On the Convergence of the Rayleigh Quotient Iteration for the Computation of the Characteristic Roots and Vectors. I

A. M. Ostrowski *

1. Let $A = (a_{\mu\nu})$ be an $(n \times n)$ real symmetric matrix. Then for a vector $\xi = (x_1, \ldots, x_n)$ we put

(1) $$Q_A(\xi) = \xi A \xi' = \sum_{\mu\nu} a_{\mu\nu} x_\mu x_\nu .$$

The quotient

(2) $$\frac{Q_A(\xi)}{|\xi|^2}$$

is called the *Rayleigh quotient* corresponding to ξ. If ξ is a characteristic vector belonging to a characteristic root λ, then the corresponding Rayleigh quotient is λ. Therefore the following procedure has been devised for obtaining a sequence of numbers $\lambda_\varkappa$ $(\varkappa = 0, 1, \ldots)$ converging to a characteristic root:

For any $\lambda_\varkappa$ of the sequence $(\varkappa = 0, 1, \ldots)$ find an approximate solution $\xi_\varkappa$ of the homogeneous system

(3) $$A \xi'_\varkappa = \lambda_\varkappa \xi'_\varkappa$$

and put

(4) $$\lambda_{\varkappa+1} = \frac{Q_A(\xi_\varkappa)}{|\xi_\varkappa|^2} \qquad (\varkappa = 0, 1, \ldots).$$

My attention was drawn to this method by John Todd, who used it in his lectures as long ago as 1945. It appears to converge fairly well in numerical practice. In what follows I give some theoretical results on the convergence of this method.

2. The crucial point in the discussion of the above method is of course a suitable rule for the computation of the "approximate solution" $\xi_\varkappa$ of (3). The rule I shall use in the first part of this discussion consists in taking an arbitrary vector $\eta \neq 0$ and in putting

(5) $$\xi'_\varkappa = (A - \lambda_\varkappa E)^{-1} \eta' .$$

The theoretical arguments in support of this rule are given in another paper**.

---

* In writing this paper I had very valuable discussions with Mr. Chr. Blatter.

** A. Ostrowski, „Über näherungsweise Auflösung von Systemen homogener linearer Gleichungen", Journal of Applied Mathematics and Physics (ZAMP), Basle, Vol. 8 (1957), pp. 280—285.

Since the formulae (4), (5) are *invariant*, for our discussion we can introduce *normal coordinates* from the beginning and therefore without loss of generality put

$$Q_A(\xi) = \sum_{\nu=1}^{n} \mu_\nu x_\nu^2, \tag{6}$$

where

$$\mu_1 \leq \mu_2 \leq \cdots \leq \mu_n \tag{7}$$

are the characteristic roots of A, ordered increasingly. Then, if

$$\eta = (y_1, \ldots, y_n), \tag{8}$$

we have from (5)

$$\xi_\varkappa = (x_1^{(\varkappa)}, \ldots, x_n^{(\varkappa)}) = \left(\frac{y_1}{\mu_1 - \lambda_\varkappa}, \ldots, \frac{y_n}{\mu_n - \lambda_\varkappa}\right),$$

$$Q_A(\xi_\varkappa) = \sum_{\nu=1}^{n} \frac{\mu_\nu y_\nu^2}{(\mu_\nu - \lambda_\varkappa)^2}, \quad |\xi_\varkappa|^2 = \sum_{\nu=1}^{n} \frac{y_\nu^2}{(\mu_\nu - \lambda_\varkappa)^2},$$

and finally

$$\lambda_{\varkappa+1} = \frac{\sum_{\nu=1}^{n} \frac{\mu_\nu y_\nu^2}{(\mu_\nu - \lambda_\varkappa)^2}}{\sum_{\nu=1}^{n} \frac{y_\nu^2}{(\mu_\nu - \lambda_\varkappa)^2}} \qquad (\varkappa = 0, 1, \ldots). \tag{9}$$

3. The expression on the right side in (9), if all products $\mu_\nu y_\nu^2$ in the numerator are replaced by $\mu_1 y_\nu^2$ or $\mu_n y_\nu^2$, reduces to μ_1 and μ_n, respectively. We see that in any case

$$\mu_1 \leq \lambda_\varkappa \leq \mu_n \qquad (\varkappa = 1, \ldots). \tag{10}$$

In the expression on the right side of (9) a μ_ν drops out if the corresponding y_ν vanishes. Denote the remaining *distinct* μ_ν in increasing order by

$$\sigma_1 < \sigma_2 < \cdots < \sigma_m. \tag{11}$$

Then the formula (9) becomes

$$\lambda_{\varkappa+1} = \frac{\sum_{\mu=1}^{m} \frac{\sigma_\mu p_\mu}{(\sigma_\mu - \lambda_\varkappa)^2}}{\sum_{\mu=1}^{m} \frac{p_\mu}{(\sigma_\mu - \lambda_\varkappa)^2}} \qquad (\varkappa = 0, 1, \ldots), \tag{12}$$

where the coefficients p_μ are all *positive*.

4. Denote one of the σ_μ by σ and the corresponding p_μ by p. By subtracting σ from both sides of (12) we obtain

$$\lambda_{\varkappa+1} - \sigma = \frac{\sum_{\mu=1}^{m} p_\mu \frac{\sigma_\mu - \sigma}{(\sigma_\mu - \lambda_\varkappa)^2}}{\sum_{\mu=1}^{m} \frac{p_\mu}{(\sigma_\mu - \lambda_\varkappa)^2}} = (\sigma - \lambda_\varkappa)^2 \frac{\sum_{\mu=1}^{m}{}' p_\mu \frac{\sigma_\mu - \sigma}{(\sigma_\mu - \lambda_\varkappa)^2}}{p + \sum_{\mu=1}^{m}{}' p_\mu \left(\frac{\sigma - \lambda_\varkappa}{\sigma_\mu - \lambda_\varkappa}\right)^2}, \tag{13}$$

where in the last sums the terms with the index μ for which $\sigma_\mu = \sigma$ are to be omitted. From this formula we see that if one $\lambda_\varkappa$ gets sufficiently near to σ, the whole sequence $\lambda_\varkappa$ tends to σ. Then dividing the first and the last term of (13) by $(\lambda_\varkappa - \sigma)^2$, we obtain

$$\frac{\lambda_{\varkappa+1} - \sigma}{(\lambda_\varkappa - \sigma)^2} \to \frac{1}{p} \sum_{\mu=1}^{m}{}' \frac{p_\mu}{\sigma_\mu - \sigma} \qquad (\lambda_\varkappa \to \sigma). \tag{14}$$

We see that in this case the convergence is at least *quadratic*. The convergence could even be faster than that, if the limit in (14) were 0. However, for $\sigma = \sigma_1$ or $\sigma = \sigma_m$ the convergence is exactly quadratic*.

5. In order to characterize the convergence neighbourhood of σ put

$$d = \underset{\sigma_\mu \neq \sigma}{\mathrm{Min}} |\sigma_\mu - \sigma|, \qquad P = \sum_{\mu=1}^{m} p_\mu. \tag{15}$$

If we then assume

$$|\lambda_\varkappa - \sigma| \leq \frac{d}{2}, \tag{16}$$

for $\sigma_\mu \neq \sigma$ it follows that

$$\frac{|\sigma_\mu - \sigma|}{(\sigma_\mu - \lambda_\varkappa)^2} = \frac{|\sigma_\mu - \sigma|}{(\sigma_\mu - \sigma + \sigma - \lambda_\varkappa)^2} \leq \frac{|\sigma_\mu - \sigma|}{\left|\frac{\sigma_\mu - \sigma}{2}\right|^2} = \frac{4}{|\sigma_\mu - \sigma|} \leq \frac{4}{d}.$$

Introducing this in the numerator of the last term in (13) and replacing the denominator by p, we get

$$\frac{|\lambda_{\varkappa+1} - \sigma|}{(\lambda_\varkappa - \sigma)^2} \leq \frac{4}{d} \frac{\sum_{p_\mu \neq p} p_\mu}{p} = \frac{4}{d} \frac{P - p}{p},$$

$$\frac{|\lambda_{\varkappa+1} - \sigma|}{(\lambda_\varkappa - \sigma)^2} \leq \frac{4}{d} \left(\frac{P}{p} - 1\right), \tag{17}$$

$$\left|\frac{\lambda_{\varkappa+1} - \sigma}{\lambda_\varkappa - \sigma}\right| \leq \frac{4}{d} \frac{P - p}{p} |\lambda_\varkappa - \sigma|. \tag{18}$$

Therefore it follows that if for one index $\varkappa$ we have

$$|\lambda_\varkappa - \sigma| < \frac{d}{2} \mathrm{Min}\left(\frac{p}{2(P - p)}, 1\right), \tag{19}$$

then the distances $|\lambda_\varkappa - \sigma|$ from that index on are strictly diminishing and therefore converge to 0. We see that a convergence neighborhood of σ is given by (19).

Of course in this way we obtain only sequences $\lambda_\varkappa$ converging to those characteristic values μ_ν of A which remain among the σ_μ. If the constant vector η is orthogonal to all characteristic vectors corresponding to a characteristic root μ_ν,

---

* This result agrees with the note of WALTER KOHN, "A Variational Iteration Method for Solving Secular Equations", Journal of Chemical Physics, **17**, 670 (1949). In this note W. KOHN discusses the application of the Rayleigh quotient method in taking one of the coordinate unit vectors for η. He says then (in our notation): "A more careful analysis shows that

$$\delta\lambda_{\varkappa+1} = K_\varkappa (\delta\lambda_\varkappa)^2,$$

where $K_\varkappa$ is in general of the order 1."

then this characteristic root drops out, since all corresponding y_ν^2 vanish. However, we obtain for each choice of η at least one characteristic root, if the starting λ_0 is chosen near enough to such a root. In particular cases it may be necessary to try out several choices of η, for instance to try each of the coordinate unit vectors.

6. In order to discuss the global convergence situation we have to consider all fixed points of the iteration $\lambda_{\varkappa+1} = \varphi(\lambda_\varkappa)$, where

$$\varphi(\lambda) = \frac{\sum_{\mu=1}^{m} \frac{\sigma_\mu p_\mu}{(\sigma_\mu - \lambda)^2}}{\sum_{\mu=1}^{m} \frac{p_\mu}{(\sigma_\mu - \lambda)^2}}. \tag{20}$$

The corresponding algebraic equation for the fixed points, $\lambda = \varphi(\lambda)$, becomes an equation of degree $\leq 2m-1$. We know already m different roots of this equation, given by the σ_μ, and we have seen that all these fixed points are points of attraction. By a theorem which we proved in another communication*, between two consecutive fixed points of attraction there is always at least one fixed point of repulsion. We see that besides the m fixed points (11) the iteration by $\varphi(\lambda)$ has at least $m-1$ further different fixed points. Therefore the iteration by $\varphi(\lambda)$ has *exactly* the $2m-1$ fixed points

$$\sigma_1 < \vartheta_1 < \sigma_2 < \vartheta_2 < \cdots < \vartheta_{m-1} < \sigma_m, \tag{21}$$

where the ϑ_μ are points of repulsion.

In order to obtain an algebraic equation of degree $m-1$ satisfied by the ϑ_μ, subtract λ from the expression (20) and multiply by the denominator. Then we obtain

$$\sum_{\mu=1}^{m} \left[\frac{\sigma_\mu p_\mu}{(\lambda - \sigma_\mu)^2} - \frac{\lambda p_\mu}{(\lambda - \sigma_\mu)^2}\right] = -\sum_{\mu=1}^{m} \frac{p_\mu}{\lambda - \sigma_\mu}. \tag{22}$$

Therefore the polynomial equation for all fixed points is given by

$$\prod_{\mu=1}^{m} (\lambda - \sigma_\mu)^2 \sum_{\mu=1}^{m} \frac{p_\mu}{\lambda - \sigma_\mu} = 0, \tag{23}$$

while the polynomial equation satisfied by the ϑ_μ is obtained in the form

$$\prod_{\mu=1}^{m} (\lambda - \sigma_\mu) \sum_{\mu=1}^{m} \frac{p_\mu}{\lambda - \sigma_\mu} = 0. \tag{24}$$

The equation for the repulsive fixed points ϑ_μ, given above in normal coordinates, can be written in the invariant form

$$Q_{(\lambda E - A)^{-1}}(\eta) = 0. \tag{25}$$

7. Since the ϑ_μ are points of repulsion, we have $|\varphi'(\vartheta_\mu)| \geq 1$. It can easily be shown that at each of the points ϑ_μ we have

$$\varphi'(\vartheta_\mu) \geq 1. \tag{26}$$

* A. Ostrowski, Mathematische Miszellen XXV, „Über das Verhalten von Iterationsfolgen im Divergenzfall", Jahresber. d. DMV, Bd. 59 (1956), pp. 69–79.

Indeed, if we take $\lambda > \sigma_\mu$, in a sufficiently small neighborhood of σ_μ we have

$$\varphi(\lambda) < \lambda.$$

If λ increases, this inequality remains valid until we get the first point for which $\varphi(\lambda) = \lambda$. This is ϑ_μ. We have therefore for sufficiently small positive ε: $\varphi(\vartheta_\mu - \varepsilon) < \vartheta_\mu - \varepsilon$. On the right side, replace ϑ_μ by $\varphi(\vartheta_\mu)$, subtract $\varphi(\vartheta_\mu)$ on both sides and divide by $-\varepsilon$. We obtain

$$\frac{\varphi(\vartheta_\mu - \varepsilon) - \varphi(\vartheta_\mu)}{-\varepsilon} > 1,$$

and from this, since $\varepsilon \downarrow 0$, (26) follows immediately.

8. The further discussion of the global convergence problem appears to present considerable difficulties if treated by the method of conjugate couples of points (*cf.* the paper cited in § 6). Indeed the determination of such couples depends on the solution of the equation

$$\lambda = \varphi(\varphi(\lambda)),$$

which reduces to an algebraic equation of degree $4m-1$; we may expect as many as $m-1$ couples of conjugate points.

Only when $m=2$ is the problem solved immediately. Indeed it follows then from the analysis given in Section 7 that as soon as a $\lambda_\varkappa$ lies in one of the open intervals (σ_1, ϑ_1), (ϑ_1, σ_2), the sequence $\lambda_\varkappa$ converges to σ_1 in the first case and to σ_2 in the second. On the other hand it follows from (10) that λ_1 lies in $<\sigma_1, \sigma_2>$. Therefore the decision in this case is possible after the first iterative step, as soon as we have determined ϑ_1. But here the equation (25) gives immediately for ϑ_1 the expression

$$\vartheta_1 = \frac{a_{11} y_2^2 - 2a_{12} y_1 y_2 + a_{22} y_1^2}{y_1^2 + y_2^2}. \tag{27}$$

9. In the foregoing discussion, in order to obtain $\xi_\varkappa$ we used an arbitrary *fixed* vector η in (5). On the other hand, in the theory of direct iteration a variant due to H. WIELANDT* and called *broken iteration* is often used and consists in forming recursively the vectors $\xi_\varkappa$ given by

$$\xi_\varkappa = A^{-1} \xi_{\varkappa-1},$$

starting with an arbitrary vector $\xi_{-1} = \eta$. In combining this idea of broken iteration with our rule (4) we obtain the following modification of our rule:

For any $\lambda_\varkappa$ define a vector $\xi_\varkappa$ by

$$\xi_\varkappa = (A - \lambda_\varkappa E)^{-1} \xi_{\varkappa-1} \qquad (\varkappa = 0, 1, \ldots), \tag{28}$$

where ξ_{-1} is an arbitrary vector $\eta \neq 0$, and then put

$$\lambda_{\varkappa+1} = \frac{Q_A(\xi_\varkappa)}{|\xi_\varkappa|^2} \qquad (\varkappa = 0, 1, \ldots). \tag{29}$$

It will turn out in this case that the convergence is indeed considerably hastened, becoming *cubic* instead of quadratic.

---

* H. WIELANDT, „Beiträge zur mathematischen Behandlung komplexer Eigenwertprobleme, V: Bestimmung höherer Eigenwerte durch gebrochene Iteration", Bericht B 44/J/37 der aerodynamischen Versuchsanstalt Göttingen, 1944.

10. In order to discuss this procedure we assume again without loss of generality that the coordinates are normal and that (6), (7), (8) hold. If we again put $\xi_\varkappa = (x_\nu^{(\varkappa)})$, it follows from (28) that $x_\nu^{(\varkappa)} = \frac{x_\nu^{(\varkappa-1)}}{\mu_\nu - \lambda_\varkappa}$, and therefore

$$x_\nu^{(k)} = \frac{y_\nu}{\prod_{\varkappa=0}^{k} (\mu_\nu - \lambda_\nu)} \qquad (k = 0, 1, \ldots) \tag{30}$$

If we put

$$N_{\nu,k} = \prod_{\varkappa=0}^{k} (\mu_\nu - \lambda_\varkappa) \qquad (\nu = 1, \ldots, n;\ k = 0, 1, \ldots),$$

we obtain

$$x_\nu^{(k)} = \frac{y_\nu}{N_{\nu,k}}, \qquad Q_A(\xi_k) = \sum_{\nu=1}^{n} \frac{\mu_\nu y_\nu^2}{N_{\nu,k}^2}$$

$$\lambda_{k+1} = \frac{\sum_{\nu=1}^{n} \frac{\mu_\nu y_\nu^2}{N_{\nu,k}^2}}{\sum_{\nu=1}^{n} \frac{y_\nu^2}{N_{\nu,k}^2}}. \tag{31}$$

Here we again disregard the μ_ν corresponding to the vanishing y_ν and denote the remaining distinct μ_ν by (11). Then (31) becomes, with appropriate positive p_μ:

$$\lambda_{k+1} = \frac{\sum_{\mu=1}^{m} \frac{\sigma_\mu p_\mu}{M_{\mu,k}^2}}{\sum_{\mu=1}^{m} \frac{p_\mu}{M_{\mu,k}^2}} \qquad (k = 0, 1, \ldots), \tag{32}$$

where

$$M_{\mu,k} = \prod_{\varkappa=0}^{k} (\sigma_\mu - \lambda_\varkappa) \qquad (\mu = 1, \ldots, m;\ k = 0, 1, \ldots). \tag{33}$$

11. Again denote by σ one of the σ_μ and by p, M_k the corresponding p_μ, $M_{\mu,k}$. Then from (32) follows

$$\lambda_{k+1} - \sigma = \frac{\sum_{\mu=1}^{m}{}' \frac{(\sigma_\mu - \sigma) p_\mu}{M_{\mu,k}^2}}{\frac{p}{M_k^2} + \sum_{\mu=1}^{m}{}' \frac{p_\mu}{M_{\mu,k}^2}}, \tag{34}$$

where in the sums $\sum'$ the terms with the index μ for which $\sigma_\mu = \sigma$ are to be omitted. From (34) putting

$$D_k = \frac{\sum_{\mu=1}^{m}{}' \frac{(\sigma_\mu - \sigma) p_\mu}{M_{\mu,k}^2}}{p + \sum_{\mu=1}^{m}{}' p_\mu \frac{M_k^2}{M_{\mu,k}^2}}, \tag{35}$$

we have again

$$\lambda_{k+1} - \sigma = D_k M_k^2. \tag{36}$$

Now put

$$d = \min_{\sigma_\mu \neq \sigma} |\sigma_\mu - \sigma|, \qquad P = \sum_{\mu=1}^{m} p_\mu, \qquad K = \frac{P}{p} (\sigma_m - \sigma_1), \tag{37}$$

and take a $\delta > 0$ such that

(38) $$\delta < \delta' = \operatorname{Min}\left(\frac{d}{2}, \frac{d^2}{4K}\right).$$

Then we have obviously

(39) $$\varepsilon \equiv \frac{\delta}{d-\delta} < 1,$$

$$K\varepsilon^2 = \delta \frac{K\delta}{(d-\delta)^2} < \delta \frac{d^2/4}{\left(d-\frac{d}{2}\right)^2},$$

(40) $$K\varepsilon^2 < \delta.$$

12. Suppose now that we have for $\varkappa = 0, 1, \ldots, k$

(41) $$|\lambda_\varkappa - \sigma| \le \delta \qquad (\varkappa = 0, 1, \ldots, k).$$

Then we have from (33), since $\sigma_\mu \neq \sigma$,

(42) $$|M_{\mu,k}| \ge (d-\delta)^{k+1}$$

and therefore by (37) and (35), since all terms in the denominator of (35) are positive,

$$|D_k| \le \frac{1}{p} P \frac{\operatorname{Max}|\sigma_\mu - \sigma|}{(d-\delta)^{2k+2}},$$

$$|D_k| \le K(d-\delta)^{-2k-2}.$$

It follows then from (39) and (40)

(43) $$|D_k \delta^{2k+2}| \le K\varepsilon^{2k+2} < \varepsilon^{2k}\delta \qquad (k = 0, 1, \ldots).$$

We have now from (36), since $M_k^2 \le \delta^{2k+2}$ by (41),

(44) $$|\lambda_{k+1} - \sigma| \le \varepsilon^{2k}\delta \qquad (k = 0, 1, \ldots).$$

We see that the sequence λ_k is convergent to σ and we have for each k: $|\lambda_k - \sigma| \le \delta$, provided only that

(45) $$|\lambda_0 - \sigma| \le \delta.$$

13. We now prove that *if $\lambda_k \to \sigma$ and if none of the λ_k is equal to σ, then*

(46) $$\frac{\lambda_{k+1} - \sigma}{(\lambda_k - \sigma)^3} \to \gamma \qquad (\varkappa \to \infty),$$

where γ is a positive constant equal to one of the quotients $\frac{1}{(\sigma_\mu - \sigma)^2}$.

We assume first that (45) is satified. Observe that from (44) and (45) by definition of M_k we have

(47) $$M_k^2 \le \delta^2 \prod_{\varkappa=1}^{k} (\delta\varepsilon^{2(\varkappa-1)})^2 = \delta^{2k+2}\varepsilon^{2k(k-1)}.$$

On the other hand, if we divide both sides of (36) by $\lambda_k - \sigma$ and use (47), it follows that

$$\left|\frac{\lambda_{k+1} - \sigma}{\lambda_k - \sigma}\right| = |D_k M_k M_{k-1}| \le |D_k| \delta^2 \left(\prod_{\varkappa=1}^{k-1} (\delta\varepsilon^{2(\varkappa-1)})^2\right) \delta\varepsilon^{2(k-1)}$$
$$= |D_k| \delta^{2k+2} \frac{1}{\delta} \varepsilon^{2(k-1)^2},$$

and therefore by (43)

$$\left|\frac{\lambda_{k+1}-\sigma}{\lambda_k-\sigma}\right| < \varepsilon^{k^2}. \tag{48}$$

On the other hand, if we write (36) for k and $k-1$ and divide, we obtain

$$\frac{\lambda_{k+1}-\sigma}{(\lambda_k-\sigma)^3} = \frac{D_k}{D_{k-1}}; \tag{49}$$

therefore we have to discuss D_k as given by (35).

14. In the formula (33) for $M_{\mu,k}$ the general factor $\sigma_\mu-\lambda_\varkappa$ can be written as $(\sigma_\mu-\sigma)\left(1+\frac{\sigma-\lambda_\varkappa}{\sigma_\mu-\sigma}\right)$. Therefore, putting

$$T_{\mu,k} = \prod_{\varkappa=0}^{k}\left(1+\frac{\sigma-\lambda_\varkappa}{\sigma_\mu-\sigma}\right), \tag{50}$$

we have

$$M_{\mu,k} = (\sigma_\mu-\sigma)^{k+1}\, T_{\mu,k}. \tag{51}$$

Since $\sum_{\varkappa=0}^{\infty}|\sigma-\lambda_\varkappa|$ is convergent by (48), we see that

$$T_{\mu,k} \to t_\mu \qquad (k\to\infty), \tag{52}$$

where t_μ is finite and positive, since $\left|\frac{\sigma-\lambda_\varkappa}{\sigma_\mu-\sigma}\right| \le \frac{\delta}{d} < \frac{1}{2}$ $(\varkappa=0, 1, \ldots)$. On the other hand we have

$$T_{\mu,k+1} - T_{\mu,k} = T_{\mu,k}\,\frac{\sigma-\lambda_{k+1}}{\sigma_\mu-\sigma} = O(\sigma-\lambda_{k+1})$$

and therefore further

$$\sum_{\varkappa=k}^{\infty}(T_{\mu,\varkappa+1}-T_{\mu,\varkappa}) = O(\sigma-\lambda_{k+1})\sum_{\varkappa=k+1}^{\infty}\left|\frac{\lambda_\varkappa-\sigma}{\lambda_{k+1}-\sigma}\right|.$$

But, by (48), the sum on the right is convergent and $<\frac{1}{1-\varepsilon}$; therefore, since the sum on the left is $t_\mu - T_{\mu,k}$, we obtain

$$T_{\mu,k} = t_\mu + O(\sigma-\lambda_{k+1}). \tag{53}$$

15. From this it follows further by (51) and (37) that

$$\frac{p_\mu}{M_{\mu,k}^2} = \frac{p_\mu}{t_\mu^2(\sigma_\mu-\sigma)^{2k+2}} + O\left(\frac{\lambda_{k+1}-\sigma}{d^{2k}}\right).$$

Therefore, if we put

$$p'_\mu = \frac{p_\mu}{t_\mu^2(\sigma_\mu-\sigma)},$$

we obtain finally for the numerator of D_k in (35) the expression

$$\sum_{\mu=1}^{m}{}' \frac{p'_\mu}{|\sigma_\mu-\sigma|^{2k}} + O\left(\frac{\lambda_{k+1}-\sigma}{d^{2k}}\right).$$

On the other hand the denominator in (35) can be written as $p+\eta_k$, where

$$\eta_k = \sum_{\mu=1}^{m}{}' p_\mu \prod_{\varkappa=0}^{k}\left(\frac{\sigma-\lambda_\varkappa}{\sigma_\mu-\lambda_\varkappa}\right)^2 \to 0 \qquad (k\to\infty),$$

since $\sigma - \lambda_\varkappa \to 0$ and $\sigma_\mu - \lambda_\varkappa \to \sigma_\mu - \sigma$. Therefore we now obtain

$$(54) \qquad (p + \eta_k) D_k = \sum_{\mu=1}^{m}{}' \frac{p'_\mu}{(\sigma_\mu - \sigma)^{2k}} + O\left(\frac{\lambda_{k+1} - \sigma}{d^{2k}}\right).$$

16. In the right hand sum in (54) it could happen that some values of $|\sigma_\mu - \sigma|$ occur twice, if there are two σ_μ symmetric with respect to σ, and it could even happen that two such terms cancel each other, if the corresponding p'_μ have the sum 0. Denote the *distinct* quotients $\frac{1}{(\sigma_\mu - \sigma)^2}$ which are not cancelled out by

$$(55) \qquad \gamma = \gamma_1 > \gamma_2 > \cdots > \gamma_r > 0.$$

Then we can write

$$(56) \qquad (p + \eta_k) D_k = \sum_{\varrho=1}^{r} s_\varrho \gamma_\varrho^k + O\left(\frac{\lambda_{k+1} - \sigma}{d^{2k}}\right),$$

where the s_ϱ are non-vanishing constants, as long as there are any terms left, that is if $r \geq 1$.

17. But if we had $r = 0$, it would follow from (56) that

$$D_k = O\left(\frac{\lambda_{k+1} - \sigma}{d^{2k}}\right);$$

introducing this into (36) yields

$$\lambda_{k+1} - \sigma = O\left(\frac{\lambda_{k+1} - \sigma}{d^{2k}} M_k^2\right), \qquad d^{2k} = O(M_k^2),$$

and therefore by (47)

$$\left(\frac{d}{\delta}\right)^{2k} = O\left(\varepsilon^{2k(k-1)}\right),$$

which is impossible, since $0 < \varepsilon < 1$. Therefore we have $r \geq 1$, and it follows from (56) and (48) that

$$(57) \qquad D_k \sim \frac{s_1}{p} \gamma^k \qquad (k \to \infty),$$

and (46) now follows from (49).

Thus far we have proved (46) only under the assumption that (45) holds. However, if we assume more generally that $\lambda_\varkappa \to \sigma$, for a certain $\varkappa_0$ we have $|\lambda_{\varkappa_0} - \sigma| \leq \delta$, and our result above applies if we put $\lambda_{\varkappa+\varkappa_0} = \lambda'_\varkappa$. The theorem stated in Section 13 is now completely proved.

It is hardly necessary to add that our results hold also for Hermitian matrices, for which the discussion above remains valid with some slight and obvious modifications.

Note added October 1957. Professor G. Forsythe has directed my attention to a paper by S. H. Crandall, "Iterative procedures related to relaxation methods for eigenvalue problems" [Proc. Royal Soc. London, **207**, 416–423 (1951)], in which the iteration rules (3), (4) and (28), (29) are discussed. In particular, Professor Crandall establishes the *cubic character* of convergence of $\xi_\varkappa$ in the rule (28), (29). However he does not arrive at our asymptotic formula (46), which is the principal result of our paper.

Mathematische Anstalt der Universität, Basel

(Received June 3, 1957)