{\lambda}_j - \lambda_l| \leq O(\overline{N}^3)\delta \parallel A \parallel, \tag{61}$$

where δ is determined in [Gr-89], pp. 22-24, 50-51.

Theorem 5.2 (St-91) *For each eigenvalue λ_i of A, there is some eigenvalue $\overline{\lambda}_k$ of $\overline{A}_n$ so that*

$$|\lambda_i - \overline{\lambda}_k| = \min_l |\lambda_i - \overline{\lambda}_l| \leq \frac{O(\overline{N})\delta \parallel A \parallel}{|(u_i, h^1)|}, \tag{62}$$

where δ is as in Theorem 5.1, $h^1 = r^0 / \parallel r^0 \parallel$.

For proofs see [G-89], p. 24, Theorem 1 (p. 23), and Theorem 1' (p. 51) and [St-91], Theorem 4.2 (p. 547). δ represents a quite complicated quantity which is hard to describe without giving details of the complicated and difficult proof of Theorem 5.1. Such exposition is hardly possible here, see [G-89].

Theorem 5.1 gives a bound on the size of intervals about the eigenvalues of A containing all the eigenvalues of $\overline{A}_n$. Numerical computations [Gr-89] suggest that this bound is a large overestimate, and that the eigenvalues of $\overline{A}_n$ are actually contained in much smaller intervals than the proven bounds would suggest. Based on the thorough discussion in [GS-92], we will assume that the size of these intervals is proportional to the machine precision ε. Though not formally justified yet, this assumption appears to be very realistic. Moreover, we believe that the formal proof is possible and we will return to this point elsewhere. Theorem 5.2 proves that for any eigenvalue of A, λ_i, there is some eigenvalue $\overline{\lambda}_k$ of $\overline{A}_n$ close to it (we suppose that $|(u_i, h^1)|$ is nonnegligible).

Let $\{J_1, J_2, \ldots, J_N\}$ be a partitioning of the set of indices $\{1, 2, \ldots, \overline{N}\}$ such that the eigenvalues $\{\overline{\lambda}_l,\ l \in J_i\}$ are within ϑ_i to λ_i, where ϑ_i is a moderate multiple of the machine precision ε, and J_i is nonempty, $i = 1, 2, \ldots, N$ (existence of such partitioning is in fact guaranteed, under the assumption mentioned above, by Theorems 5.1, 5.2). We assume that $\lambda_1 - \vartheta_1 > 0$, and the eigenvalues of $\overline{A}_n$ are ordered so that

$$0 < \overline{\lambda}_1 < \overline{\lambda}_2 < \ldots < \overline{\lambda}_{\overline{N}}. \tag{63}$$

Let $f(\overline{\lambda}_j) \in C^1(\bigcup_{i=1}^N [\lambda_i - \vartheta_i, \lambda_i + \vartheta_i])$ and $|f'(\lambda)| \leq \theta$, $\lambda \in \bigcup_{i=1}^N [\lambda_i - \vartheta_i, \lambda_i + \vartheta_i]$. Then for $j \in J_i$

$$f(\overline{\lambda}_j) = f(\lambda_i) + O(\theta)\vartheta_i, \tag{64}$$

and the sum (60) can be rewritten as

$$\sum_{j=1}^{\overline{N}} \overline{\sigma}_j^2 f(\overline{\lambda}_j) = \sum_{i=1}^{N} \sum_{j \in J_i} \overline{\sigma}_j^2 f(\overline{\lambda}_j) = \sum_{i=1}^{N} (f(\lambda_i) + O(\theta)\vartheta_i) \sum_{j \in J_i} \overline{\sigma}_j^2.$$

From [Gr-89], relation (8.21) on p. 60,

$$\sum_{j \in J_i} \overline{\sigma}_j^2 = \sum_{j \in J_i} \overline{s}_{1j}^2 = (u_i, h^1)^2 + O(\overline{N}^2)\delta \;=\; \sigma_i^2 + O(\overline{N}^2)\delta.$$

Finally, combining the above results, we have proved the following theorem:

Theorem 5.3 *The sums (58) and (60) are related by*

$$\sum_{j=1}^{\overline{N}} \overline{\sigma}_j^2 f(\overline{\lambda}_j) = \sum_{i=1}^{N} \sigma_i^2 f(\lambda_i) + P_n(f), \tag{65}$$

$$P_n(f) = O(\theta) \sum_{i=1}^{N} \vartheta_i \sigma_i^2 + O(\overline{N}^2)\delta \parallel A \parallel (\sum_{i=1}^{N} (f(\lambda_i) + O(\theta)\vartheta_i)). \tag{66}$$

As mentioned above, numerical computations justify that it is realistic to consider that ϑ_i, $i = 1, \ldots, N$, and δ in (66) are modest multiples of the machine precision ε. For any nonvanishing function f with a reasonable bounded first derivative, i.e. $\theta \ll 1/\varepsilon$, (58) is therefore close to (60).

We will characterize the **truncation error** (see [DR-84], (4.1.1)) of the **exact** n-point Gauß quadrature for (60) computed by the exact CG process for $\overline{A}_n$, $\overline{r}_n^0$.
Using (16),

$$\sum_{j=1}^{\overline{N}} \overline{\sigma}_j^2 f(\overline{\lambda}_j) = \sum_{i=1}^{n} \hat{\omega}_l f(\hat{\mu}_l) + \overline{R}_n(f), \tag{67}$$

$$\overline{R}_n(f) = \sum_{j=1}^{\overline{N}} \overline{\sigma}_j^2 \{ f(\overline{\lambda}_j, \hat{\mu}_1, \hat{\mu}_1, \ldots, \hat{\mu}_n, \hat{\mu}_n) \prod_{k=1}^{n} (\overline{\lambda}_j - \hat{\mu}_k)^2 \}. \tag{68}$$

As a consequence, we receive the following corollary to Theorem 5.3:

Corollary 5.1 *Consider the n-point Gauß quadrature for the sum (58). Then the total error of the Gauß quadrature approximation computed in finite precision arithmetic via the CG method is determined by the formula*

$$\sum_{i=1}^{N} \sigma_i^2 f(\lambda_i) = \sum_{l=1}^{n} \hat{\omega}_l f(\hat{\mu}_l) + \overline{R}_n(f) - P_n(f), \tag{69}$$

where $P(f)$ is given by (66) and $\overline{R}_n(f)$ is given by (68).

Corollary 5.1, in fact, formulates the basic result describing the effect of rounding errors to the Gauß quadrature computed via the CG process (for the Gauß-Radau quadrature the result is analogous). Formula (69) is analogous to (16) with one but very substantial difference. The main part of the error, i.e. $\overline{R}_n(f)$, is given by $f(\lambda, \hat{\mu}_1, \hat{\mu}_1, \ldots \hat{\mu}_n, \hat{\mu}_n)\hat{q}_n^2(\lambda)$ integrated not using the original "integral" (58), as in (16), but using the "integral" (60), determined by the actual values of rounding errors in steps 1 through n of the finite precision CG process for A, r^0.

Note that the **total error** (including the roundoff error) of the Gauß quadrature for (58) computed in finite precision arithmetic is expressed as the **truncation error** of the Gauß quadrature for a different problem (60). We emphasize that $R_n(f)$ and $\overline{R}_n(f)$ may differ substantially. That is as much as we are able to say for a general function f.

For the particular case $f = 1/\lambda$ the situation is more transparent. Indeed, using (51) and (52), (69) becomes

$$\sum_{i=1}^{N} \frac{\sigma_i^2}{\lambda_i} = \sum_{l=1}^{n} \frac{\hat{\omega}_l}{\hat{\mu}_l} + \overline{R}_n(1/\lambda) - P_n(1/\lambda) \tag{70}$$

or

$$(T_N^{-1})_{11} = (\hat{T}_n^{-1})_{11} + \overline{R}_n(1/\lambda) - P_n(1/\lambda), \tag{71}$$

where $P_n(1/\lambda)$ is given by (66), and

$$\overline{R}_n(1/\lambda) = (1/ \parallel \overline{r}_n^0 \parallel^2) \parallel \overline{x} - \overline{x}^n \parallel_{\overline{A}_n}^2 . \tag{72}$$

Because $\overline{A}_n$, $\overline{r}^0$ and $\overline{R}_n(1/\lambda)$ are not available during the actual finite precision computation, we want to replace $\overline{R}_n(1/\lambda)$ in (70), (71) by some quantity which is easily computable from the original data σ_i, λ_i, $i = 1, \ldots, N$, (or A, r^0). We use the bound

$$\overline{R}_n(1/\lambda) = (1/\parallel \overline{r}_n^0 \parallel^2) \parallel \overline{x} - \overline{x}^n \parallel^2_{\overline{A}_n} \leq (1/\parallel \overline{r}_n^0 \parallel^2) \parallel x - \overline{x}^{n-1} \parallel^2_{\overline{A}_n} \leq$$
$$\leq \frac{1}{\overline{\lambda}_1} \frac{\parallel \overline{r}^{n-1} \parallel^2}{\parallel \overline{r}^0 \parallel^2},$$

and the relation between the relative norm of the CG residual $\overline{\rho}_k = \parallel \overline{r}^k \parallel / \parallel \overline{r}_n^0 \parallel$ and the coefficients of the corresponding Lanczos process

$$(73)\quad \overline{\rho}_k = \frac{\hat{\beta}_{k+1}\overline{\rho}_{k-1}\overline{\rho}_{k-2}}{\hat{\alpha}_k\overline{\rho}_{k-2} - \hat{\beta}_k\overline{\rho}_{k-1}}, \qquad k = 2, \ldots, n-1, \quad \overline{\rho}_1 = \frac{\hat{\beta}_2}{\hat{\alpha}_1}, \quad \overline{\rho}_0 = 1,$$

[Gr-89], relation (4.10) on p.25 (note that all the quantities in (73) are considered as determined by the **exact** process for $\overline{A}_n$, $\overline{r}_n^0$). The coefficients $\hat{\alpha}_k, \hat{\beta}_k$ are computed by (49), from which

$$\overline{\rho}_k \equiv \hat{\rho}_k = \parallel \hat{r}^k \parallel / \parallel r^0 \parallel, \qquad k = 1, 2, \ldots, n-1.$$

As a consequence, we derive the following bound

$$(74)\qquad \overline{R}_n(1/\lambda) \leq \frac{\hat{\rho}^2_{n-1}}{\overline{\lambda}_1}.$$

Replacing $\overline{\lambda}_1$ by the smallest Ritz value $\hat{\mu}_1$, computed at the step n, we can write (74) in the final form suitable for computation

$$(75)\qquad \overline{R}_n(1/\lambda) = O(1)\frac{\hat{\rho}^2_{n-1}}{\hat{\mu}_1}.$$

Thus,

$$(76)\qquad (T_N^{-1})_{11} = (\hat{T}_n^{-1})_{11} + O(1)\frac{\hat{\rho}^2_{n-1}}{\hat{\mu}_1} - P_n(1/\lambda).$$

Using this formula (and considering $\mid P(1/\lambda) \mid$ small), one can check the accuracy of the computed quadrature (or the convergent of the continued fraction) during the computation. In our experiments the value $\hat{\rho}^2_{n-1}/\mu_1$ always gave an actual upper bound for the total error, unless it was of the order of the machine precision. Some examples of experimental results are presented below.

As mentioned in Section 4, we want to derive an analogy of (50) for actually computed quantities, i.e., to relate the accuracy of the computed Gauß quadrature for $f(\lambda) = \lambda^{-1}$ to the rate of convergence of the finite precision CG run. For that, we need to express $\overline{R}_n(1/\lambda)$ in terms of the energy norm of the error of the actual CG run for A, r^0. This can be easily done using the results proved by Greenbaum.

Let $\vartheta \leq O(\overline{N}^3)\delta \parallel A \parallel$ be the largest distance from an eigenvalue of $\overline{A}_n$ to the closest eigenvalue of A. From [Gr-89], proof of Theorem 3, pp. 28-29, and (8.21) on p. 60, and assuming that $\lambda_1 - \vartheta$ and $\lambda_N + \vartheta$ are closely approximated by the Ritz values $\hat{\mu}_1$ and $\hat{\mu}_n$ computed at the n-th step,

$$| \parallel r^0 \parallel^2 - \parallel \overline{r}^0 \parallel^2 | \; = \parallel r^0 \parallel^2 \frac{\hat{\mu}_n}{\hat{\mu}_1} O(\chi), \tag{77}$$

$$| \parallel x - \hat{x}^k \parallel_A^2 - \parallel \overline{x} - \overline{x}^k \parallel_{\overline{A}}^2 | = \parallel \hat{r}^k \parallel^2 \frac{\hat{\mu}_n}{(\hat{\mu}_1)^2} O(\chi), k = 1, \ldots, n, \tag{78}$$

where

$$\chi = \vartheta/\lambda_1 + O(1)N\overline{N}^2\delta. \tag{79}$$

If λ_1 is sufficiently well separated from zero, i.e., if ϑ/λ_1 is a modest multiple of the machine precision ε, then (77) and (78) imply that $\parallel r^0 \parallel$ is close to $\parallel \overline{r}^0 \parallel$ and $\parallel x - \hat{x}^k \parallel_A$ is close to $\parallel \overline{x} - \overline{x}^k \parallel_{\overline{A}}$.

Combining (77), (78), with (71) and (72) gives

$$(T_N^{-1})_{11} = (\hat{T}_n^{-1})_{11} + \parallel x - \hat{x}^n \parallel_A^2 / \parallel r^0 \parallel^2 + \frac{\hat{\mu}_n}{\hat{\mu}_1} O(\widetilde{\chi}) - P_n(1/\lambda), \tag{80}$$

where $O(\widetilde{\chi})$ counts for the resulting small terms which are proportional to χ. (80) is the desired finite precision analogy of (50).

As a consequence, the **total error** of the Gauß quadrature for (58) computed via the CG method (or, equivalently, the convergence of the continued fraction $\hat{C}_n = (\hat{T}_n^{-1})_{11}$, computed by the finite precision CG process, to the value $C_N = (T_N^{-1})_{11}$, cf. (44)-(46)) is determined by the **actual A-norm of the error** in the finite precision conjugate gradient process.

Using (80), one can also explain the effect of rounding errors in computing Gauß quadrature for $f(\lambda) = \lambda^{-1}$ and the distribution function $\omega(\lambda)$ with

finite points of increase (or, equivalently, in computing corresponding continued fraction). Computed results of the Gauß quadrature (or the continued fraction calculation) are affected considerably by rounding errors if and only if the same is true for the energy norm of the error of the underlying conjugate gradient run. For a general function $f(\lambda)$, the effect of rounding errors is described by the difference between the values of $R_n(f)$ and $\overline{R}_n(f)$ given by (16) and (68).

This approach can be used for "reconstructing" the A-norm of the error in the finite precision CG run. Rewriting (67) for the special case $f(\lambda) = \lambda^{-1}$, and using (72), we obtain

$$\| \overline{x} - \overline{x}^n \|_A^2 = \| \overline{r}^0 \|^2 \left[(\overline{T}_N^{-1})_{11} - (\hat{T}_n^{-1})_{11} \right]. \tag{81}$$

Considering (77)-(79),

$$\| x - \hat{x}^n \|_A^2 = \| r^0 \|^2 \left[(\overline{T}_N^{-1})_{11} - (\hat{T}_n^{-1})_{11} \right] + \frac{\hat{\mu}_n}{\hat{\mu}_1} O(\tilde{\tilde{\chi}}), \tag{82}$$

where $O(\tilde{\tilde{\chi}})$ counts for the small terms proportional to χ. The value $(\overline{T}_N^{-1})_{11}$ is unknown, but it can be approximated by $(\hat{T}_k^{-1})_{11}$ for some $k > n$

$$(\overline{T}_N^{-1})_{11} = (\hat{T}_k^{-1})_{11} + O(1) \frac{\hat{\rho}_{k-1}^2}{\hat{\mu}_1}, \tag{83}$$

where $\hat{\mu}_1$ is, from now on, the smallest Ritz value computed at the step k (we assume $\hat{\mu}_1 > 0$. Note that the Ritz values are not double-indexed, i.e., we omit the index of the iteration step. This simplification is possible because the meaning is clear from the context). Finally, (82) and (83) give (with replacing $\hat{\mu}_n$ by $\hat{\mu}_k$)

$$\| x - \hat{x}^n \|_A^2 = \| r^0 \|^2 \left[(\hat{T}_k^{-1})_{11} - (\hat{T}_n^{-1})_{11} \right] + O(1) \frac{\| \hat{r}^{k-1} \|^2}{\hat{\mu}_1} + \frac{\hat{\mu}_k}{\hat{\mu}_1} O(\tilde{\tilde{\chi}}). \tag{84}$$

$(\hat{T}_k^{-1})_{11}$ and $(\hat{T}_n^{-1})_{11}$ can be easily computed as the continued fractions $\hat{C}_k$ and $\hat{C}_n$.

The relation (84) can be interpreted in the following way. Let for $n = 1, 2, \ldots, n_k$ the first term in (84) is dominating, i.e.

$$(85) \qquad \hat{t}_{n,k} = \| r^0 \|^2 [(\hat{T}_k^{-1})_{11} - (\hat{T}_n^{-1})_{11}] \gg \frac{\| \hat{r}^{k-1} \|^2}{\hat{\mu}_1} + \frac{\hat{\mu}_k}{\hat{\mu}_1} O(\tilde{\tilde{\chi}}).$$

Then $(\hat{t}_{n,k})^{1/2}$ is a good approximation to $\| x - \hat{x}^n \|_A$, $n = 1, \ldots, n_k$, and it "restores" the behavior of the energy norm of the actual error in the first n_k steps. For $n_k < n < k$, the value of $(\hat{t}_{n,k})^{1/2}$ can be considered as a lower bound, and $\| \hat{r}^n \| / \sqrt{\hat{\mu}_1}$ as an upper bound for $\| x - \hat{x}^n \|_A$.

This is, of course, only a curiosity, without a practical application. Indeed,

$$\| x - \hat{x}^k \|_A = O(1) \frac{\| \hat{r}^{k-1} \|}{\sqrt{\hat{\mu}_1}}$$

implies that $n_k < k$, and we cannot gain a reasonable information about the energy norm of the error at the k-th step from (84). Moreover, (84) does not provide us with any approximation to $\| x - \hat{x}^n \|_A$ under the level, say,

$$(86) \qquad O(\| x - \hat{x}^0 \|_A) \sqrt{\varepsilon},$$

which is the square root of the rough estimate of the roundoff in computing the value of $\hat{t}_{n,k}$ from $\hat{\alpha}_l, \hat{\beta}_l$.

We illustrate the behavior of $\hat{t}_{n,k}$ by the following experiments. For the first example we consider a diagonal matrix $A = diag(\lambda_1, \ldots, \lambda_N)$ from the class of matrices described in [St-91] and used in [GS-92]. We used the dimension $N = 48$; $\lambda_1 = 0.1$, $\lambda_N = 100$. The distribution of eigenvalues was choosen so that the effect of rounding errors had been crucial (it corresponded to the parameter ρ from [St-91] set to $\rho = 0.875$). Figures 1-3 show the actual A-norm of the error $\| x - \hat{x}^n \|_A$ (solid lines), the computed value of $(\hat{t}_{n,k})^{1/2}$ (dashed lines) and the "precise" value of $(\hat{t}_{n,k})^{1/2}$ (dotted lines). The last one is computed from (85) with $(\hat{T}_k^{-1})_{11}$ replaced by T_N^{-1}, which is determined from the "exact", i.e. doubly reorthogonalized, process, see [GS-92]. Figure 1 corresponds to $k = 59$, Figure 2 to $k = 79$ and Figure 3 to $k = 99$. The dashed-dotted lines denotes

$$(87) \qquad \max(\| \hat{r}^{k-1} \| / \sqrt{\mu_1}, \ 10^{-6}),$$

where 10^{-6} substitutes for the ultimate limiting precision level (86).

Figures 1-3 were plotted for a randomly choosen solution vector; for the other choices the situation was similar. We point out that above the level

(86) $(\hat{t}_{n,k})^{1/2}$ "restores" the value of $\| x - \hat{x}^n \|_A$ precisely (dashed, dotted and solid lines almost coincide), while below this level the computed value of $(\hat{t}_{n,k})^{1/2}$ gives for $\| \hat{r}^{k-1} \| / \sqrt{\hat{\mu}_1} > 10^{-6}$ the lower bound for $\| x - \hat{x}^n \|_A$. For $\| \hat{r}^{k-1} \| / \sqrt{\hat{\mu}_1} < 10^{-6}$ it contains no useful information.

As a second example we used the model problem with the 5-point difference approximation to Laplacian on the mesh 20×20 ($N = 400$). Results are for the solution vector $(1, 1, \ldots, 1)^T$, $k = 24, 34$ and 44 shown on Figures 4-6 (the dotted line was not computed).

6 Concluding remarks

In this paper we emphasized relations between quadratures, orthogonal polynomials, Jacobi matrices, continued fractions, Lanczos and conjugate gradient methods. We have shown that the problem of estimating quadratic forms can be considered as a problem of computing Gauß and Gauß-Radau quadratures, and that it can be solved in an elegant way using the Lanczos method. Exploiting relations between the areas mentioned above, we can easily reformulate a question from one area into the language of the other area. This approach can simplify the original problem and make its solution very easy.

We analyzed the effect of rounding errors to the Gauß quadrature calculations for the distribution function with finite points of increase. Using the backward error analysis of the conjugate gradient method (developed by Greenbaum), we proved that the **total error** of the Gauß quadrature computed via the CG method in finite precision arithmetic can be described as the **truncation error** of the Gauß quadrature for a different distribution function. As a consequence, we have shown that for the particular case $f(\lambda) = 1/\lambda$, the total accuracy of the Gauß quadrature, computed by the finite precision CG process, is determined by the energy norm of the error of the CG process.

Acknowledgment

Zdeněk Strakoš would like to thank Anne Greenbaum for helpful discussions and advice, and Jitka Drkošová, Miroslav Rozložník and Ladislav Lukšan for useful comments.

Both authors thank the referee for his/her thorough report, which helped to improve the presentation of the paper.

References

[AB-88] K. Amos , L. Berge, F.A. Brieva, A. Katsogiannis, L. Petris and L. Rikus, Off-shell Properties of the Two-Nucleon T Matrix, Physical Review C **37**, 934-948.

[D-93] P. Deufhard, Cascadic Conjugate Gradient Methods for Elliptic Partial Differential Equations I. Algorithm and Numerical Results, Preprint SC93-23, Konrad Zuse Zentrum, Berlin, 1993.

[DEG-72] G. Dahlquist, S.C. Eisenstat and G.H. Golub, Bounds for the Error of Linear Systems of Equations Using the Theory of Moments, J. of Math. Anal. and Appl. **37**, 151-166,1972.

[DGN-78] G. Dahlquist, G.H. Golub and S.G. Nash, Bounds for the Error in Linear Systems, in Proc. of the Workshop on Semi-Infinite Programming, R.Hettich ed., Springer, 154-172, 1978.

[DR-84] P.J. Davis and P. Rabinowitz, Numerical Integration, Academic Press, N.Y., 1984.

[FG-92] B. Fischer and G.H. Golub, On the Error Computation for Polynomial Based Iterative Methods, Manuscript NA 92-21, Dept. of Comp. Sc., Stanford University, 1992.

[FF-93] B. Fischer and R. Freund, On Adaptive Weighted Polynomial Preconditioning for Hermitian Positive Definite Matrices, SIAM J. Sci. Comput. (to appear).

[FH-93] R. Freund and M. Hochbruck, Gauß Quadratures Associated with the Arnoldi Process and the Lanczos Algorithm, In: Linear Algebra for Large Scale and Real-Time Applications, M.S. Moonen et al. (eds.), Kluwer Academic Publishers, Amsterdam, 377-380, 1993.

[Ga-68] W. Gautschi, Construction of Gauß -Christoffel Quadrature Formulas, Math. Comp. **22**, 251-270, 1968.

[Ga-81] W. Gautschi, A Survey of Gauß -Christoffel Quadrature Formulae, In: E.B. Christoffel - The Influence of His Work on Mathematics and the Physical Sciences, P.L. Bultzer and F. Fehér (eds.), Birkhauser, Boston, 73 - 157, 1981.

[Ga-82] W. Gautschi, On Generating Orthogonal Polynomials, SIAM J. Sci. Stat. Comput. **3**, 289-317, 1982.

[GM-91] G.H. Golub and U.v. Matt, Quadratically Constrained Least Squares and Quadratic Problems, Numer. Math. 59, 561 - 580, 1991.

[GM-93] G.H. Golub and G. Meurant, Matrices, Moments and Quadrature, Manuscript SCCM93-07, Dept. of Comp. Sc., Stanford University, 1993.

[GW-69] G.H. Golub and J.H. Welsch, Calculation of Gauß Quadrature Rules, Math. Comp. **23**, 221-230, 1969.

[Go-73] G.H. Golub, Some Modified Matrix Eigenvalue Problems, SIAM Review **15**, 318-334, 1973.

[Go-74] G.H. Golub, Bounds for Matrix Moments, Rocky Mountain J. of Math. **4**, 207-211, 1974.

[GK-89] G.H. Golub and M.D. Kent, Estimates of Eigenvalues for Iterative Methods, Math. Comp. **53**, 619-626, 1989.

[Gr-89] A. Greenbaum, Behavior of Slightly Perturbed Lanczos and Conjugate Gradient Reccurences, Linear Algebra and its Appl. **113**, 7-63, 1989.

[GS-92] A. Greenbaum and Z. Strakoš, Predicting the Behavior of Finite Precision Lanczos and Conjugate Gradient Computations, SIAM J. Matrix Anal. Appl. **13**, 121-137, 1992.

[GS-93] G.H. Golub and Z. Strakoš, Estimates in Quadratic Formulas, Manuscript SCCM93-08, Dept. of Comp. Sc., Stanford University, 1993.

[HY-81] L.A. Hageman and D.M. Young, Applied Iterative Methods, Academic Press, N.Y., 1981.

[HS-52] M.R. Hestenes and E. Stiefel, Methods of Conjugate Gradients for Solving Linear Systems, J. of Research of the Natl. Bureau of Standards **49**, 409-436, 1952.

[HSa-84] J. Horáček and T. Sasakawa, Method of Continued Fractions with Application to Atomic Physics, Physical Review A **30**, 2274-2277, 1984.

[HSa-85] J. Horáček and T. Sasakawa, Method of Continued Fractions for On-and off-shell T Matrix of Local and Nonlocal Potentials, Physical Review C **37**, 70-75, 1985.

[H-87] J. Horáček, A Note on Iterative Solution of Scatering Equations, Acta Physica Polonica A **71**, 3-8, 1987.

[H-89] J. Horáček, Extrapolation Technique for Solution of Scattering Integral Equations, J. Phys. A: Math, Gen. **22**, 1-9, 1989.

[K-94] L.Knizhnerman, The Simple Lanczos Process: Error Estimates for The Gaussian Quadrature Formula and Their Applications (in preparation).

[K-89] M.D. Kent, Chebyshev, Krylov and Lanczos, Matrix Relationship and Computations, Ph.D.Thesis, Stanford University, Stanford, 1989.

[KG-83] J. Kautský and G.H. Golub, On the Calculation of Jacobi Matrices, Linear Alg. Appl. **52/53**, 439-455, 1983.

[KS-53] S. Karlin and L.S. Shapley, Geometry of Moment Spaces, Memoirs of the American Mathematical Society 12, Amer. Math. Soc., Providence, 1953.

[MF-86] G. Moro and J.H. Freed, The Lanczos Algorithm in Molecular Dynamics: Calculation of Spectral Densities, In: Large Scale Eigenvalue Problems, J. Cullum and R.A. Willoughby (eds.), Elsevier (North Holland), Amsterdam, 143 - 161, 1986.

[MP-89] H.D. Mayer and S. Pal, A Band-Lanczos Method for Computing Matrix Elements of a Resolvent, J. Chem. Phys. **91**, 6195-6204, 1989.

[N-93] Y. Notay, On the Convergence Rate of the Conjugate Gradients in Presence of Rounding Errors, Numer. Math. (to appear).

[Pa-87] B.N. Parlett, Misconvergence in the Lanczos Algorithm, Res. Rep. PAM-404, U. of California, Berkeley, 1987.

[Pg-71] C.C. Paige, The Computation of Eigenvalues and Eigenvectors of Very Large Sparse Matrices, Ph.D. Thesis. U. of London, 1971.

[Pg-72] C.C. Paige, Computational Variants of the Lanczos Method for the Eigenproblem, J. Inst. Math. Applics **10**, 373-381, 1972.

[Pg-76] C.C. Paige, Error Analysis of the Lanczos Algorithm for Tridiagonalizing a Symmetric Matrix, J. Inst. Math. Applics. **18**, 341-349, 1976.

[Pg-80] C.C. Paige, Accuracy and Effectiveness of the Lanczos algorithm for the Symmetric Eigenproblem, Linear Algebra Appl. **34**, 235-258, 1980.

[R-79] W. Reinhardt, L^2 Discretization of Atomic and Molecular Electronic Continua: Moment, Quadrature and J-Matrix Techniques, Comp. Physics Communications **17**, 1 - 21, 1979.

[RW-92] P.D. Robinson and A.J. Wathen, Variational Bounds on the Entries of the Inverse of a Matrix, IMA J. of Numer. Anal. **12**, 1992.

[S-55] E. Stiefel, Relaxationsmethoden bester Strategie zur Lösung Linearer Gleichungsysteme, Comm. Math. Helv. **29**, 157-179, 1955.

[Si-84a] H.D. Simon, The Lanczos Algorithm with Partial Reorthogonalization, Math. Comp. **42**, 115-136, 1984(a).

[Si-84b] H.D. Simon, Analysis of the Symmetric Lanczos Algorithm with Reorthogonalization Methods, Linear Algebra and Its Appl. **61**, 101-131, 1984(b).

[St-91] Z. Strakoš, On the Real Convergence Rate of the Conjugate Gradient Method, Linear Algebra and Its Appl. **154-156**, 535-549, 1991.

[SG-92] Z. Strakoš and A. Greenbaum, Open Questions in the Convergence Analysis of the Lanczos Process for the Real Symmetric Eigenvalue Problem, IMA Preprint Series 934, IMA, U of Minnesota, 1992.

[Sz-39] G. Szegö, Orthogonal Polynomials, American Mathematical Society Colloq. Publ. 23, Providence, 1939.

[W-62] H.S. Wilf, Mathematics for the Physical Sciences, J. Wiley and Sons, N.Y., 1962.