STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING TOEGEPASTE WISKUNDE

Report TW 107

On the acceleration of Richardson's method II

Numerical aspects

by

P.J. van der Houwen

May 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

In this second report on Richardson's method [1] some numerical aspects with which one is faced in actual computation are discussed. First we will consider the numerical stability of the first order version of Richardson's method. It is well-known that the order of the relaxation parameters must be chosen very carefully in order to get convergence. In fact, the distribution recommended by Young failed for the problems we tested on the computer. The factorization method proposed in section 3 proved to be numerically stable for very ill-conditioned matrix problems. Secondly we will prove that the second order process is strongly stable in the sense of O'Brien, Hyman and kaplan and finally we will discuss the effect of not estimating exactly the first eigenvalues.

In a third report [5] we will apply the theory to a number of matrix problems arising from the numerical solution of Laplace's equation.

---

1) In a preceding report (see reference list [4]) we discussed theoretical aspects of Richardson's method.

1. The elimination method

In this section we recall some of the main features of the theory developped in [4].

In [4] the solution of the matrix equation

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

was approximated by a sequence of functions u_k, which were calculated by the formula

$$u_{k+1} = (\alpha_k - \omega_k L)u_k + (1 - \alpha_k)u_{k-1} + \omega_k f. \tag{1.2}$$

The function u_k may be expressed in terms of the solution u and the initial error $v_0 = u_0 - u$, i.e.

$$u_k = u + P_k(L)v_0, \tag{1.3}$$

where $P_k(L)$ is a polynomial of degree k in L satisfying $P_k(0) = 1$. Let us assume that L has a complete set of eigenfunctions e_i, with eigenvalues λ_i of which most are positive and no more than a few are negative. We develop v_0 in the series

$$v_0 = \sum_i c_i e_i. \tag{1.4}$$

The essence of the method proposed in [4] was the reduction of the late eigenfunctions of L corresponding to the larger eigenvalues $\lambda_i \in [a,b]$, followed by the elimination of the remaining eigenfunctions. The reduction was achieved by means of the Chebyshef operator $C_K(a,b,L)$ adjusted to the interval $[a,b]$ and scaled to satisfy the condition $C_K(a,b,0) = 1$. The remaining eigenfunctions were eliminated by Chebyshef operators $C_K*(a^*,b,L)$, where a^* is chosen to satisfy $C_K*(a^*,b,\lambda_1) = 0$, λ_1 being the eigenvalue of the eigenfunction to be eliminated.

2. Numerical stability of iteration processes

In this section we discuss the numerical stability of iteration processes of the type

$$u_{k+1} = H_k u_k + g_k, \quad k = 0, 1, 2, \ldots \tag{2.1}$$

where g_k is a known vector depending on k and H_k is the iteration matrix also depending on k.

Such a process is called strongly stable [1] if for every k

$$||H_k|| < 1. \tag{2.2}$$

From this condition one may derive that the rounding errors occurring in actual computation cannot accumulate.

Suppose that with the application of the operator H_k we have the rounding error ε_k. Instead of the exact solution u_k we then get the numerical solution u_k^* satisfying the scheme

$$u_{k+1}^* = H_k u_k^* + g_k + \varepsilon_k. \tag{2.3}$$

After K iterations we have the numerical error

$$u_K^* - u_K = \prod_{k=0}^{K-1} H_k (u_0^* - u_0) + \sum_{k=0}^{K-2} \prod_{l=k+1}^{K-1} H_l \varepsilon_k + \varepsilon_{K-1}.$$

Evidently $u_0^* = u_0$, hence

$$u_K^* - u_K = \sum_{k=0}^{K-2} \prod_{l=k+1}^{K-1} H_l \varepsilon_k + \varepsilon_{K-1}. \tag{2.4}$$

We define

$$||H|| = \max_{0 \le k \le K-1} ||H_k|| \tag{2.5}$$

and

$$||\varepsilon|| = \max_{0 \le k \le K-1} ||\varepsilon_k||. \tag{2.6}$$

From (2.4) we obtain

$$(2.7) \qquad ||u_K^* - u_K|| \leq \sum_{k=0}^{K-1} ||H||^{K-k-1} ||\varepsilon|| = \frac{1 - ||H||^K}{1 - ||H||} ||\varepsilon||.$$

It may be concluded that a strongly stable iteration process has a final numerical error of the same order as the maximal round-off of the individual iterations. In those cases where $||\varepsilon||$ is small, the process is numerically stable.

Let us now consider iterative processes where $||H_k||$ also assumes values larger than 1.
We define the operators

$$(2.8) \qquad A_k = \prod_{l=k+1}^{K-1} H_l$$

for k = 0, 1, ..., K-1, where $A_{K-1} = 1$.
From (2.4) we see that $||A_k|| \cdot ||\varepsilon_k||$ is an upper bound for the growth of the round-off ε_k. Therefore we should require that

(a) The operator norms $||A_k||$ have to be small for k = 0,1,...,K-1.

A second condition is obtained by requiring that each round-off ε_k will be small, i.e. there will be no growth of numbers.
Let us define the operators

$$(2.9) \qquad B_k = \prod_{l=0}^{k} H_l$$

for k = 0, 1, ..., K-1.
To prevent a growth of numbers we require

(b) The operator norms $||B_k||$ have to be small for k = 0,1,...,K-1.

If we are able to arrange the operators H_k in such a way that the conditions (a) and (b) are satisfied, we may expect that the iteration process (2.1) is numerically stable.

3. Numerical stability of the first order Richardson process

As far as the author knows there is no detailed study of the stability against rounding errors of the first order Richardson process (cf. Forsythe and Wasow [2]p. 233). In this section the stability problem is investigated, which results in an arrangement of the relaxation parameters stable on the Electrologica X8 computer for very ill-conditioned matrix problems and for relative large values of K.

The first order Richardson process is of type (2.1) with

$$H_k = 1 - \omega_k L. \tag{3.1}$$

We shall assume that L has a complete set of eigenfunctions e_i with eigenvalues λ_i. Since the orders of the elimination operators $C_{K_i^*}(a_i^*,b,L)$ are small we will only be concerned with the effect of rounding errors associated with the application of the operator $C_K(a,b,L)$. It is sufficient to restrict the considerations to the space S(a,b), which is generated by the eigenfunctions e_i corresponding to the eigenvalues $\lambda_i \in [a,b]$, i.e. we will consider the numerical stability with respect to the space S(a,b) (compare [4] section 6). The eigenvalues of H_k are given by $1 - \omega_k\lambda_i$, so that the operators H_k are instable with respect to S(a,b) when

$$\omega_k \geq \frac{2}{b}\ . \tag{3.2}$$

In figure 1 the spectra of a stable and an instable operator H_k are illustrated.

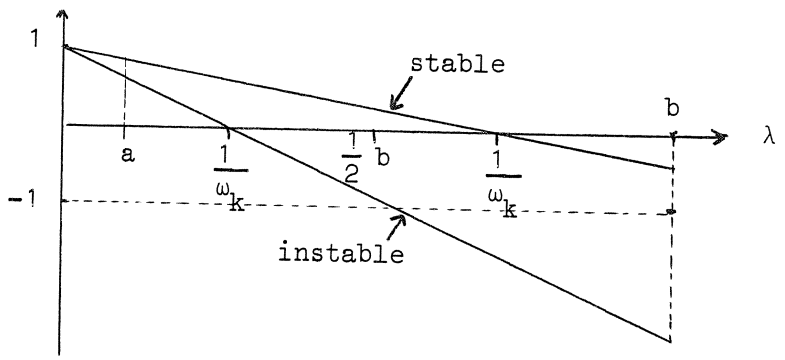

fig. 1

The norms $||H_k||$ may become very large, hence the relaxation parameters ω_k must be carefully distributed over the K iterations of the process.

Results of Young-Warlick

The experiments of Young and Walick ([6], [7]) proved that a descending or ascending order of the norms $||H_k||$ was unfavourable. Theoretically this may be concluded from the structure of the operators A_k and B_k defined in the preceding section.
Let us suppose that $||H_k||$ or ω_k is decreasing with k. The norms $||A_k||$ are small for every k. However, $||B_k||$ increases strongly in the first part of the iteration process (figure 2) giving rise to a strong growth of the error v_k in the middle of the process.

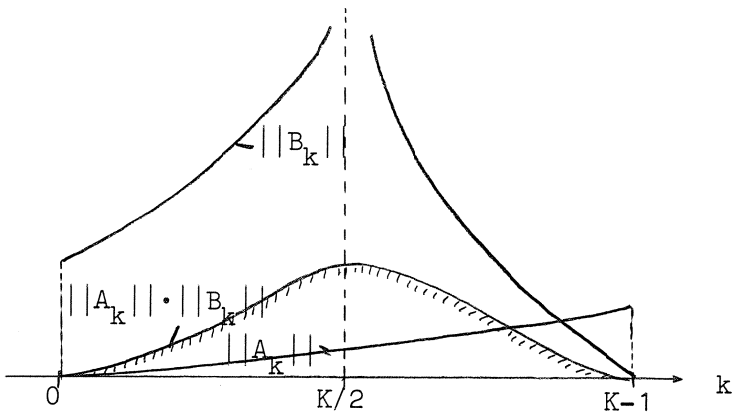

fig. 2

The descending order gives rise to a strong growth of numbers, but there is no growth of the individual round-offs ε_k. The area below the curve $||A_k||\cdot||B_k||$ (figure 2) gives an upper bound for the final numerical error $||u_K^* - u_K||$. If we invert the direction of k in figure 2 and if we interchange $||A_k||$ and $||B_k||$ we get a situation corresponding to an increasing order of the norms $||H_k||$.
The ascending order gives rise to a strong growth in the individual rounding errors ε_k, however, there is no growth of numbers.

In table I we have listed our results for the model problem I discussed in [5]. For the sake of completeness we have mentioned the parameters γ and stv respectively indicating the type of the difference formulae used and the number of the starting vector u_0.

TABLE I. Numerical stability of the first order Richardson process applied to the model problem I (see [5]).

	K	a	b	γ	stv	$R^*(K)$
Process without round-off	27	2	162	1.5		$\geq$ 0,20 (theor.)
Ascending order	27	2	162	1.5	4	- 0,05
Ascending order	27	2	162	1.5	5	0,03
Descending order	27	2	162	1.5	4	0,21
Descending order	27	2	162	1.5	5	0,23
Young-Warlick order	27	2	162	1.5	4	0,21
Second degree process	27	2	162	1.5	4	0,21

The number $R^*(K)$ is defined by

$$R^*(K) = -\frac{1}{K} \ln \frac{||Lu_K^* - f||}{||Lu_0 - f||}, \tag{3.3}$$

which will serve as an estimate for the convergence of the iterative process with respect to the space $S(a,b) = S(2,162)$. This follows from the following argument. If there are no rounding errors we may write

$$R^*(K) = -\frac{1}{K} \ln \frac{||Lu - Lv_K - f||}{||Lu - Lv_0 - f||} = -\frac{1}{K} \ln \frac{||C_K(a,b,L)Lv_0||}{||Lv_0||}. \tag{3.4}$$

From this we conclude that for $\varepsilon_k = 0$ the following inequality holds

$$R^*(K) \geq R(K), \tag{3.5}$$

where $R(K)$ is the average rate of convergence for K iterations. In the column denoted $R^*(K)$, the first number corresponds with the value of the lower bound $R(K)$ of the numerical values $R^*(K)$.

It follows from table I that the descending order is superior to the ascending order; the descending order yields the same rate of convergence as the strongly stable second order process (see the following section), while the ascending order is divergent. In connection with this, it is interesting to remark that Young [6] stated that the ascending order is superior to the descending one.

Young and Warlick tested an arrangement of the relaxation parameters, which starts with a middle value of ω_k, and then alternatively proceeds from the next lower to the next higher. This order was satisfactory for a problem characterized by (K,a,b) = (40,2,324) [6]. If larger values of K are required, Young recommends repeating the whole process. However, for ill-conditioned problems the rate of convergence is reduced considerable (see [4]). Therefore it is highly desirable to look for still better orders of the relaxation parameters.

Improvements of the Young-Warlick order

A first improvement of the Young-Warlick order is obtained if we invert the order of the relaxation parameters in the Young-Warlick method. This is suggested by the observation that a descending order of $||H_k||$ is superior to an ascending order. The Young-Warlick method may be interpreted as a sequence of quadratic operators

$$H_k H_{k'} = (1 - \omega_k L)(1 - \omega_{k'} L)$$

with

$$\omega_k + \omega_{k'} = \frac{1}{2}(a + b).$$

The norms of these quadratic operators are increasing, hence we expect that arranging them with decreasing norms will be a better strategy. Table II justifies this conclusion, although the method diverges for K = 81.

TABLE II. Numerical stability of the first order Richardson process applied to model problem I [5].

	K	a	b	γ	stv	$R^{*}(K)$
Process without round-off	81	2	162	1.5		$\geq$ 0,21 (theor.)
Ascending order	81	2	162	1.5	4	- 0,63
Descending order	81	2	162	1.5	4	- 0,46
Ascending-descending order	81	2	162	1.5	4	- 0,23
Young-Warlick method	81	2	162	1.5	4	- 0,05
Inverted Y-W method	81	2	162	1.5	4	- 0,00
Ascending-descending Y-W method	81	2	162	1.5	4	+ 0,19
Y-W method applied to quadr.operators	81	2	162	1.5	4	+ 0,21
Second degree process	81	2	162	1.5	4	+ 0,21

A second and better improvement is obtained if the quadratic operators are arranged with increasing-decreasing norms (see table II). This is clarified by the structure of the operators A_k and B_k defined in the preceding section. Let us first arrange the linear operators H_k with increasing-decreasing norms. In figure 3 the behaviour of $||A_k||$ and $||B_k||$ is illustrated. We see that $||A_k||$ and $||B_k||$ have not such an explosive behaviour as in figure 2. Moreover $||A_k||$ is small when $||B_k||$ is large and vice versa. This means that a large round-off does not grow strongly and a rounding error that does grow strongly is a small one.

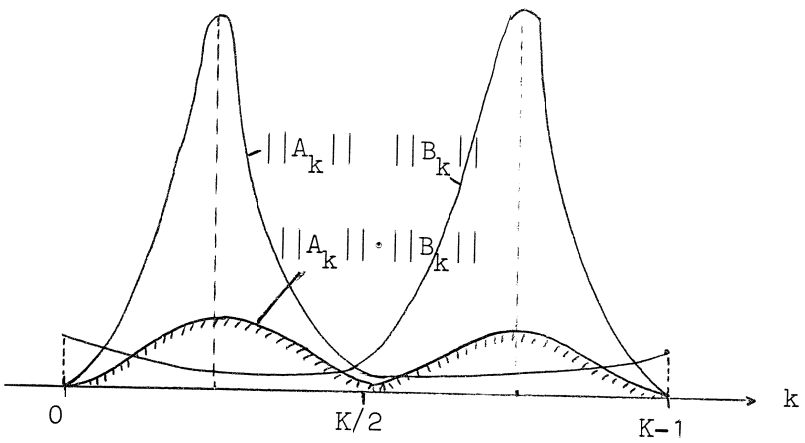

fig. 3

The same argument holds for an increasing-decreasing order of the quadratic operators of Young-Warlick.
The results in table II agree with these conclusions.

The Young-Warlick method and the improvements discussed above are obtained by replacing the operator H_k with $H_kH_{k'}$, and applying the methods found for the linear operators to these quadratic operators. This process may be continued. For instance one may apply the Young-Warlick method to the operators $H_kH_{k'}$, which results in a still better rate of convergence as is seen in table II.
However, this is an experimental manner of investigation. We shall attempt to give a more analytical description of the reduction of numerical errors.

Factorization into perturbed Chebyshef operators

We start with the following theorem

Theorem I

Let the relaxation parameters $\{\omega_k\}_{k=0}^{K-1}$ of the operator $C_K(a,b,L)$ be arranged in decreasing order. The relaxation parameters of the operators $C_{K_0}(a,b,L)$ are then given by the set $\{\omega_k \mid k = md_0 + \frac{1}{2}(d_0 - 1), m = 0, 1, \ldots, K_0-1, d_0 = K/K_0\}$, where d_0 is an odd divisor of K.

Proof

According to formula [4](2.3) the k-th zero of $C_K(a,b,\lambda)$ is given by

$$z_k = \frac{1}{2}(b + a) - \frac{1}{2}(b - a)\cos\left(\frac{2k+1}{2K}\pi\right). \tag{3.6}$$

Writing

$$\frac{2k+1}{2K}\pi = \frac{2d_0^{-1}\left(k - \frac{1}{2}(d_0 - 1)\right) + 1}{2K_0}\pi$$

we see that the $\left[(k - \frac{1}{2}(d_0 - 1))/d_0\right]^{th}$ zero of $C_{K_0}(a,b,\lambda)$ is a zero of $C_K(a,b,\lambda)$, which proves the theorem.

If d_0 is an even divisor of K it is easily verified from (3.6) that the zeros ζ_m of $C_{K_0}(a,b,\lambda)$ satisfy the inequality

$$z_{md_0+\frac{1}{2}(d_0-2)} < \zeta_m < z_{md_0+\frac{1}{2}d_0}. \tag{3.7}$$

To each zero z_m or ζ_m of $C_{K_0}(a,b,\lambda)$ we choose a neighbouring zero of $C_K(a,b,\lambda)$. In this manner we may factor $C_K(a,b,\lambda)$ into d_0 polynomials $P^{(i)}_{K_0}(\lambda)$ of degree K_0, which may be interpreted as perturbed Chebyshef polynomials. If d_0 is small and K_0 is large the polynomials $P^{(i)}_{K_0}(\lambda)$ will approximate the Chebyshef polynomial $C_{K_0}(a,b,\lambda)$.

If we take for the operators H_k of the preceding section the operators $P^{(i)}_{K_0}(L)$, we see that the iteration process is numerically stable if the operator $C_{K_0}(a,b,L)$ is numerically stable.

The perturbation of the operator $C_{K_0}(a,b,L)$ may be chosen in different ways but we shall consider only the following two cases:

(a) symmetric perturbations.

(b) one-sided perturbations.

(a) <u>symmetric perturbations</u>

For the sake of simplicity we will consider only perturbations of polynomials $C_{K_0}(a,b,\lambda)$ of even degree.

We define quadratic polynomials $Q_m(\lambda)$, satisfying $Q_m(0) = 1$ (see figure 4), with zeros z_m and z_{K_0-m-1}.

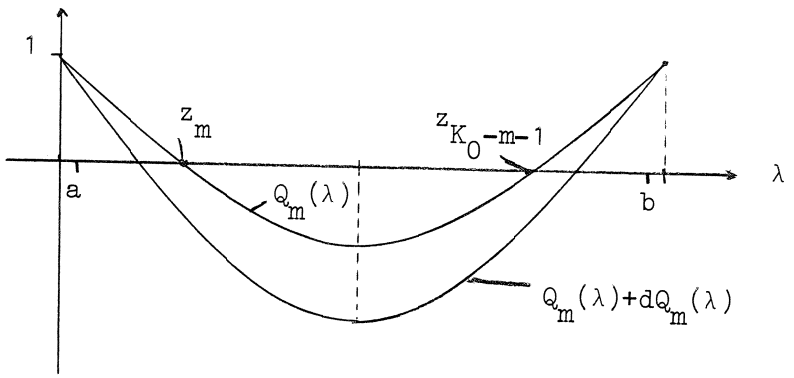

fig. 4

$C_{K_0}(a,b,\lambda)$ may be written as

$$C_{K_0}(a,b,\lambda) \equiv \prod_{m=0}^{\frac{1}{2}K_0-1} Q_m(\lambda). \tag{3.8}$$

We change z_m to z_m+dz_m and z_{K_0-m-1} to $z_{K_0-m-1}-dz_m$, where dz_m depends not only on m but also on i. The polynomial $Q_m(\lambda)$ is changed to $Q_m(\lambda) + dQ_m(\lambda)$ and $C_{K_0}(a,b,\lambda)$ to $P^{(i)}_{K_0}(\lambda)$.
From (3.8) we obtain

$$P^{(i)}_{K_0}(\lambda) = C_{K_0}(a,b,\lambda)\left[1 + \sum_{m=0}^{\frac{1}{2}K_0-1} \frac{dQ_m(\lambda)}{Q_m(\lambda)}\right]. \tag{3.9}$$

To keep the perturbation term small we choose dz_m with alternating sign. Then $dQ_m(\lambda)$ changes sign with m, so that the summation in (3.9) contains positive and negative terms in each point λ.
(It seems impossible to give a simpler expression for the perturbation term.)

(b) <u>one-sided perturbation</u>

If we choose dz_m definite for all m it is possible to construct an simple expression for $P^{(i)}_{K_0}(\lambda)$.
We assume the following form of $P^{(i)}_{K_0}(\lambda)$

$$P^{(i)}_{K_0}(\lambda) = \alpha_i\, C_{K_0}(a,b,\lambda) - \delta_i, \tag{3.10}$$

where α_i and δ_i are parameters to be determined.
From the condition $P^{(i)}_{K_0}(0) = 1$ we find

$$\alpha_i = 1 + \delta_i. \tag{3.11}$$

The parameters δ_i are obtained from the fact that the zeros of $C_K(a,b,\lambda)$ and the zeros of $\prod P^{(i)}_{K_0}(\lambda)$ are the same.
Let us write

$$C_K(a,b,\lambda) = \prod_{i=1}^{d_0} \left[(1 + \delta_i)C_{K_0}(a,b,\lambda) - \delta_i\right]. \tag{3.12}$$

If z_k is a zero of $C_K(a,b,\lambda)$ we have the relation

$$C_{K_0}(a,b,z_k) = \frac{\delta_i}{1+\delta_i},$$

hence

$$\cos\left[K_0 \arccos\frac{a+b-2z_k}{b-a}\right] = \frac{\delta_i}{1+\delta_i} T_{K_0}(y_0).$$

Solving for z_k yields

$$z_k = \frac{1}{2}(b+a) - \frac{1}{2}(b-a)\cos\left[\frac{\arccos\left(\frac{\delta_i}{1+\delta_i} T_{K_0}(y_0) + m\pi\right)}{K_0}\right], \tag{3.13}$$

where $m = 0, \pm 1, \ldots$.

Comparing (3.6) and (3.13) leads to

$$k = \frac{d_0}{\pi}\arccos\left(\frac{\delta_i}{1+\delta_i} T_{K_0}(y_0)\right) - \frac{1}{2} + md_0. \tag{3.14}$$

Putting k equal to i-1 gives the following expression for δ_i

$$\delta_i = \frac{\cos\frac{2i-1}{2d_0}\pi}{T_{K_0}(y_0) - \cos\frac{2i-1}{2d_0}\pi}. \tag{3.15}$$

We collect these results in the following theorem.

Theorem II

The polynomial $P^{(i)}_{K_0}(\lambda) \equiv (1+\delta_i)C_{K_0}(a,b,\lambda) - \delta_i$, where δ_i is given by (3.15) has the following zeros

$$\lambda = z_{i-1+md_0}, \quad m = 0, 1, \ldots, K_0-1. \tag{3.16}$$

The polynomial $P_{K_0}^{(i)}(\lambda)$ is completely determined by the values of K, K_0 and i. Ine may obtain these polynomials by selecting those zeros z_k of $C_K(a,b,\lambda)$ with k congruent modulo d_0 (,which means a one-sided perturbation of the zeros of $C_{K_0}(a,b,\lambda)$).
For large values of K_0 the parameters δ_i are small, so that the operators $P_{K_0}^{(i)}(L)$ may be interpreted as Chebyshef operators of degree K_0 with a small perturbation term. In particular if d_0 is an odd number, $P_{K_0}^{(\frac{d_0-1}{2})}(L)$ is exactly the operator $C_{K_0}(a,b,L)$ (cf. theorem I).

Theorem III

The sequence of operators $\{P_{K_0}^{(i)}(L)\}_{i=1}^{d_0}$ as defined by theorem II is strongly stable if

$$K_0 \geq \frac{1}{2}\sqrt{\frac{b}{a}} \cdot \ln(3 + 2\sqrt{2}). \tag{3.17}$$

Proof

The norm of $P_{K_0}^{(i)}(L)$ is given by

$$||P_{K_0}^{(i)}(L)|| = \frac{1 + |\cos(\frac{2i-1}{2d_0}\pi)|}{T_{K_0}(y_0) - \cos(\frac{2i-1}{2d_0}\pi)} \quad . \tag{3.18}$$

The stability condition is

$$||P_{K_0}^{(i)}(L)|| < 1. \tag{3.19}$$

In figure 5 the behaviour of $||P_{K_0}^{(i)}(L)||$ is illustrated as a function of i. For large values of $T_{K_0}(y_0)$ the curve is almost symmetrical with respect to $i = \frac{1}{2}(d_0 - 1)$. For small values of $T_{K_0}(y_0)$ there is a strong increase for $i \to 1$. In both cases $P_{K_0}^{(1)}(L)$ has the largest norm.

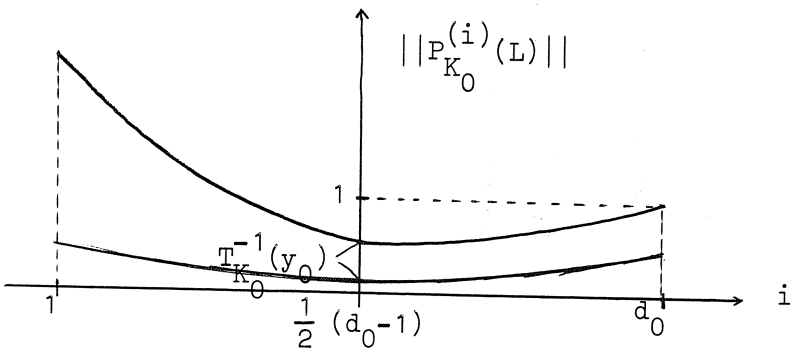

fig. 5

Condition (3.19) is satisfied for all i if

$$T_{K_0}(y_0) > 1 + 2\cos(\frac{K_0}{2K}\pi). \tag{3.20}$$

Using [4] (7.8) we get

$$K_0 > \frac{\ln\left[1 + 2\cos(\frac{K_0}{2K}\pi) + 2\sqrt{\cos(\frac{K_0}{2K}\pi)(\cos(\frac{K_0}{2K}\pi) - 1)}\right]}{\ln(\frac{\sqrt{b} + \sqrt{a}}{\sqrt{b} - \sqrt{a}})}. \tag{3.21}$$

This condition is certainly satisfied by K_0 if

$$K_0 \geq \frac{\ln(3 + 2\sqrt{2})}{\ln(\frac{\sqrt{b} + \sqrt{a}}{\sqrt{b} - \sqrt{a}})}. \tag{3.22}$$

From the Taylor expansion

$$\ln(\frac{\sqrt{b} + \sqrt{a}}{\sqrt{b} - \sqrt{a}}) = 2\frac{\sqrt{a}}{\sqrt{b} - \sqrt{a}} - \frac{2a}{(\sqrt{b} - \sqrt{a})^2} + \cdots \tag{3.23}$$

it may be deduced that for b >> a the following relation holds

$$\ln(\frac{\sqrt{b} + \sqrt{a}}{\sqrt{b} - \sqrt{a}}) \sim 2\sqrt{\frac{a}{b}}. \tag{3.24}$$

Formula (3.24) proves the theorem.

With the aid of theorem III one may draw certain conclusions about the numerical stability of the operators $P_{K_0}^{(i)}(L)$.
We apply formula (2.7) to the operators $\{P_{K_0}^{(i)}(L)\}_{i=1}^{d_0}$ where K_0 satisfies (3.17). We have

$$K = d_0 \ , \quad H = P_{K_0}^{(1)}(L)$$

and $||\varepsilon||$ is equal to the maximal numerical error associated with the operators $P_{K_0}^{(i)}(L)$. We find for the final numerical error the bound

$$\frac{1 - ||P_{K_0}^{(1)}(L)||^{d_0}}{1 - ||P_{K_0}^{(1)}(L)||} ||\varepsilon|| < d_0 ||\varepsilon||. \tag{3.25}$$

A scheme for arranging the relaxation parameters

We conclude this section with a description of a distribution of the relaxation parameters, which proved to be numerical stable on the X8 computer for very ill-conditioned matrix problems (see table III).
Let us suppose that the prime factorization of K is given by

$$K = d_1 \, d_2 \, \ldots \, d_n \tag{3.26}$$

and let us define the numbers

$$K_\nu = (d_1 \, d_2 \, \ldots \, d_\nu)^{-1} K. \tag{3.27}$$

We now choose $d_0 = d_1$ and $K_0 = K_1$ and we construct the polynomials $P_{K_1}^{(i)}(L)$ according to theorem II. Each of these polynomials may again be factored into d_2 polynomials $\overline{P}_{K_2}^{(j)}(L)$ by selecting those zeros z_m of $P_{K_1}^{(i)}(\lambda)$ with m congruent modulo d_2. This process is continued. We call this the factorization method.

Theorem IV

If the operators $P_{K_2}^{(l_0)}(L)$ and $\bar{P}_{K_2}^{(j_0)}(L)$ are obtained by the factorization method, then they are identical if $l_0 = j_0$.

Proof

According to theorem II the zeros of $P_{K_1}^{(i_0)}(\lambda)$ and $P_{K_2}^{(l_0)}(\lambda)$ coincide with the zeros z_k of $C_K(a,b,\lambda)$ for $k = i_0 - 1 + md_1$ $(m = 0, 1, \ldots, K_1-1)$ and for $k = l_0 - 1 + md_2$ $(m = 0, 1, \ldots, K_2-1)$ respectively. The zeros of $\bar{P}_{K_2}^{(j_0)}(\lambda)$ coincide with the zeros z_k for $k = i_0 - 1 + m_0 d_1 + md_1 d_2$ $(m = 0, 1, \ldots, K_2-1)$.

Therefore the operators $P_{K_2}^{(l_0)}(L)$ and $\bar{P}_{K_2}^{(j_0)}(L)$ are the same for $l_0 = i_0 + m_0 d_1 = j_0$.

Using the factorization method we obtain a product of perturbed Chebyshef operators of degree K_{n-1}, or degree K_{n-2}, or etc. There remains the problem of the order of the prime factors d_ν, the order of the d_ν operators $P_{K_\nu}^{(i)}(L)$ and the order of the relaxation parameters of $P_{K_{n-1}}^{(i)}(L)$. For the first problem it seems best to choose

(3.28) $$d_1 \leq d_2 \leq \ldots \leq d_n.$$

The last two order problems are of minor importance. One may choose for instance an ascending-descending order.

TABLE III. Numerical stability of Richardson's process applied to model problem I [5]

	K	a	b	γ	stv	$R^*(K)$
Factorization method	81	2	162	1.5	4	0,21
Process without round-off	81	.5	162	1.5		$\geq$ 0,103 (theor.)
Factorizat-on method	81	.5	162	1.5	4	0,105
Second degree process	81	.5	162	1.5	4	0,105
Process without round-off	81	.125	162	1.5		$\geq$ 0,047 (theor.)
Factorization method	81	.125	162	1.5	4	0,052
Second degree process	81	.125	162	1.5	4	0,052

4. Numerical stability of the second order Richardson process

We recall that the second order process is of the form (compare |4|, section 2)

$$u_{k+1} = (\alpha_k - \omega_k L)u_k + (1-\alpha_k)u_{k-1} + \omega_k f, \tag{4.1}$$

where

$$\alpha_k = 2y_0 \frac{T_k(y_0)}{T_{k+1}(y_0)}, \quad \omega_k = \frac{4}{b-a} \frac{T_k(y_0)}{T_{k+1}(y_0)}. \tag{4.2}$$

As in the preceding section we restrict our stability considerations to the space S(a,b) generated by the eigenfunction e_i of L with eigenvalues λ_i in the interval $[a,b]$.

We introduce the vectors $\vec{w}_k$ and $\vec{g}$ with components u_k, u_{k-1} and f, 0 respectively. Formula (4.1) may be written as a first order iteration process of type (2.1), i.e.

$$\vec{w}_{k+1} = H_k \vec{w}_k + \vec{w}_k \vec{g}, \tag{4.3}$$

where

$$H_k = \begin{pmatrix} \alpha_k - \omega_k L & 1 - \alpha_k \\ 1 & 0 \end{pmatrix}. \tag{4.4}$$

Theorem V

The second order Richardson process is strongly stable in the space S(a,b).

Proof

We require that for each $\lambda \in [a,b]$ the matrix H_k, with L replaced by λ, has eigenvalues within the unit circle. These eigenvalues satisfy the equation

$$z^2 - Sz + P = 0, \tag{4.5}$$

where

$$S = \alpha_k - \omega_k\lambda \ , \quad P = \alpha_k - 1. \tag{4.6}$$

The roots of (4.5) are within the unit circle when

$$P < 1, \ 1 + S + P > 0 \text{ and } 1 - S + P > 0. \tag{4.7}$$

These conditions lead to

$$\alpha_k < 2, \ \omega_k\lambda > 0, \ 2\alpha_k - \omega_k\lambda > 0. \tag{4.7'}$$

In figure 6 that domain of points (ω,α), which guarantees stability is indicated by the shaded region.

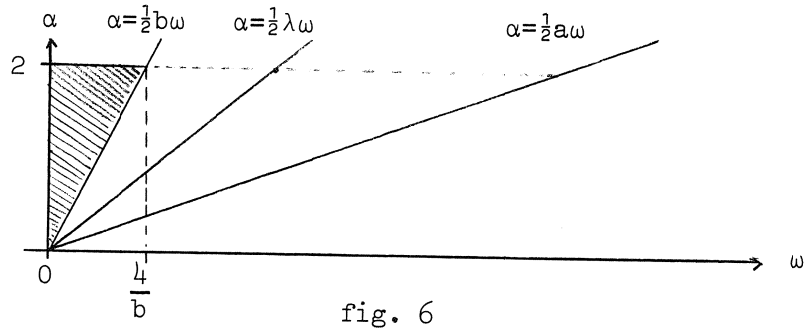

fig. 6

With the aid of the recurrence relation for the Chebyshef polynomials $T_k(y)$ one may easily verify that the values given by (4.2) satisfy (4.7').

According to this theorem there will be no accumulation of rounding errors in actual application. Therefore the final rounding error will be of the same order as the order of the round-off associated with one iteration. For numerical results we refer to the experiments of Frank [3] and table III.

5. The accuracy of elimination operators in numerical applications

In actual computation the eigenvalues of the eigenfunctions to be eliminated are only known as an approximate value, so that these eigenfunctions cannot be eliminated exactly. In this section we discuss the accuracy of the elimination process.

Let $\bar{\lambda}_1$ be an estimate for the eigenvalue λ_1 of L. We eliminate the corresponding eigenfunctions e_1 by means of the operator $C_{K_1^*}(a_1^*, b, L)$.

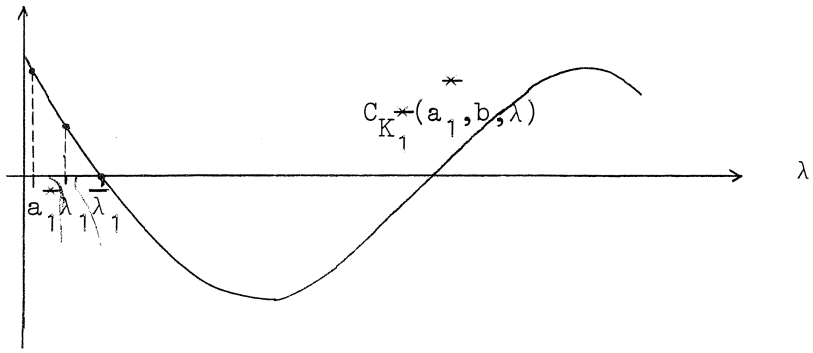

fig. 7

In figure 7 the curve $C_{K_1^*}(a_1^*, b, \lambda)$ is illustrated. We approximate the value of $C_{K_1^*}(a_1^*, b, \lambda)$ in $\lambda = \lambda_1$ by the formula

$$C_{K_1^*}(a_1^*, b, \lambda) = (\lambda_1 - \bar{\lambda}_1)\left[\frac{d}{d\lambda} C_{K_1^*}(a_1^*, b, \lambda)\right]_{\lambda=\bar{\lambda}_1} . \tag{5.1}$$

From the definition of the Chebyshef polynomials we find

$$C_{K_1^*}(a_1^*, b, \lambda_1) = \frac{(\bar{\lambda}_1 - \lambda_1)K_1^*}{T_{K_1^*}\left(\frac{b+a_1^*}{b-a_1^*}\right)\sqrt{(b-\bar{\lambda}_1)(\bar{\lambda}_1 - a_1^*)}} . \tag{5.1'}$$

Now a_1^* is given by (cf. [4] formula (4.4))

$$a_1^* = \frac{2\bar{\lambda}_1 + b\left(\cos\frac{\pi}{2K_1^*} - 1\right)}{\cos\frac{\pi}{2K_1^*} - 1} .$$

Substituting this into (5.1') leads to

$$C_{K_1^*}(a_1^*,b,\lambda_1) \simeq \frac{(\bar{\lambda}_1-\lambda_1)K_1^*}{(b-\bar{\lambda}_1)\operatorname{tg}\frac{\pi}{4K_1^*}\,T_{K_1^*}(y_0^*)}\,, \tag{5.2}$$

where y_0^* is given by

$$y_0^* = \frac{b\cos\frac{\pi}{2K_1^*} + \bar{\lambda}_1}{b - \bar{\lambda}_1}\,. \tag{5.3}$$

In practice we may write

$$C_{K_1^*}(a_1^*,b,\lambda_1) \simeq \frac{4K_1^{*2}}{b\pi T_{K_1^*}(y_0^*)}\,(\bar{\lambda}_1 - \lambda_1). \tag{5.2'}$$

For large values of K_1^* the value of $C_{K_1^*}(a_1^*,b,\lambda_1)$ will be small. However, we wish to use relative small values for K_1^*. For the elimination operators discussed in [4] section 5 we have, for example:

$$K_1^* \simeq \frac{1}{4}\,\pi\sqrt{\frac{b}{\bar{\lambda}_1}}\,,$$

yielding

$$C_{K_1^*}(a_1^*,b,\lambda_1) \sim \frac{1}{4}\,\pi\,\frac{\bar{\lambda}_1 - \lambda_1}{\bar{\lambda}_1}\,, \tag{5.4}$$

which is approximately of the same order as the relative error of the estimation $\bar{\lambda}_1$.

The estimation $\bar{\lambda}_1$ is accurate enough when

$$C_{K_1^*}(a_1^*,b,\lambda_1) \leq ||C_K(a,b,\lambda)||. \tag{5.5}$$

Using (5.2') we obtain for large values of K and b >> a

$$\bar{\lambda}_1 - \lambda_1 \leq \frac{1}{2}\,\pi\,\frac{b}{(K_1^*)^2}\,T_{K_1^*}(y_0^*)\exp(-2K\sqrt{\frac{a}{b}}). \tag{5.6}$$

In table IV the values of $T_K^*(y_0^*)$ for a number of values for y_0^* are listed. In view of later applications to the Dirichlet problem for the Poisson equation

$$\Delta U + F = 0, \tag{5.7}$$

we have chosen the lowest eigenvalues of the operator $-\Delta$, given by $\lambda_e = n^2 + m^2$, where n and m are integers, to be eliminated. Further, the matrix approximation L to $-\Delta$ is assumed to have the spectral norm

$$\sigma(L) = 162. \tag{5.8}$$

TABLE IV. Inverse norms of the elimination operators

	$\lambda_e = 2$		$\lambda_e = 5$		$\lambda_e = 8$		$\lambda_e = 10$	
K^*	T_K^*	$\ln T_K^*$	T_K^*	$\ln T_K^*$	T_K^*	$\ln T_K^*$	T_K^*	$\ln T_K^*$
1	0.013	-4.34	0.032	-3.44	0.052	-2.96	0.066	-2.72
2	0.061	-2.80	0.16	-1.83	0.27	-1.31	0.34	-1.08
3	0.15	-1.90	0.39	-0.94	0.68	-0.39	0.90	-0.11
4	0.27	-1.31	0.77	-0.26	1.4	0.34	1.9	0.64
5	0.44	-0.82	1.3	0.26	2.6	0.95	3.7	1.29
6	0.68	-0.39	2.2	0.79	4.6	1.52	6.8	1.92
7	0.98	-0.02	3.5	1.25	7.8	2.05	12	2.48
8	1.4	0.34	5.3	1.67	13	2.56	21	3.04

If the estimation λ_1 does not satisfy (5.6), that is the approximation if λ_i is too rough, we may apply the operator $C_{K_1}^*(a_1^*, b, L)$ again.
It may be remarked that the norm of the operator

$$C_{K_1}^*(a_1^*, b, L) C_{K_2}^*(a_2^*, b, L),$$

where $K_1^* + K_2^*$ is a fixed even number, is minimal for $K_1^* = K_2^*$. This follows from the fact that $\ln T_K^*(y_0^*)$ is a concave function of K^*, as seen from table IV.

References

[1] O'Brien, G., Hyman, M.A. and Kaplan, S. [1950], "A study of the numerical solution of partial differential equations". J. of Math. and Phys., V.XXIX.

[2] Forsythe, G.E. and Wasow, W.R. [1960], "Finite difference methods for partial differential equations". John Wiley and Sons, Inc. New York, London.

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

[4] Houwen, P.J. van der [1967], "On the acceleration of Richardson's method I" (Theoretical part). Report TW 104, Math. Centrum, Amsterdam.

[5] Houwen, P.J. van der [1967], "On the acceleration of Richardson's method II" (Applications). Report TW 108, Math. Centrum, Amsterdam.

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

[7] Young, David M. and Warlick, H. [1953], "On the use of Richardson's method for the numerical solution of Laplace's equation on the ORDVAC". Memorandum Report No 707, Ballistics Research Laboratories, Aberdeen Proving Ground, Md., 32 pp.