STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING TOEGEPASTE WISKUNDE

Report TW 104

On the acceleration of Richardson's method I

Theoretical part

by

P.J. van der Houwen

March 1967

The Mathematical Centre at Amsterdam, founded the 11th of February, 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.) and the Central Organization for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

Introduction

Richardson's method is used to solve iteratively matrix equations of the type

$$Lu = f,$$

where L is a symmetric matrix with positive eigenvalues. In applications of this method one needs the values of the lowest eigenvalue λ_1 of L and the spectral norm $\sigma(L)$ of L. For ill-conditioned matrices L, i.e. $\sigma(L) >> \lambda_1$, the rate of convergence is very slow and an accelerating process is highly desirable. In reference [3] Frank described two accelerations of Richardson's method. However, when tried on a computer the method turned out to be unsatisfactory.

In this paper a different accelerating procedure is proposed which was used successfully on a computer. One advantage over Richardson's method is the fact that no apriori knowledge of the first eigenvalue λ_1 is needed. This eigenvalue is estimated during the first phase of the method. Further, one or more negative eigenvalues λ of L are also admitted.

As a consequence the method can also be used to estimate the smallest eigenvalues of symmetric matrices L.

In the last section the process is adapted to find upper and lower bounds for the estimated eigenvalues.

This paper contains theoretical results only. There will appear a second paper in the near future dealing with applications of this method to elliptic boundary value problems wherein numerical results are given.

1. Definition of iterative processes

In this section we give definitions concerning iterative methods for solving the matrix equation

$$Lu = f, \tag{1.1}$$

where L is a symmetric matrix, u the unknown vector and f a known vector. For a detailed discussion of these definitions we refer to the literature [2] (see also the appendix to this paper).
To (1.1) we associate the following iterative process

$$u_{k+1} = (\alpha_k - \omega_k L)u_k + (1 - \alpha_k)u_{k-1} + \omega_k f, \quad k = 1, 2, \ldots , \tag{1.2}$$

where the vectors u_0 and u_1 are the beginapproximations of the process. When the sequence u_k, u_{k+1}, ... converges, the limitvector will be the solution of (1.1).
The iterative scheme (1.2) is called of first degree (or order) if $\alpha_k = 1$ for all k, and of second degree if $\alpha_k \neq 1$. We shall consider mainly non-stationary or semi-iterative processes i.e. the parameters α_k and ω_k depend on k. It is convenient to write

$$u_k = u + v_k, \tag{1.3}$$

where v_k can be considered as the error of the approximation u_k. Then v_k satisfies the homogeneous recurrence relation

$$v_{k+1} = (\alpha_k - \omega_k L)v_k + (1 - \alpha_k)v_{k-1}. \tag{1.4}$$

If we suppose that v_1 is obtained from v_0 as $v_1 = (1 - \omega_0 L)v_0$, i.e. by applying to v_0 an operator which is linear in L then v_k is obtained from v_0 by applying a polynomial-operator of degree k in L. Thus we may write

$$u_k = u + P_k(L)v_0. \tag{1.5}$$

In connection with this expression one defines the average rate of convergence for K iterations of the iteration process (1.5) as the quantity [2]

$$R(K) = - \frac{\ln \sigma(P_K(L))}{K} , \tag{1.6}$$

where $\sigma(P_K(L))$ is the spectral norm of the matrix $P_K(L)$.
In the following sections we construct polynomials $P_K(L)$ with small spectral norms, which can be used to obtain fast converging iterative schemes.

2. Richardson's method

We shall briefly describe Richardson's method for positive definite matrices L. For a more detailed discussion we refer again to the literature [2].

When in the polynomial operator $P_K(L)$ the operator L is replaced by the real variable λ, we obtain a real polynomial $P_K(\lambda)$ with the property $P_K(0) = 1$. The eigenvalues of $P_K(L)$ are given by $P_K(\lambda_i)$, $i = 1, 2, \ldots, M$, where λ_i is an eigenvalue of L. In figure 1 the dots on the curve $P_K(\lambda)$ correspond to the eigenvalues $P_K(\lambda_i)$. In this section we assume that $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_M = \sigma(L)$.

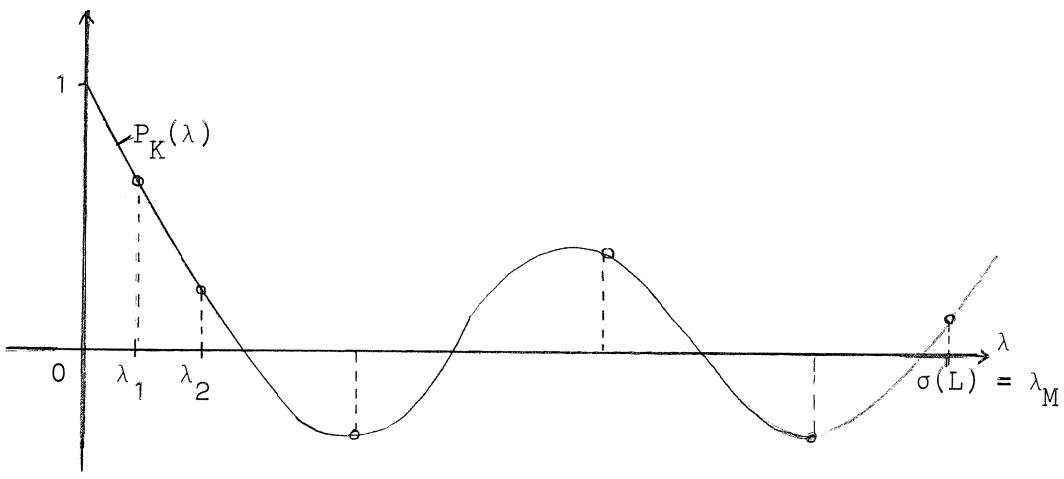

fig. 1

If we know the eigenvalues λ_i, we may take the zeros of $P_K(\lambda)$ to coincide with the values $\lambda = \lambda_i$, resulting in a zero spectral norm for $P_K(L)$. In actual application, however, there exists a large number of eigenvalues λ_i, thus we must perform many iterations. Moreover in most cases we only have a rough estimate for the first eigenvalue λ_1 and the last eigenvalue $\lambda_M = \sigma(L)$.
Another method to keep $\sigma(P_K(L))$ small is to minimize the polynomial $P_K(\lambda)$ over the continuous interval $[\lambda_1, \sigma(L)]$. For this we only need to know the first and last eigenvalue of L.
Richardson (1910) [5] chose the zeros of $P_K(\lambda)$ to coincide with

$$\lambda = \lambda_1 + i \frac{\sigma(L) - \lambda_1}{K + 1} \quad , \quad i = 1, 2, \ldots, K,$$

but there is a better operator $P_K(L)$, based on the following theorem of W. Markoff (cited in [2]).

Theorem I. The polynomial

$$C_K(a,b,\lambda) = \frac{T_K\left(\frac{b + a - 2\lambda}{b - a}\right)}{T_K\left(\frac{b + a}{b - a}\right)} \quad ,$$

where $T_K(y) = \cos(K \arccos y) = \begin{cases} \cos(K \arccos y) & \text{for } |y| \leq 1 \\ \cosh(K \operatorname{arccosh} y) & \text{for } |y| > 1 \end{cases}$,

has, of all the polynomials $P_K(\lambda)$ of degree K in λ satisfying $P_K(0) = 1$, a minimal maximum-norm over the interval $a \leq \lambda \leq b$.

The function $T_K(y)$ is the Chebyshef-polynomial of degree K.
One defines the Richardson method with respect to the operator L by the formulae

(2.1) $$P_K(L) = C_K(a,b,L) \ , \quad a = \lambda_1 \ , \quad b = \sigma(L).$$

For applications we must know the expressions for the parameters α_k and ω_k. First we consider the linear Richardson method i.e. $\alpha_k = 1$. The zeros of $P_K(\lambda)$ are given by

(2.2) $$\lambda = 1/\omega_k \ , \quad k = 0, 1, \ldots, K-1,$$

and $C_K(a,b,\lambda)$ assumes a zero value for the points

(2.3) $$\lambda = \frac{1}{2}(a + b) + \frac{1}{2}(a - b)\cos\frac{2l+1}{2K}\pi, \quad l=0,1,\ldots,K-1.$$

Hence the parameters ω_k (the so-called relaxation parameters) are given by the following values

(2.4) $$\left(\frac{1}{2}(a + b) + \frac{1}{2}(a - b)\cos\left(\frac{2l+1}{2K}\pi\right)\right)^{-1}, \quad l=0,1,\ldots,K-1,$$

where $a = \lambda_1$, $b = \sigma(L)$ and k is not necessarily equal to 1.

Next we consider the non-linear process with $a_k \neq 1$.
The polynomials $P_k(\lambda)$ must satisfy the relation

(2.5) $$P_{k+1}(\lambda) = (\alpha_k - \omega_k\lambda)P_k(\lambda) + (1 - \alpha_k)P_{k-1}(\lambda),$$

obtained from (1.4) by constituting $v_k = P_k(L)v_0$ and replacing L with λ.
On the other hand we derive from the well-known recurrence relation

(2.6) $$T_{k+1}(y) = 2y\,T_k(y) - T_{k-1}(y) \ , \quad k \geq 1,$$

the following formula for $C_k(a,b,\lambda)$

$$(2.7)\quad C_{k+1}(a,b,\lambda) = \left(2y_0 - \frac{4\lambda}{b-a}\right)\frac{T_k(y_0)}{T_{k+1}(y_0)}\,C_k(a,b,\lambda)$$

$$- \frac{2y_0\,T_k(y_0) - T_{k+1}(y_0)}{T_{k+1}(y_0)}\,C_{k-1}(a,b,\lambda),$$

where $y_0 = (b + a)/(b - a)$ and $k \geq 1$.
If we define for $k \geq 1$

$$\alpha_k = 2y_0\frac{T_k(y_0)}{T_{k+1}(y_0)} \ , \quad \omega_k = \frac{4}{b-a}\frac{T_k(y_0)}{T_{k+1}(y_0)} \ ,$$

the relations (2.5) and (2.7) are exactly the same. Therefore, if

$$P_0(\lambda) = 1, \quad P_1(\lambda) = 1 - \frac{2}{b+a}\lambda,$$

we obtain polynomials $P_k(\lambda)$, which are identical to the polynomials $C_k(a,b,\lambda)$ for every k. Putting $a = \lambda_1$ and $b = \sigma(L)$ we get Richardson's method of the second degree.

The linear form of Richardson's method was first used by Young (1953) [8]. It has the advantage in being simple and it requires less storage space than the second degree iteration scheme. The numerical stability, however, depends strongly on the distribution of the relaxation parameters ω_k, particularly when K is large. Young avoids this problem by repeating the iteration process with relaxation parameters $\{\omega_k\}_{k=0}^{K-1}$ for a stable order of k, but this reduces the average rate of convergence (section 3).

In a forthcoming paper we shall discuss the dependence of the stability on the distribution of the relaxation parameters.

The non-linear case was developed by Varga (1957) [7] and tested by Frank (1960) [3]. This procedure is obviously stable if $\lambda_i > 0$ for all i. Further one needs no apriori knowledge of K as was required in the first order process.

3. The rate of convergence

According to the definition of the operator $C_K(a,b,L)$ we have

$$\sigma(C_K(a,b,L)) \leq T_K^{-1}\left(\frac{b+a}{b-a}\right). \tag{3.1}$$

For large K we find (approximately)

$$T_K\left(\frac{b+a}{b-a}\right) = \cosh\left(K \operatorname{arcosh}\left(\frac{b+a}{b-a}\right)\right) \approx \frac{1}{2}\exp\left(K \operatorname{arcosh}\left(\frac{b+a}{b-a}\right)\right), \tag{3.2}$$

thus

$$R(K) \geq -\frac{\ln\left(T_K^{-1}\left(\frac{b+a}{b-a}\right)\right)}{K} \approx \operatorname{arcosh}\left(\frac{b+a}{b-a}\right) - \frac{\ln 2}{K}\,. \tag{3.3}$$

Putting $a = \lambda_1$ and $b = \sigma(L)$ we obtain a lower bound for the average rate of convergence for K iterations of Richardson's method.

We consider the behaviour of $R(\infty)$ as a function of $y_0 = (b+a)/(b-a)$.

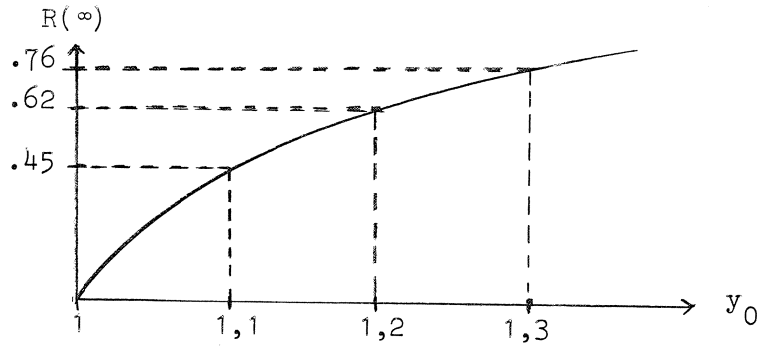

fig. 2

From figure 2 we see that the asymptotic rate of convergence has the largest increase in the neighbourhood of $y_0 = 1$.
If $a << b$ the term $\ln 2/K$ decreases the rate of convergence considerably in actual computation. Hence K must be as large as possible. This is the reason that repeating the iteration process with a lower K is very disadvantageous for the average rate of convergence. For example, repeating Richardson's process over K/3 iterations three times yields an average rate of convergence for K iterations which is given by

$$R(K) \geq \operatorname{arccosh}\left(\frac{b + a}{b - a}\right) - 3\,\frac{\ln 2}{K} .$$

4. The elimination method

In this section we propose a variation of Richardson's method, which has a considerably larger asymptotic rate of convergence and which is applicable not only to positive matrix equations but also to equations where L may have negative eigenvalues as well.

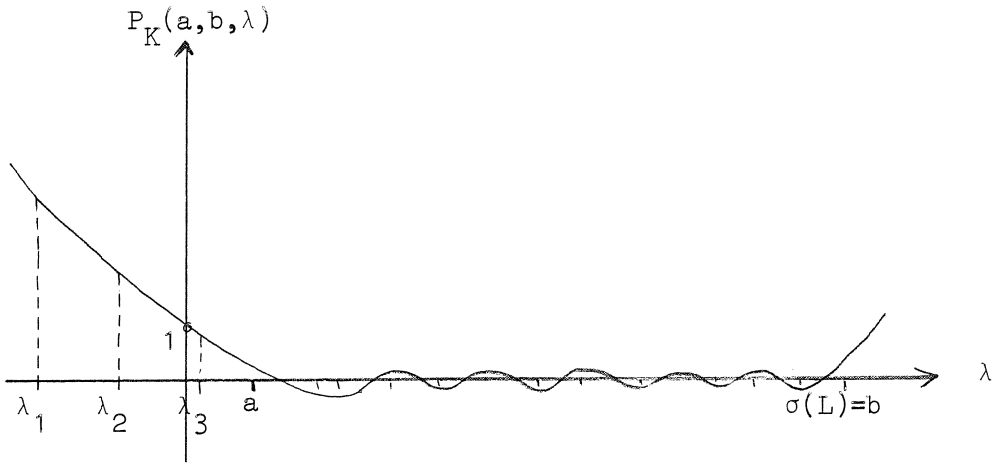

fig. 3

The essence of the method is the reduction of the late eigenfunctions of L corresponding to eigenvalues λ inside the interval $[a,b]$, where $a > \lambda_1$ and $a > 0$, followed by the elimination of the remaining eigenfunctions of L. This may be achieved by means of an operator $C_K(a,b,L)$ and an elimination operator $E_{K^*}(L)$, where the eigenvalues λ_i outside the interval $[a,b]$ are zeros of $E_{K^*}(\lambda)$.
K^* is the degree of the operator $E_{K^*}(L)$.
Using (1.6) and (3.2) we derive the average rate of convergence for this method

$$(4.1) \qquad R(K + K^*) = \operatorname{arccosh} y_0 - \frac{K^* \operatorname{arccosh} y_0 + \ln \sigma(E_{K^*}) + \ln 2}{K + K^*} ,$$

where $y_0 = (b+a)/(b-a)$.
As in Richardson's method, we choose $b = \sigma(L)$.
We now discuss the value of a for $\lambda_1 > 0$ and $\lambda_1 << b$: The asymptotic rate of convergence for $k \to \infty$ of the elimination method is $\operatorname{arccosh} (\frac{b+a}{b-a})$. For Richardson's method it is $\operatorname{arccosh} (\frac{b + \lambda_1}{b - \lambda_1})$.
Let $a = n\lambda_1$, then we find for not too large values of n

$$(4.2)\qquad \frac{\operatorname{arccosh}\left(\frac{b+a}{b-a}\right)}{\operatorname{arccosh}\left(\frac{b+\lambda_1}{b-\lambda_1}\right)} \approx \frac{2\sqrt{\frac{a}{b}}}{2\sqrt{\frac{\lambda_1}{b}}} = \sqrt{n}.$$

Thus using instead of $a = \lambda_1$ the value $a = n\lambda_1$, we gain a factor $\sqrt{n}$ in the asymptotic rate of convergence. However, the number of eigenfunctions to be eliminated becomes larger for increasing n. In practice the optimal value for n is determined from the distribution of the lower eigenvalues of L, bearing in mind that the gainfactor increases most rapidly for small values of n.

Next we consider the elimination of the lower eigenfunctions of L. We assume that the eigenvalues of the eigenfunctions to be eliminated are known (see the following section). Suppose we wish to eliminate the eigenfunction e_i with eigenvalue λ_i. This may be done by means of an operator $E_{K_i^*}(\lambda_i, L)$ of degree K_i^* in L, satisfying the conditions

$$(4.3)\qquad E_{K_i^*}(\lambda_i, 0) = 1 \quad , \quad E_{K_i^*}(\lambda_i, \lambda_i) = 0.$$

In this connection the following is useful.

Theorem II. The polynomial $E_{K_i^*}(\lambda_i, \lambda)$ defined by

$$E_{K_i^*}(\lambda_i, \lambda) \equiv C_{K_i^*}(a_i^*, b, \lambda),$$

$$(4.4)\qquad a_i^* = \frac{2\lambda_i + b\left(\cos\frac{\pi}{2K_i^*} - 1\right)}{\cos\frac{\pi}{2K_i^*} + 1}$$

satisfies the conditions (4.3).

Of all polynomials of degree K_i^* satisfying (4.3), this polynomial has the smallest maximum-norm over the interval $[c_i, b]$ when

(4.5)
$$\begin{cases} c_i > 0\ , \quad c_i > \lambda_i, \\ \frac{1}{4}\gamma_i\left(1 - \sqrt{1 + 8/\gamma_i}\right) \le \cos\frac{\pi}{2K_i^*} \le \frac{1}{4}\gamma_i\left(1 + \sqrt{1 + 8/\gamma_i}\right), \end{cases}$$

where $\gamma_i = (b - c_i)/(b - \lambda_i)$.

Proof.

It is clear that $E_{K_i^*}(\lambda_i, 0) = 1$.

The second condition of (4.3) follows from the fact that the zeros of $C_{K_i^*}(a_i^*, b, \lambda)$ are given by

(4.6)
$$\lambda = \frac{1}{2}(b + a_i^*) - \frac{1}{2}(b - a_i^*)\cos\left[(2n+1)\frac{\pi}{2K_i^*}\right], \quad n=0,1,\ldots .$$

The smallest zero is assumed for n = 0. Substituting (4.4) into (4.6) and putting n = 0 gives λ_i as the first zero of $E_{K_i^*}(\lambda_i, \lambda)$.

To prove the minimax-property we assume the existence of a polynomial $S_{K_i^*}(\lambda)$ of degree K_i^* in λ satisfying (4.3) and the inequality

$$||S_{K_i^*}(\lambda)|| < ||C_{K_i^*}(a_i^*, b, \lambda)||,$$

where $||\ ||$ means the maximum-norm over the interval $[c_i, b]$.
Consider the polynomial

$$Q(\lambda) = S_{K_i^*}(\lambda) - C_{K_i^*}(a_i^*, b, \lambda).$$

$Q(\lambda)$ has positive values for those points of the interval $[c_i, b]$, where $C_{K_i^*}(a_i^*, b, \lambda)$ is minimal and negative values in the points where $C_{K_i^*}(a_i^*, b, \lambda)$ is maximal.

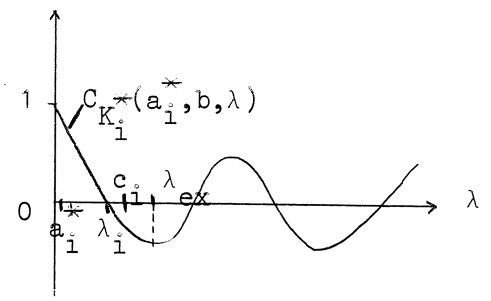

fig. 3a

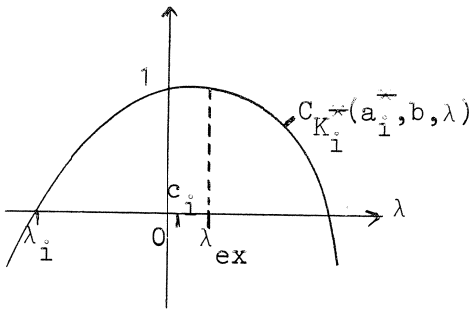

fig. 3b

If $C_{K_i^*}(a_i^*,b,\lambda)$ has n extrema in the interval $[c_i,b]$, then $Q(\lambda)$ has at least (n-1) zeros in the interval $[c_i,b]$. The first extremum is assumed for the point $\lambda = a_i^*$, the second for the point

$$\lambda_{ex} = \frac{1}{2}(b + a_i^*) - \frac{1}{2}(b - a_i^*)\cos\frac{\pi}{K_i^*}. \tag{4.7}$$

Suppose $c_i < \lambda_{ex}$, then $C_{K_i^*}(a_i^*,b,\lambda)$ has K_i^* extrema in the interval $[c_i,b]$, hence $Q(\lambda)$ has (K_i^*-1) zeros in $[c_i,b]$. In addition $Q(\lambda)$ has two other zeros in the points $\lambda = 0$ and $\lambda = \lambda_i$ (see (6.5)), therefore $Q(\lambda)$ has $(K_i^* + 1)$ different zeros. On the other hand $Q(\lambda)$ is at most of degree K_i^*, implying at most K_i^* zeros. This contradiction eliminates the existence of a polynomial $S_{K_i^*}(\lambda)$. Hence the last part of the theorem is proved.

We now prove that $c_i < \lambda_{ex}$. Substituting (4.4) into (4.7) and writing $\cos\frac{\pi}{K_i^*}$ as $2\cos^2\frac{\pi}{2K_i^*} - 1$ yields

$$\lambda_{ex} = \frac{2(\lambda_i - b)\cos^2\frac{\pi}{2K_i^*} + b\cos\frac{\pi}{2K_i^*} + b}{\cos\frac{\pi}{2K_i^*} + 1}. \tag{4.7'}$$

Using (4.7') the inequality $c_i < \lambda_{ex}$ becomes

$$2(b - \lambda_i)\cos^2\frac{\pi}{2K_i^*} - (b - c_i)\cos\frac{\pi}{2K_i^*} - (b - c_i) \leq 0.$$

From this inequality we find the second part of (4.5).

We shall now investigate condition (4.5) for large values of b (in most applications b is very large with respect to $|\lambda_i|$ and c_i). If $b >> |\lambda_i|$ and $b >> c_i$ we can approximate

$$\gamma_i \tilde{=} 1 - \frac{c_i - \lambda_i}{b}$$

$$\sqrt{1 + 8/\gamma_i} \tilde{=} 3 + \frac{4}{3}\frac{c_i - \lambda_i}{b} .$$

Substituting this into (4.5) we obtain

$$(4.5') \qquad -\frac{1}{2} \leq \cos\frac{\pi}{2K_i^*} \leq 1 - \frac{2}{3}\frac{c_i - \lambda_i}{b}$$

or equivalently

$$(4.5'') \qquad 1 \leq K_i^* \leq \frac{1}{4}\pi\sqrt{\frac{3b}{c_i - \lambda_i}} .$$

In practice we want the minimax-property to be valid over the interval $[a,b]$. The eigenfuncties corresponding to eigenvalues outside the interval $[a,b]$ need no reduction, since they may be eliminated successively. We find for K_i^* the condition

$$(4.8) \qquad K_i^* \leq \frac{1}{4}\pi\sqrt{\frac{3b}{a - \lambda_i}} .$$

In practice the value for K_i^* is determined by stability considerations (see section 5) or by the requirement that the average rate of convergence is as large as possible (section 6). Because of the usually large values of b, (4.8) is satisfied in most cases. If we use large values for K_i^* the operator $E_{K_i^*}(\lambda_i,L)$ still eliminates the eigenfunction e_i, but the theorem doesn't indicate if that operator is the "best" operator to eliminate the eigenfunction e_i.

A second stability condition is

$$b \geq \sigma(L).$$

In general we cannot choose $b = \sigma(L)$ without violating (5.4). For large values of $\sigma(L)$ this condition reduces to

$$K_i^* \geq \frac{1}{4}\pi\sqrt{\frac{3\sigma(L)}{-\lambda_i}}. \tag{5.5}$$

(Compare the derivation of (4.5")).
Evidently inequality (4.8) cannot be satisfied.

6. Optimal elimination operators

In this section we define the average rate of convergence with respect to the interval $[a,b]$, i.e.

$$-\frac{\ln\left(\underset{a\leq\lambda_i\leq b}{\text{Max}}\ C_K(a,b,\lambda_i)\right)}{K}.$$

This expression has the lower bound

$$r(0) = -\frac{\ln||C_K(a,b,\lambda)||}{K}, \tag{6.1}$$

where $||\ ||$ means the maximum-norm over the interval $[a,b]$.
After application of the elimination operators $E_{K_i^*}(\lambda_i,L)$, we have the lower bound

$$r(K^*) = -\frac{\ln||P_K(a,b,\lambda)|| + \ln||E_{K^*}(\lambda)||}{K + K^*}. \tag{6.2}$$

This expression is also a lower bound for the average rate of convergence as defined in section 1, thus we want (6.2) to be as large as possible.

5. Stable elimination operators with respect to the interval [a,b]

The danger of applying elimination operators ($E_{K_i^*}(\lambda_i,L)$) of small degree is that they have large spectral-norms ($\sigma(E_{K_i^*}(\lambda_i,L))$). This nullifies the effect of the minimizing operator $C_K(a,b,L)$. We can avoid this by requiring that the elimination operators are stable, i.e.

$$\sigma(E_{K_i^*}(\lambda_i,L)) \leq 1.$$

$\underline{\lambda_i > 0}$

The stability condition is simply

$$a_i^* \geq 0.$$

From (4.4) we obtain for $\lambda_i \ll b$

$$K_i^* \geq \frac{1}{2}\pi \arccos^{-1}\left(1 - 2\frac{\lambda_i}{b}\right) \simeq \frac{1}{4}\pi\sqrt{\frac{b}{\lambda_i}}. \tag{5.1}$$

The smallest admissable value for K_i^* is given by

$$K_i^* = \text{entier}\left(\frac{1}{4}\pi\sqrt{\frac{b}{\lambda_i}}\right) + 1. \tag{5.2}$$

(4.8) is satisfied for large values of b, if

$$a \leq 4\lambda_1. \tag{5.3}$$

$\underline{\lambda_i < 0}$

The stability condition is

$$\lambda_{ex} = 0.$$

From (4.7') we find the relation

$$b = \frac{2\lambda_i \cos^2\frac{\pi}{2K_i^*}}{2\cos^2\frac{\pi}{2K_i^*} - \cos\frac{\pi}{2K_i^*} - 1}. \tag{5.4}$$

We suppose that the eigenfunctions $e_1, e_2, \ldots, e_{i-1}$ are successively eliminated by the operators $E_{K_1^*}(\lambda_1,L)$, $E_{K_2^*}(\lambda_2,L)$, ..., $E_{K_{i-1}^*}(\lambda_{i-1},L)$. The average rate of convergence with respect to the interval $[a,b]$ is bounded below by

$$r(S_{i-1}^*) = - \frac{\ln||C_K|| + \ln||E_{K_1^*}|| + \ldots + \ln||E_{K_{i-1}^*}||}{K + S_{i-1}^*}, \tag{6.3}$$

where S_i^* is defined by $K_1^* + \ldots + K_i^*$ for $i = 1, 2, \ldots$.
The optimal value for the degree of the next elimination operator $E_{K_i^*}(\lambda_i,L)$, with the only knowledge of the preceding $K_1^*, \ldots, K_{i-1}^*$, is apparently the value which maximizes the expression $r(S_i^*)$. In this way the average rate of convergence with respect to $[a,b]$ remains as good as possible during the elimination process.

Theorem III. The values of K_i^*, $i = 1, 2, \ldots$, which maximize the expressions $r(S_i^*)$ are defined by the inequalities

$$r(S_i^*+1) < \ln \frac{||E_{K_i^*}(\lambda_i,\lambda)||}{||E_{K_i^*+1}(\lambda_i,\lambda)||} < r(S_i^*).$$

Proof. Suppose $\overline{K_i^*}$ is the value of K_i^* we are looking for, then an increase of $\overline{K_i^*}$ by one yields a smaller value for $r(S_i^*)$, thus

$$r(S_i^* + 1) < r(S_i^*). \tag{6.4}$$

Let us write $r(S_i^*) = \frac{P}{Q}$ and $r(S_i^* + 1) = \frac{P + p}{Q + q}$, where

$$p = \ln \frac{||E_{K_i^*}||}{||E_{K_i^*+1}||}, \quad q = 1.$$

From (6.4) we obtain

$$\frac{P}{Q} > \frac{p}{q},$$

which proves the right member of the inequalities of the theorem. In the same way the other inequality can be proved.

We investigate the values of K_i^* defined by the theorem for large values of K. In the same way as we derived (4.1) we find for $r(S_i^*)$ the expression

$$(6.5)\quad r(S_i^*) = \operatorname{arccosh} y_0 - \frac{S_i^* \operatorname{arccosh} y_0 + \ln 2 + \ln||E_{K_i^*}|| + \ldots + \ln||E_{K_i^*}||}{K + S_i^*}.$$

If $K \gg S_i^*$ only the first term remains i.e.

$$r(S_i^*) \approx \operatorname{arccosh} y_0.$$

From the foregoing theorem we obtain for K_i^* the relation

$$(6.6)\qquad ||E_{K_i^*+1}(\lambda_i,\lambda)|| \approx e^{-\operatorname{arccosh} y_0}||E_{K_i^*}(\lambda_i,\lambda)||.$$

We recall that $||E_{K_i^*}||$ is given by

$$(6.7)\qquad ||E_{K_i^*}|| = T^{-1}_{K_i^*}\left(\frac{b+a_i^*}{b-a_i^*}\right) = T^{-1}_{K_i^*}\left(\frac{b\cos\frac{\pi}{2K_i^*} + \lambda_i}{b - \lambda_i}\right).$$

7. Evaluation of the eigenvalues of L

In this section we give some methods to find the dominating eigenvalues of L during the iteration process.

In actual computation the iterates u_k are known. They are related to the errors v_k by the relation

$$u_k = u + v_k = u + P_k(L)v_0.$$

Forming the difference $u_{k+1} - u_k$ we can eliminate the unknown function u:

$$(7.1)\qquad u_{k+1} - u_k = (P_{k+1}(L) - P_k(L))v_0.$$

For $k = K \gg 1$ and $\lambda < a$ we have

$$(7.2)\qquad P_K(\lambda) = C_K(a,b,\lambda) = \frac{T_K(y_1(\lambda))}{T_K(y_0)} = \frac{\cosh\,(K \ln(y_1+\sqrt{y_1^2-1}))}{\cosh\,(K \ln(y_0+\sqrt{y_0^2-1}))}$$

$$\tilde{=} \left(\frac{y_1 + \sqrt{y_1^2-1}}{y_0 + \sqrt{y_0^2-1}} \right)^K ,$$

where $y_1(\lambda) = y_0 - 2\frac{\lambda}{b-a}$.
We compare $C_K(a,b,\lambda_1)$ with $C_K(a,b,\lambda_2)$. From (7.2) we obtain

$$(7.3)\qquad C_K(a,b,\lambda_1) \tilde{=} \left(\frac{y_1(\lambda_1) + \sqrt{y_1^2(\lambda_1)-1}}{y_1(\lambda_2) + \sqrt{y_1^2(\lambda_2)-1}} \right)^K C_K(a,b,\lambda_2).$$

If $\lambda_1 < \lambda_2$ the term between brackets is > 1, therefore, by choosing K large enough, $C_K(a,b,\lambda_1)$ is strongly dominating. In that case we have for k in the neighbourhood of K

$$u_{k+1} - u_k \tilde{=} (P_{k+1}(\lambda) - P_k(\lambda))e,$$

where λ is the dominating eigenvalue with eigenfunction e.
Forming the quotient

$$(7.4)\qquad q_{k+1} = \frac{||u_{k+1} - u_k||}{||u_k - u_{k-1}||} ,$$

where $||\ ||$ denotes an arbitrary norm, we obtain the following fundamental formula

$$(7.5)\qquad q_{k+1} \tilde{=} \frac{|P_{k+1}(\lambda) - P_k(\lambda)|}{|P_k(\lambda) - P_{k-1}(\lambda)|} .$$

In the <u>first order Richardson process</u> we have

$$P_{k+1}(\lambda) = (1 - \omega_k\lambda)P_k(\lambda) = (1 - \omega_k\lambda)(1 - \omega_{k-1}\lambda)P_{k-1}(\lambda).$$

Substituting this into (7.5) yields the estimate

$$\lambda \tilde{=} \frac{1}{\omega_{k-1}} - \frac{q_{k+1}}{\omega_k} . \tag{7.6}$$

In the second order Richardson process we have for every k, $P_k(\lambda) = C_k(a,b,\lambda)$. Substituting (7.2) into (7.5) results in the relation

$$q_{k+1} \tilde{=} \frac{y_1(\lambda) + \sqrt{y_1^2(\lambda) - 1}}{y_0 + \sqrt{y_0^2 - 1}} .$$

Solving this for λ gives the estimate

$$\lambda \tilde{=} \frac{1}{2}(b-a)\left(y_0 - \frac{1 + q_{k+1}^2(y_0 + \sqrt{y_0^2 - 1})^2}{2q_{k+1}(y_0 + \sqrt{y_0^2 - 1})}\right). \tag{7.7}$$

Using the formula

$$y_0 + \sqrt{y_0^2 - 1} = \frac{(\sqrt{a} + \sqrt{b})^2}{b-a} , \tag{7.8}$$

we obtain in terms of a and b

$$\lambda \tilde{=} \frac{1}{2}\left(b + a - \frac{(b-a)^2 + q_{k+1}^2(\sqrt{a} + \sqrt{b})^4}{2q_{k+1}(\sqrt{a} + \sqrt{b})^2}\right) \tag{7.7'}$$

$$= \frac{1}{4q_{k+1}}\left(-(\sqrt{a} + \sqrt{b})^2 q_{k+1}^2 + 2(b+a)q_{k+1} - (\sqrt{a} - \sqrt{b})^2\right).$$

There is another independent method to estimate dominating eigenvalues, which uses the relation $Lu = f$.

We define the quantities

$$r_{k+1} = \frac{||Lu_{k+1} - f||}{||u_{k+1} - u_k||} , \quad s_{k+1} = \frac{||Lu_k - f||}{||u_{k+1} - u_k||} . \tag{7.8}$$

As q_{k+1} the quantities r_{k+1} and s_{k+1} can be calculated during the iteration process. For sufficiently large k we have two more fundamental relations

$$(7.9) \qquad r_{k+1} \doteq \frac{|\lambda P_{k+1}(\lambda)|}{|P_{k+1}(\lambda) - P_k(\lambda)|}, \quad s_{k+1} \doteq \frac{|\lambda P_k(\lambda)|}{|P_{k+1}(\lambda) - P_k(\lambda)|},$$

where λ is again the dominating eigenvalue.

For the first order Richardson process we obtain from the first relation of (7.9)

$$(7.10) \qquad \lambda \doteq \frac{1}{\omega_k} - r_{k+1}.$$

The second relation of (7.9) results in an identity.

In the second order case we obtain the formulae

$$y_0 + \sqrt{y_0^2 - 1} = (1 + \lambda/r_{k+1})(y_1(\lambda) + \sqrt{y_1^2(\lambda) - 1})$$

$$y_1(\lambda) + \sqrt{y_1^2(\lambda) - 1} = (1 - \lambda/s_{k+1})(y_0 + \sqrt{y_0^2 - 1}).$$

If we substitute $y_1(\lambda) = y_0 - \frac{2}{b-a}\lambda$, we find the estimates

$$\lambda \doteq 2r_{k+1} \frac{b + a - \dfrac{b - a}{y_0 + \sqrt{y_0^2 - 1}} - 2r_{k+1}}{\dfrac{b - a}{y_0 + \sqrt{y_0^2 - 1}} + 4r_{k+1}}$$

$$\lambda \doteq 2s_{k+1} \frac{(b - a)\sqrt{y_0^2 - 1} - 2s_{k+1}}{(b - a)(\sqrt{y_0^2 - 1} + y_0) - 4s_{k+1}}.$$

In terms of a and b we have finally

(7.11a) $$\lambda = 4r_{k+1}\,\frac{\sqrt{ab} - r_{k+1}}{(\sqrt{a} - \sqrt{b})^2 + 4r_{k+1}}$$

and

(7.11b) $$\lambda = 4s_{k+1}\,\frac{\sqrt{ab} - s_{k+1}}{(\sqrt{a} + \sqrt{b})^2 - 4s_{k+1}} \quad .$$

We remark that the estimates (7.6), (7.7), (7.10) and (7.11) hold for positive as well as for negative eigenvalues λ.

8. Upper and lower bounds for the eigenvalues of L

The second order Richardson method may be used to compute upper and lower bounds for the first eigenvalue of L.

Suppose $\overline{\lambda}_1$ is an estimate for the eigenvalue λ_1 of L and v is a function in which the eigenfunction e_1 corresponding to λ_1 is strongly dominating. Such a situation can be obtained by the method described in the preceding sections. We try to eliminate the eigenfucntion e_1 from v by applying to v the operator $C_{K_1^*}(a_1^*,b,L)$, with

(8.1) $$a_1^* = \frac{2\overline{\lambda}_1 + b(\cos\frac{\pi}{2K_1^*} - 1)}{\cos\frac{\pi}{2K_1^*} + 1} \quad , \quad b = \sigma(L).$$

Using the second order process, the first zero of the polynomial $P_k(\lambda) = C_k(a_1^*,b,\lambda)$ is given by

$$\lambda_0(k) = \frac{1}{2}\,(b + a_1^*) - \frac{1}{2}\,(b - a_1^*)\cos\frac{\pi}{2k} \ ,$$

hence substituting (8.1) yields

$$(8.2)\qquad \lambda_0(k) = \frac{\bar{\lambda}_1(1 + \cos\frac{\pi}{2k}) + b(\cos\frac{\pi}{2K_1^*} - \cos\frac{\pi}{2k})}{1 + \cos\frac{\pi}{2K_1^*}}.$$

We consider the difference $\Delta\lambda = \lambda_0(k_1) - \lambda_0(k_2)$ i.e. the distance between the first zeros of the polynomial $C_k(a_1^*,b,\lambda)$ after k_1 and k_2 iterations. From (8.2) we obtain

$$(8.3)\qquad \Delta\lambda = 2(b - \bar{\lambda}_1)\,\frac{\sin[\frac{1}{4}\pi(\frac{1}{k_1} + \frac{1}{k_2})] - \sin[\frac{1}{4}\pi(\frac{1}{k_2} - \frac{1}{k_1})]}{1 + \cos\frac{\pi}{2K_1^*}}.$$

When the zero λ_0 "passes" the eigenvalue λ_1 an estimate of the dominating eigenvalue, such as given in section 7, will show a maximum, for at that moment λ_2 is dominating. Suppose that this maximum is assumed between the k_1^{th} and $k_2^{th} = (k_1 + 2)^{th}$ iteration, then we have

$$(8.4)\qquad \lambda_0(k_1 + 2) < \lambda_1 < \lambda_0(k_1),$$

where $\lambda_0(k_1 + 2)$ and $\lambda_0(k_1)$ follow from (8.2).
From (8.3) we can derive an estimate for $\Delta\lambda = \lambda_0(k_1) - \lambda_0(k_1 + 2)$ if $k_1 >> 1$. We obtain

$$(8.5)\qquad \Delta\lambda \doteq \frac{1}{8}(b - \bar{\lambda}_1)\,\pi^2(k_1)^{-3}.$$

If $\bar{\lambda}_1$ is areasonable estimate we have $K_1^* \doteq k_1$, hence

$$(8.5')\qquad \Delta\lambda \doteq \frac{1}{8}(b - \bar{\lambda}_1)\,\pi^2(K_1^*)^{-3}.$$

Therefore if we desire a certain accuracy $\Delta\lambda$, we must choose

$$K_1^* \doteq \frac{1}{2}\left(\frac{b - \lambda_1}{\Delta\lambda}\,\pi^2\right)^{1/3}.$$

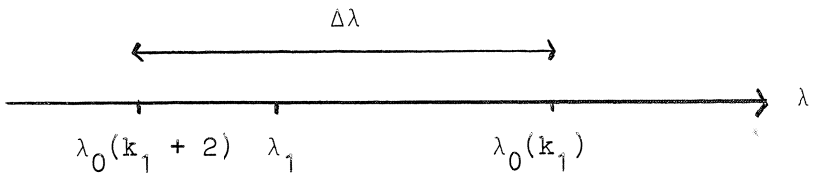

fig. 4

If one desires to determine also higher eigenvalues, one has to calculate λ_1 with great accuracy to accomplish a reasonable exact elimination of the first eigenfunction.

Appendix

In the preceding sections we have assumed that the matrix L was symmetric. Then the operator $P_K(L)$ was also symmetric, hence

$$\sigma(P_K(L)) = ||P_K(L)||,$$

where $||\ ||$ denotes the inner-product norm of the matrix $P_K(L)$. The Euclidean norm of the error v_k satisfies the inequality

$$||v_K|| \leq ||P_K(L)||\ ||v_0|| = \sigma(P_K(L))\ ||v_0||.$$

Therefore it was important to construct polynomials $P_K(L)$ with small spectral norms (compare section 1).

In our method we used two properties of L, namely that L had real eigenvalues and a complete set of eigenfunctions. Hence our method is applicable not only to symmetric matrices L, but also to non-symmetric matrices with the two properties mentioned above.

Let us now consider matrices L, which have positive eigenvalues (and possibly one or two negative eigenvalues), and which have not necessarily a complete set of eigenfunctions. We may again construct an operator $P_K(L)$ with a small spectral norm, however this doesn't guarantee that the inner-product norm (or another norm) is small. Let us apply the operator $P_K(L)$ n times to v_0 to get

$$v_{nK} = P_K^n(L)v_0,$$

so that

$$||v_{nK}|| \leq ||P_K^n(L)||\ ||v_0||.$$

It is well-known that $||P_K^n(L)||$ converges to zero for $n \to \infty$ if and only if $\sigma(P_K(L)) < 1$. Hence by repeating the operator $P_K(L)$ we may solve matrix equations $Lu = f$, where L is only required to have positive and some negative eigenvalues. (We remark that in this case it is not always possible to estimate the dominating eigenvalues during the first phase of our method.)

References

[1] Flanders, Donald A. and Shortley, George [1950], "Numerical determination of fundamental modes". J. Appl. Phys. 21, 1326-1332.

[2] Forsythe, G.E. and Wasow, W.R. [1960], "Finite Difference Methods for Partial Differential equations". John Wiley and Sons, Inc. New York, London 444 pp.

[3] Frank, Werner [1960], "Solution of linear systems by Richardson's method". J. Assoc. Comput. Mach. 7, 274-286.

[4] Frankel, S.P. [1950], "Convergence rates of iterative treatments of partial differential equations". Math. Tables Aids Comput. 4, 65-75.

[5] Richardson, L.F. [1910], "The approximate arithmetical solution by finite differences of physical problems involving differential equations, with application to the stress in a masonry dam". Philos. Trans. Roy. Soc. London, Ser. A 210, 307-357.

[6] Varga, Richard [1962], "Matrix iterative analysis". Prentice-Hall, Inc. Englewood Cliffs, New Yersey.

[7] Varga, Richard [1957], "A comparison of the successive overrelaxation method and semi-iterative methods using Chebyshev polynomials". J. Soc. Indust. Appl. Math. 5, 39-46.

[8] Young, David M. [1954], "On Richardson's method for solving linear systems with positive definite matrices". J. Math. Phys. 32, 243-255.