11. ERROR BOUNDS FOR ITERATIVE REFINEMENTS

A customary way for approximating eigenelements λ, φ of $T \in BL(X)$ is to consider a nearby simpler operator T_0, solve the eigenvalue problem

$$T_0\varphi_0 = \lambda_0\varphi_0 \ , \ 0 \neq \varphi_0 \in X \ , \ \lambda_0 \in \mathbb{C} \ ,$$

and refine the eigenelements λ_0, φ_0 of T_0 successively to obtain approximations of λ, φ .

In this section we develop some refinement schemes of this type when λ_0 is *simple*. We also show that two main iteration schemes lead to a *simple* eigenvalue λ of T ; a region of isolation for λ from the rest of $\sigma(T)$ is also found. We conclude this section with a discussion of the power method, the inverse iteration and the Rayleigh quotient iteration.

We shall assume throughout this section that λ_0 is a simple eigenvalue of $T_0 \in BL(X)$, and φ_0 (resp., φ_0^*) is an eigenvector of T_0 (resp., T_0^*) corresponding to λ_0 (resp., $\bar{\lambda}_0$) such that $\langle \varphi_0, \varphi_0^* \rangle = 1$. Let P_0 and S_0 denote, as usual, the spectral projection and the reduced resolvent associated with T_0 and λ_0 , respectively. We let $V_0 = T - T_0$, so that $T = T_0 + V_0$, and seek an eigenvector φ of T which satisfies the same condition : $\langle \varphi, \varphi_0^* \rangle = 1$.

We recall the notations introduced in (10.16):

$$\eta_0 = \|V_0\varphi_0\| \ , \ p_0 = \|\varphi_0^*\| \ , \ s_0 = \|S_0\| \ ,$$
$$\alpha_0 = \|V_0S_0\| \ , \ \gamma_0 = \max\{\eta_0 p_0 s_0 \ , \ \alpha_0\} \ .$$

Note that if $\gamma_0 = 0$, then $\eta_0 = 0 = \alpha_0$, so that $V_0P_0 = 0 = V_0S_0$; this implies $V_0 = 0$. We discard this trivial case.

The following function will prove to be very useful:

$$(11.1)\qquad g(t) = \begin{cases} (1-\sqrt{1-4t})/2t\ , & \text{if}\quad 0 < |t| \le 1/4\ , \\ 1\ , & \text{if}\quad t = 0\ , \end{cases}$$

where $\sqrt{\ }$ denotes the principal value of the square root function. It can be seen that g has the power series expansion

$$g(t) = \sum_{k=0}^{\infty} a_k t^k\ ,$$

which converges for $|t| \le 1/4$, and

$$(11.2)\qquad a_0 = 1\ ,\ a_k = \sum_{i=1}^{k} a_{i-1}a_{k-i} = \frac{(2k)!}{(k+1)!k!}\ .$$

Also, we note that $g(1/4) = 2$, and

$$(11.3)\qquad \begin{cases} 1 < |g(t)| \le 2 \\ |[g(t)-1]/t| \le 4 \\ |[g(t)-1-t]/t^2| \le 12 \end{cases} \quad \text{for}\quad 0 < |t| \le 1/4\ .$$

Note that $g(t)$ is a real-valued increasing function of t .

Often it is possible, and also desirable, to develop an iteration scheme which approximates an error vector $\varphi - \varphi_0$ rather than an eigenvector φ itself. We now study two schemes of this type. For the first, we take a clue from the Rayleigh-Schrödinger approach developed in Section 10, and in particular, the formula

$$\varphi_{(k)} = S_0\left[-V_0\varphi_{(k-1)} + \sum_{i=1}^{k-1} \lambda_{(i)}\varphi_{(k-i)}\right]$$

for the k-th coefficient of the series (10.7).

LEMMA 11.1 Let

(11.4) $$\psi_{(1)} = -V_0\varphi_0 \ , \text{ and } \psi_{(k)} = -V_0S_0\psi_{(k-1)} + \sum_{i=1}^{k-1} \lambda_{(i)}S_0\psi_{(k-i)} \ ,$$

for $k = 2,3,\ldots$, where

(11.5) $$\lambda_{(1)} = \langle V_0\varphi_0, \varphi_0^* \rangle \ , \text{ and } \lambda_{(k)} = \langle V_0S_0\psi_{(k-1)}, \varphi_0^* \rangle \ .$$

For $j = 1,2,\ldots$, let

(11.6) $$\psi_j = \psi_{(1)} + \ldots + \psi_{(j)} \ ,$$
$$\lambda_j = \lambda_0 + \lambda_{(1)} + \ldots + \lambda_{(j)} \ , \quad \varphi_j = \varphi_0 + S_0\psi_j \ .$$

Then λ_j and φ_j are the j-th partial sums of the Rayleigh-Schrödinger series (10.4) and (10.7) for $T = T_0 + V_0$, respectively.

If (ψ_j) converges in X to ψ , then (φ_j) converges in X to an eigenvector φ of T satisfying $\langle \varphi, \varphi_0^* \rangle = 1$, and (λ_j) converges to the corresponding eigenvalue $\lambda = \langle T\varphi, \varphi_0^* \rangle$. For $j = 1,2,\ldots,$

(11.7) $$\varphi_j = \varphi_{j-1} + S_0[-(T-\lambda_1 I)\varphi_{j-1} + \sum_{i=2}^{j} (\lambda_i - \lambda_{i-1})\varphi_{j-i}] \ ,$$
$$\lambda_j = \langle T\varphi_{j-1}, \varphi_0^* \rangle \ .$$

Proof Let $\varphi_{(0)} = \varphi_0$ and for $k = 1,2,\ldots$,

$$\varphi_{(k)} = S_0\psi_{(k)} \ .$$

Then for $k = 1,2,\ldots$,

(11.8) $$\psi_{(k)} = -V_0\varphi_{(k-1)} + \sum_{i=1}^{k-1} \lambda_{(i)}\varphi_{(k-1)} \ .$$

Clearly, φ_j and λ_j are the j-th partial sums of (10.7) and (10.4).

Now, let the sequence (ψ_j) converge to ψ in X, i.e., let the sum of the series $\sum_{k=1}^{\infty} \psi_{(k)}$ be ψ. Since S_0 and $V_0 S_0$ are continuous linear operators, it follows that the two series

$$\varphi_0 + \sum_{k=1}^{\infty} \varphi_{(k)} \quad \text{and} \quad \lambda_0 + \sum_{k=1}^{\infty} \lambda_{(k)}$$

converge to $\varphi = \varphi_0 + S_0\psi$ and to λ, say, in X and $\mathbb{C}$, respectively. We show that λ and φ are, in fact, eigenelements of T and that $\langle\varphi,\varphi_0^*\rangle = 1$.

First,

$$\langle\varphi,\varphi_0^*\rangle = \langle\varphi_0,\varphi_0^*\rangle + \langle S_0\psi,\varphi_0^*\rangle = 1 + 0 = 1 .$$

Next,

$$\begin{aligned} T\varphi &= (T_0+V_0)\sum_{k=0}^{\infty} \varphi_{(k)} \\ &= \sum_{k=1}^{\infty} (T_0-\lambda_0 I)\varphi_{(k)} + \lambda_0 \sum_{k=0}^{\infty} \varphi_{(k)} + \sum_{k=1}^{\infty} V_0\varphi_{(k-1)} . \end{aligned}$$

But by (11.8), we have for $k = 1,2,\ldots$,

$$\begin{aligned} (T_0-\lambda_0 I)\varphi_{(k)} &= (T_0-\lambda_0 I)S_0\psi_{(k)} \\ &= (I-P_0)\psi_{(k)} \\ &= \psi_{(k)} + \lambda_{(k)}\varphi_0 \\ &= -V_0\varphi_{(k-1)} + \sum_{i=1}^{k-1} \lambda_{(i)}\varphi_{(k-i)} + \lambda_{(k)}\varphi_0 \\ &= -V_0\varphi_{(k-1)} + \sum_{i=1}^{k} \lambda_{(i)}\varphi_{(k-1)} . \end{aligned}$$

Hence

$$\begin{aligned} T\varphi &= \sum_{k=1}^{\infty} \left[\sum_{i=1}^{k} \lambda_{(i)}\varphi_{(k-i)}\right] + \lambda_0 \sum_{k=0}^{\infty} \varphi_{(k)} \\ &= \sum_{k=0}^{\infty} \left[\sum_{i=0}^{k} \lambda_{(i)}\varphi_{(k-i)}\right] . \end{aligned}$$

Now, series $\sum_{k=0}^{\infty} \varphi_{(k)}$ and $\sum_{k=0}^{\infty} \lambda_{(k)}$ converge in X and $\mathbb{C}$ to φ and λ respectively, and the Cauchy product series $\sum_{k=0}^{\infty} \left[\sum_{i=0}^{k} \lambda_{(i)}\varphi_{(k-i)}\right]$ converges in X to $T\varphi$. Hence by Abel's theorem, $T\varphi = \lambda\varphi$, i.e., φ is an eigenvector of T corresponding to the eigenvalue λ.

To prove (11.7), we note that by (11.8)

$$\begin{aligned}
\varphi_j &= \varphi_0 + S_0\left[\sum_{k=1}^{i} \psi_{(k)}\right] \\
&= \varphi_0 + \sum_{k=1}^{i} S_0\left[-V_0\varphi_{(k-1)} + \sum_{i=1}^{k-1} \lambda_{(i)}\varphi_{(k-i)}\right] \\
&= \varphi_0 - S_0V_0 \sum_{k=1}^{j} \varphi_{(k-1)} + \sum_{k=1}^{j}\sum_{i=1}^{k} \lambda_{(i)}S_0\varphi_{(k-i)} \\
&= \varphi_0 - S_0V_0\varphi_{j-1} + \sum_{i=1}^{j} \lambda_{(i)}S_0\left[\sum_{k=i}^{j} \varphi_{(k-i)}\right] \\
&= \varphi_0 - S_0(T-T_0)\varphi_{j-1} + \sum_{i=1}^{j} (\lambda_i-\lambda_{i-1})S_0\varphi_{j-i} \\
&= \varphi_0 + S_0(T_0-\lambda_0 I)\varphi_{j-1} + S_0\left[-(T-\lambda_1 I)\varphi_{j-1} + \sum_{i=2}^{j} (\lambda_i-\lambda_{i-1})\varphi_{j-i}\right] \\
&= \varphi_{j-1} + S_0\left[-(T-\lambda_1 I)\varphi_{j-1} + \sum_{i=2}^{j} (\lambda_i-\lambda_{i-1})\varphi_{j-i}\right],
\end{aligned}$$

since $S_0(T_0-\lambda_0 I)\varphi_{j-1} = (I-P_0)\varphi_{j-1} = \varphi_{j-1} - \varphi_0$. Also, since $\langle T_0S_0\psi_{j-1}, \overset{*}{\varphi}_0\rangle = \langle S_0\psi_{j-1}, \overset{*}{T}_0\overset{*}{\varphi}_0\rangle = \lambda_0\langle S_0\psi_{j-1}, \overset{*}{\varphi}_0\rangle = 0$,

$$\begin{aligned}
\lambda_j &= \lambda_0 + \lambda_{(1)} + \cdots + \lambda_{(j)} \\
&= \langle T_0\varphi_0, \overset{*}{\varphi}_0\rangle + \langle V_0\varphi_0, \overset{*}{\varphi}_0\rangle + \sum_{i=2}^{j} \langle V_0S_0\psi_{(i-1)}, \overset{*}{\varphi}_0\rangle \\
&= \langle (T_0+V_0)\varphi_0, \overset{*}{\varphi}_0\rangle + \langle (T-T_0)S_0\psi_{j-1}, \overset{*}{\varphi}_0\rangle \\
&= \langle T\varphi_0, \overset{*}{\varphi}_0\rangle + \langle TS_0\psi_{j-1}, \overset{*}{\varphi}_0\rangle \\
&= \langle T(\varphi_0+S_0\psi_{j-1}), \overset{*}{\varphi}_0\rangle \\
&= \langle T\varphi_{j-1}, \overset{*}{\varphi}_0\rangle .
\end{aligned}$$

This completes the proof. //

PROPOSITION 11.2 For $k = 1,2,\ldots$, let $\psi_{(k)}$ be defined by (11.4) and (11.5). Then

$$\|\psi_{(k)}\| \le a_k \eta_0 \gamma_0^{k-1} . \tag{11.9}$$

Let ψ_j, $j = 1,2,\ldots$, be defined by (11.6). If $0 < \gamma_0 \le 1/4$, then (ψ_j) converges to some ψ in X, and we have

$$\|\psi\| \le \eta_0[g(\gamma_0)-1]/\gamma_0 \le 4\eta_0 , \tag{11.10}$$

$$\|\psi-\psi_j\| \le \eta_0[g(\gamma_0)-a_0-\ldots-a_j\gamma_0^j]/\gamma_0 \le 3\eta_0(4\gamma_0)^j , \tag{11.11}$$

for $j = 1,2,\ldots$.

Proof We prove (11.9) by induction on k . Since

$$\|\psi_{(1)}\| = \|V_0\varphi_0\| = \eta_0 ,$$

we see that (11.7) holds for $k = 1$. Now, let $k \ge 2$ and assume that (11.7) holds for all positive integers $\le k - 1$. By the definition of $\psi_{(k)}$,

$$\|\psi_{(k)}\| \le \|V_0 S_0 \psi_{(k-1)}\| + \sum_{i=1}^{k-1} |\lambda_{(i)}| \, \|S_0\| \, \|\psi_{(k-i)}\| .$$

The induction hypothesis now gives

$$\|V_0 S_0 \psi_{(k-1)}\| \le \alpha_0 a_{k-1} \eta_0 \gamma_0^{k-2} \le a_{k-1} \eta_0 \gamma_0^{k-1} ,$$

$$|\lambda_{(1)}| \le \eta_0 p_0 , \quad |\lambda_{(i)}| \le \alpha_0 a_{i-1} \eta_0 \gamma_0^{i-2} p_0 , \quad i = 2,\ldots,k-1 .$$

Hence for $i = 1,\ldots,k - 1$, we have by (11.2)

$$|\lambda_{(i)}| \, \|S_0\| \, \|\psi_{(k-i)}\| \le a_{i-1} a_{k-i} \eta_0 \gamma_0^{k-1} ,$$

$$\|\psi_{(k)}\| \le \left[a_{k-1} + \sum_{i=1}^{k-1} a_{i-1} a_{k-i}\right] \eta_0 \gamma_0^{k-1} = a_k \eta_0 \gamma_0^{k-1} .$$

Thus, (11.9) is established for all $k = 1,2,\dots$.

Let, now, $0 < \gamma_0 \le 1/4$. Then

$$\sum_{k=1}^{\infty} \|\psi_{(k)}\| \le \eta_0 \sum_{k=1}^{\infty} a_k\gamma_0^{k-1} = \eta_0\Big[\sum_{k=1}^{\infty} a_k\gamma_0^{k}\Big]/\gamma_0$$
$$= \eta_0[g(\gamma_0)-1]/\gamma_0 \le 4\eta_0$$

by (11.3) . Since X is a Banach space , every absolutely convergent series is convergent in X ([L], 8.2). Hence $\sum_{k=1}^{\infty} \psi_k$ converges to some ψ in X . The bound given in (11.10) for $\|\psi\|$ is now immediate. Also, for $j = 1,2,\dots$,

$$\|\psi-\psi_j\| \le \sum_{k=j+1}^{\infty} \|\psi_{(k)}\|$$
$$\le \eta_0 \sum_{k=j+1}^{\infty} a_k\gamma_0^{k-1} = \eta_0 \Big[\sum_{k=j+1}^{\infty} a_k\gamma_0^{k}\Big]/\gamma_0$$
$$= \eta_0[g(\gamma_0)-a_0-\dots-a_j\gamma_0^j]/\gamma_0 .$$

But since $0 < \gamma_0 \le 1/4$ and $[g(t) - a_0 - \dots - a_jt^j]/t^{j+1}$ is an increasing function of $t \in (0,1/4]$, we have for $j = 1,2,\dots$,

$$[g(\gamma_0)-a_0-\dots-a_j\gamma_0^j]/\gamma_0^{j+1} \le [g(1/4)-a_0-\dots-a_j\gamma_0^j]/(1/4)^{j+1}$$
$$\le 4^{j+1}[g(1/4)-a_0-a_1/4]$$
$$= 4^{j+1}[2-1-1/4] = 3(4^j) .$$

This proves (11.11). //

The above estimates were first considered in [R]. See also [LN], Proposition 3.1.

Before we turn to another iteration scheme which approximates $\varphi - \varphi_0$, we prove a lemma which shows a connection between the existence of an eigenvector of T and a fixed point of an appropriate function.

LEMMA 11.3 (a) Let $\varphi \in X$. Then the following conditions are equivalent:

(i) φ is an eigenvector of T and $\langle \varphi, \varphi_0^* \rangle = 1$

(ii) φ is a fixed point of the function

$$F(x) = \varphi_0 + S_0[-V_0 x + \langle V_0 x, \varphi_0^* \rangle x] \ , \ x \in X \tag{11.12}$$

(iii) $\varphi = \varphi_0 + S_0 \psi$ for some fixed point ψ of the function

$$\tilde{F}(x) = -V_0(\varphi_0 + S_0 x) + \langle V_0(\varphi_0 + S_0 x), \varphi_0^* \rangle S_0 x \tag{11.13}$$

(b) Let $\psi_1 \in X$,

$$\psi_j = \tilde{F}(\psi_{j-1}) \ , \ j = 2,3,\ldots, \quad \varphi_j = \varphi_0 + S_0 \psi_j \ , \ j = 1,2,\ldots .$$

If (ψ_j) converges in X to ψ, then (φ_j) converges in X to an eigenvector φ of T satisfying $\langle \varphi, \varphi_0^* \rangle = 1$. For $j = 1,2,\ldots$,

$$\begin{aligned} \varphi_j &= \varphi_{j-1} + S_0[-T\varphi_{j-1} + \lambda_j \varphi_{j-1}] \ , \quad \text{where} \\ \lambda_j &= \langle T\varphi_{j-1}, \varphi_0^* \rangle . \end{aligned} \tag{11.14}$$

Proof (a) Let (i) hold. Then $T\varphi = \lambda\varphi$ for some $\lambda \in \mathbb{C}$. Taking scalar product with φ_0^* on both sides, we have

$$\langle T\varphi, \varphi_0^* \rangle = \lambda \langle \varphi, \varphi_0^* \rangle = \lambda \ ,$$

so that $T\varphi = \langle T\varphi, \varphi_0^* \rangle \varphi$, i.e.,

$$\begin{aligned} (T_0 + V_0)\varphi &= \langle (T_0 + V_0)\varphi, \varphi_0^* \rangle \varphi \\ &= \langle \varphi, T_0^* \varphi_0^* \rangle \varphi + \langle V_0 \varphi, \varphi_0^* \rangle \varphi \\ &= \lambda_0 \varphi + \langle V_0 \varphi, \varphi_0^* \rangle \varphi . \end{aligned}$$

Hence

$$(T_0 - \lambda_0 I)\varphi = -V_0 \varphi + \langle V_0 \varphi, \varphi_0^* \rangle \varphi .$$

Applying S_0 on both sides, we see that

$$\varphi - \varphi_0 = (I-P_0)\varphi = S_0(T_0-\lambda_0 I)\varphi = S_0[-V_0\varphi + \langle V_0\varphi,\varphi_0^*\rangle\varphi] \;,$$

i.e., $\varphi = F(\varphi)$. Thus (ii) holds.

If (ii) holds, and we let

$$\psi = -V_0\varphi + \langle V_0\varphi,\varphi_0^*\rangle(\varphi-\varphi_0) \;,$$

then

$$\varphi_0 + S_0\psi = \varphi_0 + S_0[-V_0\varphi + \langle V_0\varphi,\varphi_0^*\rangle\varphi] = F(\varphi) = \varphi \;,$$

and also,

$$\tilde{F}(\psi) = -V_0\varphi + \langle V_0\varphi,\varphi_0^*\rangle S_0\psi = -V_0\varphi + \langle V_0\varphi,\varphi_0^*\rangle(\varphi-\varphi_0) = \psi \;,$$

i.e., (iii) holds.

Next, let (iii) hold. Then

$$\langle\varphi,\varphi_0^*\rangle = \langle\varphi_0+S_0\psi,\varphi_0^*\rangle = \langle\varphi_0,\varphi_0^*\rangle = 1 \;.$$

Also,

$$T\varphi = (T-T_0)\varphi + (T_0-\lambda_0 I)\varphi + \lambda_0\varphi = V_0\varphi + (T_0-\lambda_0 I)\varphi + \lambda_0\varphi \;.$$

Now,

$$\begin{aligned}(T_0-\lambda_0 I)\varphi &= (T_0-\lambda_0 I)(\varphi_0+S_0\psi)\\ &= (T_0-\lambda_0 I)S_0\tilde{F}(\psi)\\ &= (T_0-\lambda_0 I)S_0[-V_0\varphi+\langle V_0\varphi,\varphi_0^*\rangle(\varphi-\varphi_0)]\\ &= (I-P_0)(-V_0\varphi+\langle V_0\varphi,\varphi_0^*\rangle\varphi)\\ &= -V_0\varphi + \langle V_0\varphi,\varphi_0^*\rangle\varphi \;.\end{aligned}$$

Hence

$$T\varphi = \langle V_0\varphi,\varphi_0^*\rangle\varphi + \lambda_0\varphi = \langle T\varphi,\varphi_0^*\rangle\varphi \;,$$

since $V_0 = T - T_0$, and $\langle T_0\varphi,\varphi_0^*\rangle = \langle\varphi,T_0^*\varphi_0^*\rangle = \lambda_0\langle\varphi,\varphi_0^*\rangle = \lambda_0$. Thus, φ is an eigenvector of T and $\langle\varphi,\varphi_0^*\rangle = 1$, i.e., (i) holds.

(b) Let $\psi_1 \in X$, $\psi_j = \tilde{F}(\psi_{j-1})$, $j = 2,3,\ldots$. If $\psi_j \to \psi$ in X , then clearly, $\psi = \tilde{F}(\psi)$, i.e., ψ is a fixed point of $\tilde{F}$. Now,

$\varphi_j = \varphi_0 + S_0\psi_j$, converges to $\varphi = \varphi_0 + S_0\psi$, which is an eigenvector of T satisfying $\langle\varphi,\varphi_0^*\rangle = 1$, by part (a).

Finally, for j = 1,2,... ,

$$\begin{aligned}\varphi_j &= \varphi_0 + S_0[-V_0(\varphi_0+S_0\psi_{j-1}) + \langle V_0(\varphi_0+S_0\psi_{j-1}),\varphi_0^*\rangle S_0\psi_{j-1}] \\ &= \varphi_0 + S_0[-V_0\varphi_{j-1} + \langle V_0\varphi_{j-1},\varphi_0^*\rangle\varphi_{j-1}] \\ &= \varphi_0 + S_0[-T\varphi_{j-1} + \langle T\varphi_{j-1},\varphi_0^*\rangle\varphi_{j-1} + (T_0-\lambda_0 I)\varphi_{j-1}] \\ &= \varphi_0 + S_0[-T\varphi_{j-1} + \langle T\varphi_{j-1},\varphi_0^*\rangle\varphi_{j-1}] + (I-P_0)\varphi_{j-1} \\ &= \varphi_{j-1} + S_0[-T\varphi_{j-1} + \langle T\varphi_{j-1},\varphi_0^*\rangle\varphi_{j-1}] ,\end{aligned}$$

which proves (11.14). //

PROPOSITION 11.4 Let $0 < \gamma_0 < 1/4$. Then the function $\tilde{F}$ given by (11.13) has a unique fixed point ψ in X such that

$$\|\psi\| \le \eta_0[g(\gamma_0)-1]/\gamma_0 \le 4\eta_0 . \tag{11.15}$$

Let

$$\psi_1 = -V_0\varphi_0 , \quad \psi_j = \tilde{F}(\psi_{j-1}) , \quad j = 2,3,\dots . \tag{11.16}$$

Then for j = 1,2,...,

$$\|\psi-\psi_j\| \le \eta_0[g(\gamma_0)-1-\gamma_0][2\gamma_0 g(\gamma_0)]^{j-1}/\gamma_0 \le 3\eta_0(4\gamma_0)^j , \tag{11.17}$$

so that $\psi_j \to \psi$ as $j \to \infty$.

Proof Let $r = \eta_0[g(\gamma_0)-1]/\gamma_0$ and $E = \{x \in X : \|x\| \le r\}$. If $\eta_0 = 0$, then $\psi = 0$ is the unique fixed point of $\tilde{F}$ in E . Now, assume $\eta_0 \ne 0$. Then for $x \in E$,

$$\begin{aligned}\|\tilde{F}(x)\| &\le \eta_0 + \alpha_0 r + \eta_0 p_0 s_0 r + \alpha_0 p_0 s_0 r^2 \\ &\le \eta_0 + 2\gamma_0 r + \gamma_0^2 r^2/\eta_0 \\ &\le r ,\end{aligned}$$

if $$\frac{\eta_0[(1-2\gamma_0) - \sqrt{1-4\gamma_0}]}{2\gamma_0^2} \le r \le \frac{\eta_0[(1-2\gamma_0) + \sqrt{1-4\gamma_0}]}{2\gamma_0^2} ,$$

i.e., $\eta_0[g(\gamma_0)-1]/\gamma_0 \le r \le \eta_0[1-\gamma_0-\gamma_0 g(\gamma_0)]/\gamma_0^2$. Thus, for $\|x\| \le r$, we have $\|\tilde{F}(x)\| \le r$, i.e., $\tilde{F}$ maps E into E . Now, for $x,y \in E$,

$$\tilde{F}(x) - \tilde{F}(y) = -V_0 S_0(x-y) + \langle V_0\varphi_0, \varphi_0^*\rangle S_0(x-y) \\ + \langle V_0 S_0(x-y), \varphi_0^*\rangle S_0 y + \langle V_0 S_0 x, \varphi_0^*\rangle S_0(x-y) ,$$

so that

$$\begin{aligned} \|\tilde{F}(x) - \tilde{F}(y)\| &\le (\alpha_0+\eta_0 p_0 s_0+\alpha_0 p_0 s_0 r+\alpha_0 p_0 s_0 r)\|x-y\| \\ &\le 2(\gamma_0+\gamma_0^2 r/\eta_0)\|x-y\| \\ &= 2(\gamma_0+\gamma_0[g(\gamma_0)-1])\|x-y\| \\ &= 2\gamma_0 g(\gamma_0)\|x-y\| = (1-\sqrt{1-4\gamma_0})\|x-y\| . \end{aligned}$$

Since $\gamma_0 < 1/4$, we have $1 - \sqrt{1-4\gamma_0} < 1$ and $\tilde{F}$ is a contraction from E to E . By Banach's contraction mapping theorem ([L], p.322), $\tilde{F}$ has a unique fixed point ψ in E . Then $\|\psi\| \le r = \eta_0[g(\gamma_0)-1]/\gamma_0$, proving (11.15).

Next, $\psi_1 = -V_0\varphi_0 = \tilde{F}(0)$ lies in E . Also, for $j = 1,2,\dots$,

$$\begin{aligned} \|\psi-\psi_j\| &= \|\tilde{F}(\psi)-\tilde{F}(\psi_{j-1})\| \\ &\le [2\gamma_0 g(\gamma_0)]\|\psi-\psi_{j-1}\| \\ &\dots \\ &\le [2\gamma_0 g(\gamma_0)]^{j-1}\|\psi-\psi_1\| . \end{aligned}$$

Now,

$$\psi - \psi_1 = \tilde{F}(\psi) - \psi_1 = -V_0 S_0\psi + \langle V_0\varphi_0, \varphi_0^*\rangle S_0\psi + \langle V_0 S_0\psi, \varphi_0^*\rangle S_0\psi ,$$

and hence

$$\|\psi-\psi_1\| \le (\alpha_0+\eta_0 p_0 s_0)r + \alpha_0 p_0 s_0 r^2$$
$$\le 2\gamma_0 r + \gamma_0^2 r^2/\eta_0$$
$$\le r - \eta_0$$
$$= \eta_0[g(\gamma_0)-1-\gamma_0]/\gamma_0 \ .$$

Thus, the first inequality in (11.17) holds for $j = 1,2,\ldots$. The remaining part follows from (11.3). //

We are now ready to state and prove an important result about the two iteration schemes, one based on the Rayleigh-Schrödinger series and the other on the fixed point principle.

THEOREM 11.5 For $j = 1,2,\ldots,$ let φ_j be defined either by (11.7):

(11.18) $$\varphi_j = \varphi_{j-1} + S_0\left[-(T-\lambda_1 I)\varphi_{j-1} + \sum_{i=2}^{j} (\lambda_i-\lambda_{i-1})\varphi_{j-i}\right]$$

or by (11.14):

(11.19) $$\varphi_j = \varphi_{j-1} + S_0[-T\varphi_{j-1} + \lambda_j\varphi_{j-1}] \ ,$$

where, in both cases,

$$\lambda_j = \langle T\varphi_{j-1}, \varphi_0^*\rangle \ .$$

Let $0 < \gamma_0 < 1/4$. Then (φ_j) converges to an eigenvector φ of T satisfying $\langle\varphi,\varphi_0^*\rangle = 1$, and (λ_j) converges to the corresponding eigenvalue $\lambda = \langle T\varphi,\varphi_0^*\rangle$. We have

(11.20)
$$\|\varphi-\varphi_0\| \le \eta_0 s_0[g(\gamma_0)-1]/\gamma_0 \le 4\eta_0 s_0 \ ,$$
$$\|\varphi-\varphi_j\| \le 3\eta_0 s_0(4\gamma_0)^j \ , \ j = 1,2,\ldots$$

$$|\lambda-\lambda_0| \le \eta_0 p_0\left[1 + \frac{\alpha_0}{\gamma_0}[g(\gamma_0)-1]\right] \le \eta_0 p_0 g(\gamma_0) \le 2\eta_0 p_0 \ ,$$

$$\text{(11.21)} \qquad |\lambda-\lambda_1| \le \eta_0 p_0 \frac{\alpha_0}{\gamma_0}[g(\gamma_0)-1] \le \eta_0 p_0[g(\gamma_0)-1] \le 4\eta_0 p_0 \gamma_0 \ ,$$

$$|\lambda-\lambda_j| \le 3\eta_0 p_0 \alpha_0 (4\gamma_0)^{j-1} \le \frac{3}{4}\eta_0 p_0 (4\gamma_0)^j \ , \ j = 2,3,\dots \ .$$

Proof Since $\gamma_0 < 1/4$, it follows by Proposition 11.2 and Lemma 11.1 in case the φ_j's are defined by (11.18), and by Proposition 11.4 and Lemma 11.3 in case the φ_j's are defined by (11.19), that $\varphi_j \to \varphi$ and $\lambda_j \to \lambda$ such that $T\varphi = \lambda\varphi$ and $\langle \varphi, \varphi_0^* \rangle = 1$.

The bounds in (11.20) are immediate from (11.10) and (11.11) in the first case, and from (11.15) and (11.17) in the second, since $\varphi = \varphi_0 + S_0\psi$ and $\varphi_j = \varphi_0 + S_0\psi_j$, $j = 1,2,\dots$. (Since $\psi_1 = -(T-T_0)\varphi_0$ and $S_0T_0\varphi_0 = \lambda_0 S_0\varphi_0 = 0$, the case $j = 1$ follows.) Similarly, the bounds in (11.21) follow if we observe that

$$\begin{aligned} \lambda - \lambda_0 &= \langle T\varphi, \varphi_0^* \rangle - \langle T_0\varphi_0, \varphi_0^* \rangle \\ &= \langle (T-T_0)\varphi_0, \varphi_0^* \rangle + \langle T(\varphi-\varphi_0), \varphi_0^* \rangle \\ &= \langle (T-T_0)\varphi_0, \varphi_0^* \rangle + \langle (T-T_0)S_0\psi, \varphi_0^* \rangle \ , \end{aligned}$$

and for $j = 1,2,\dots$,

$$\begin{aligned} \lambda - \lambda_j &= \langle T(\varphi-\varphi_{j-1}), \varphi_0^* \rangle \\ &= \langle (T-T_0)(\varphi-\varphi_{j-1}), \varphi_0^* \rangle \\ &= \begin{cases} \langle (T-T_0)S_0\psi, \varphi_0^* \rangle \ , & \text{if } j = 1 \\ \langle (T-T_0)S_0(\psi-\psi_{j-1}), \varphi_0^* \rangle \ , & \text{if } j = 2,3,\dots \ . \end{cases} \end{aligned}$$

//

The iteration scheme (11.19) was considered along with some error estimates, and its connection with Newton's method was discussed in [RO] . See also [A], p.145, where the iteration scheme (11.19) is denoted by DCA1.

Now we consider the question about the simplicity of λ and its isolation from the rest of $\sigma(T)$. In this connection, we first prove some preliminary results.

LEMMA 11.6 ([LN], Lemma 3.3) Let $T \in BL(X)$ and φ be an eigenvector of T corresponding to an eigenvalue λ . Let $x_0^* \in X$ with $\langle \varphi, x_0^* \rangle = 1$. Consider the projection

$$Qx = \langle x, x_0^* \rangle \varphi \ , \quad x \in X \ .$$

If we let $(I-Q)(X) = Z$, then

$$\sigma(T) \subset \{\lambda\} \cup \sigma((I-Q)T|_Z) \ .$$

If $\lambda \in \rho((I-Q)T|_Z)$, then λ is a simple eigenvalue of T .

Proof Note that Q is a projection since $\langle \varphi, x_0^* \rangle = 1$. As φ is an eigenvector of T corresponding to λ , we have $TQ = \lambda Q$. Hence $QTQ = \lambda Q$ and $(I-Q)TQ = 0$, so that

$$\begin{aligned} T &= [Q + (I-Q)]T[Q + (I-Q)] \\ &= \lambda Q + QT(I-Q) + (I-Q)T(I-Q) \ . \end{aligned}$$

Let $A = (I-Q)T|_Z$. If $z \neq \lambda$ and $z \in \rho(A)$, then we can verify that $z \in \rho(T)$; in fact,

$$R(T,z) = \frac{Q}{\lambda - z} + \frac{QT\,R(A,z)(I-Q)}{z - \lambda} + R(A,z)(I-Q) \ .$$

Hence $\sigma(T) \subset \{\lambda\} \cup \sigma(A)$, as desired. (Cf. Problem 6.6.)

Let, now, $\lambda \in \rho(A)$. Since $\sigma(A)$ is a closed set, we see that λ is an isolated spectral value of T . Let a curve Γ in $\rho(T)$ separate λ from $\sigma(A)$. Then by integrating the above expression for $R(T,z)$ over Γ , we see that the spectral projection P_λ associated with T and λ is given by

$$P_\lambda = Q + Q\left[\frac{-1}{2\pi i}\int_\Gamma \frac{T\,R(A,z)(I-Q)}{z-\lambda}dz\right] .$$

Hence $P_\lambda(X) \subset Q(X)$. But Q is of rank 1 by definition, and $P_\lambda \neq 0$. Thus, $P_\lambda(X)$ is also of rank 1 , i.e., λ is a simple eigenvalue of T . //

PROPOSITION 11.7 (Cf. [LN], Theorem 3.4.) Let φ be an eigenvector of T corresponding to an eigenvalue λ satisfying $\langle\varphi,\varphi_0^*\rangle = 1$. Assume that $\alpha_0 + \alpha_0 p_0\|\varphi-\varphi_0\| < 1$. Then the disk

$$(11.22) \qquad \Delta_0 = \left\{z \in \mathbb{C} : |z-\lambda_0| < \frac{1-\alpha_0-\alpha_0 p_0\|\varphi-\varphi_0\|}{s_0}\right\}$$

contains no spectral point of T other than λ . If $\lambda \in \Delta_0$, then λ is simple.

Proof Let

$$Qx = \langle x,\varphi_0^*\rangle\varphi \ , \ x \in X \ , \ (I-Q)(X) = Z \ , \quad \text{and} \quad A = (I-Q)T|_Z \ .$$

By Lemma 11.6, it is enough to show that

$$\Delta_0 \subset \rho(A) \ .$$

First we show that the centre λ_0 of the disk Δ_0 belongs to $\rho(A)$. Note that $(I-P_0)(X) = (I-Q)(X) = Z$, and hence

$$\begin{aligned} A &= [(I-P_0)-(Q-P_0)][T_0+(T-T_0)]|_Z \\ &= A_1 + A_2 + A_3 \ , \end{aligned}$$

where

$$A_1 = T_0|_Z \ , \ A_2 = (I-P_0)(T-T_0)|_Z \ , \quad \text{and} \quad A_3 = (Q-P_0)P_0(T_0-T)|_Z \ ,$$

as $(Q-P_0)(I-P_0) = 0$ and T_0 commutes with $I - P_0$. Now, by the spectral decomposition theorem (Theorem 6.3) we see that $\lambda_0 \in \rho(A_1)$. In fact,

$$(A_1-\lambda_0 I)^{-1} = S_0|_Z .$$

But

$$r_\sigma(A_2(A_1-\lambda_0 I)^{-1}) \le r_\sigma((T-T_0)S_0) \le \alpha_0 < 1 ,$$

so that by Theorem 9.1, $\lambda_0 \in \rho(A_1+A_2)$, and by (9.10),

$$(A_1+A_2-\lambda_0 I)^{-1} = (A_1-\lambda_0 I)^{-1} \sum_{k=0}^{\infty} [-A_2(A_1-\lambda_0 I)^{-1}]^k .$$

Hence

(11.23) $$(A_1+A_2-\lambda_0 I)^{-1}(I-P_0) = S_0 \sum_{k=0}^{\infty} [(T_0-T)S_0]^k .$$

This shows that

(11.24) $$\|(A_1+A_2-\lambda_0 I)^{-1}(I-P_0)\| \le s_0/(1-\alpha_0) .$$

Since $\alpha_0 < 1$, we see by (11.23),

$$\|(T_0-T)(A_1+A_2-\lambda_0 I)^{-1}(I-P_0)\| = \|(T_0-T)S_0 \sum_{k=0}^{\infty} [(T_0-T)S_0]^k\|$$
$$\le \alpha_0/(1-\alpha_0) .$$

Also,

$$\|A_3(A_1+A_2-\lambda_0 I)^{-1}\| \le \|A_3(A_1+A_2-\lambda_0 I)^{-1}(I-P_0)\|$$
$$= \|(Q-P_0)P_0(T-T_0)(A_1+A_2-\lambda_0 I)^{-1}(I-P_0)\|$$

(11.25) $$\le \|\varphi-\varphi_0\| p_0 \alpha_0/(1-\alpha_0) .$$

But $\beta_0 \equiv \|\varphi-\varphi_0\| p_0\alpha_0/(1-\alpha_0) < 1$, by assumption. This shows that $\lambda_0 \in \rho(A_1+A_2+A_3) = \rho(A)$, and

$$(A-\lambda_0 I)^{-1} = (A_1+A_2-\lambda_0 I)^{-1} \sum_{k=0}^{\infty} [-A_3(A_1+A_2-\lambda_0 I)^{-1}]^k .$$

Hence

(11.26) $$(A-\lambda_0 I)^{-1}(I-P_0) = (A_1+A_2-\lambda_0 I)^{-1}(I-P_0) \sum_{k=0}^{\infty} [-A_3(A_1+A_2-\lambda_0 I)^{-1}]^k .$$

Let, now, $z \in \Delta_0$. To conclude $z \in \rho(A)$, it is enough to prove that

$$|z-\lambda_0| < 1/r_\sigma((A-\lambda_0 I)^{-1}) .$$

But by (11.26), (11.24) and (11.25), we have

$$r_\sigma((A-\lambda_0 I)^{-1}) = r_\sigma((A-\lambda_0 I)^{-1}(I-P_0)) \leq \|(A-\lambda_0 I)^{-1}(I-P_0)\|$$
$$\leq s_0/(1-\alpha_0)(1-\beta_0) = s_0/(1-\alpha_0-\alpha_0 p_0\|\varphi-\varphi_0\|) .$$

Since $z \in \Delta_0$, we have $|z-\lambda_0| < (1-\alpha_0-\alpha_0 p_0\|\varphi-\varphi_0\|)/s_0 \leq 1/\|(A-\lambda_0 I)^{-1}\|$. The proof of the proposition is now complete. //

THEOREM 11.8 Let $0 < \gamma_0 < 1/4$. Both the iteration schemes (11.18) and (11.19) give the same eigenelements λ and φ of T ; λ is a simple eigenvalue of T ,

$$|\lambda-\lambda_0| \leq \frac{1 - \sqrt{1-4\gamma_0}}{2s_0} ,$$

and there is no other spectral value of T lying in the disk

(11.27) $$D_0 = \left\{z \in \mathbb{C} : |z-\lambda_0| < \frac{1 + \sqrt{1-4\gamma_0}}{2s_0}\right\} .$$

In particular, λ is the nearest spectral value of T from λ_0 .

Proof For both the iteration schemes, we have by (11.20),

$$\|\varphi-\varphi_0\| \leq \eta_0 s_0[g(\gamma_0)-1]/\gamma_0 ,$$

so that

$$1 - \alpha_0 - \alpha_0 p_0\|\varphi-\varphi_0\| \geq 1 - \gamma_0 - \alpha_0\eta_0 p_0 s_0[g(\gamma_0)-1]/\gamma_0$$
$$\geq 1 - \gamma_0 - \gamma_0[g(\gamma_0)-1]$$
$$= 1 - \gamma_0 g(\gamma_0) = (1 + \sqrt{1-4\gamma_0})/2 > 0 .$$

Thus, we see that the disk D_0 is contained in the disk Δ_0 of Proposition 11.7. Also, for both the schemes, we have by (11.21),

$$|\lambda-\lambda_0| \leq \eta_0 p_0 g(\gamma_0) = \eta_0 p_0 \frac{1-\sqrt{1-4\gamma_0}}{2\gamma_0} \leq \frac{1-\sqrt{1-4\gamma_0}}{2s_0},$$

since $\eta_0 p_0 s_0 \leq \gamma_0$. This shows that $\lambda \in D_0$. Hence λ is a simple eigenvalue of T and there is no other spectral point of T in D_0 . In particular, this says that both the iteration schemes yield the same eigenvalue. Also, since this eigenvalue is simple and the corresponding eigenvector φ satisfies the same constraint $\langle\varphi,\varphi_0^*\rangle = 1$, we see that the two schemes yield the same eigenvector as well. //

REMARKS 11.9 (i) It is interesting to note that although the iteration scheme based on the Rayleigh-Schrödinger procedure and the one based on the fixed point principle are completely different in their approach to the eigenvalue problem, Theorem 11.5 gives the same condition $\gamma_0 < 1/4$ for the convergence as well as the error estimates for both of them. Also the isolation region for λ as given in Theorem 11.8 is identical for the two schemes. It is worthwhile to notice that the essential part of Theorem 11.8 was proved in Theorem 10.5 by an entirely different method.

(ii) If the perturbation operator $V_0 = T - T_0$ satisfies the conditions $P_0V_0P_0 = 0 = S_0V_0S_0$, then one can obtain convergence of the iteration scheme (11.18) under the weaker condition $\eta_0 p_0 s_0 \alpha_0 < 1/4$ (or $\gamma_0 < 1/2$), and sharper error estimates are available. We leave these considerations to Problem 11.6.

(iii) Note that the first iterate $\lambda_1 = \langle T\varphi_0,\varphi_0^*\rangle$ is the generalized Rayleigh quotient of T based at (φ_0,φ_0^*) . If we let

$$T_1 = P_0TP_0 + (I-P_0)T(I-P_0) ,$$

then it is easy to see that $T_1\varphi_0 = \lambda_1\varphi_0$ and $T_1^*\varphi_0^* = \bar{\lambda}_1\varphi_0^*$, so that λ_1 (resp., $\bar{\lambda}_1$) is an eigenvalue of T_1 (resp., T_1^*) with φ_0 (resp., φ_0^*) as a corresponding eigenvector such that $\langle\varphi_0, \varphi_0^*\rangle = 1$. In Lemma 11.6 if we let $T = T_1$, $\varphi = \varphi_0$, $x_0^* = \varphi_0^*$ and $Z = (I-P_0)(X)$, then λ_1 would be a simple eigenvalue of T_1 , provided $\lambda_1 \in \rho((I-P_0)T|_Z)$. Since $\lambda_0 \in \rho(T_0|_Z)$, λ_1 would also belong to $\rho(T_0|_Z)$, if it is sufficiently close to λ_0 . Finally, since

$$(I-P_0)T|_Z = T_0|_Z + (I-P_0)(T-T_0)|_Z ,$$

λ_1 would be in $\rho((I-P_0)T|_Z)$ as well, if $r_\sigma((I-P_0)(T-T_0)|_Z) < 1$. In practice this is often the case when T_0 is sufficiently close to T . Let us then assume that λ_1 is a simple eigenvalue of T_1 . Then the spectral projection associated with T_1 and λ_1 is P_0 itself. Now,

$$T = T_1 + V_1 ,$$

where $V_1 = P_0T(I-P_0) + (I-P_0)TP_0 = P_0T + TP_0 - 2P_0TP_0$, which has rank at most 2 , although T_1 may not be of finite rank. We can carry out the two iterative processes discussed earlier with λ_1 and φ_0 as the initial terms. In this case, we have $P_0V_1P_0 = 0 = S_1V_1S_1$, where S_1 is the reduced resolvent associated with T_1 and λ_1 . (Note that $S_1P_0 = 0 = P_0S_1$.) Accordingly, a better convergence criterion and sharper error estimates are available, as pointed out in (ii) above.

(iv) We have seen in (11.21) that

$$|\lambda-\lambda_0| \le 2\eta_0 p_0 , \quad \text{while} \quad |\lambda-\lambda_1| \le 4\eta_0 p_0 \gamma_0 .$$

Thus, if γ_0 is small, we have a better estimate for $|\lambda-\lambda_1|$ than for $|\lambda-\lambda_0|$. We give another estimate for $\lambda - \lambda_1$ as follows. We have

$$\lambda - \lambda_1 = \langle (T-T_0)(\varphi-\varphi_0), \varphi_0^* \rangle = \langle \varphi-\varphi_0, (T^*-T_0^*)\varphi_0^* \rangle .$$

Let

$$\eta_0^* = \|(T^*-T_0^*)\varphi_0^*\| \ , \ \alpha_0^* = \|(T^*-T_0^*)S_0^*\| \ , \ \gamma_0^* = \max\{\eta_0^* p_0 s_0, \alpha_0^*\}$$

then

$$|\lambda-\lambda_1| \le \|\varphi-\varphi_0\|\eta_0^* \le 4\eta_0\eta_0^* s_0 \ .$$

Again, if η_0^* is small, then the above upper bound for $|\lambda-\lambda_1|$ is better than the one for $|\lambda-\lambda_0|$. This suggests that if we are interested in a higher order accuracy for eigenvalue approximation but are satisfied with a lower order accuracy for eigenvector approximation, then we should carry out two iteration processes simultaneously: one on λ_0, φ_0 and the other on $\bar{\lambda}_0, \varphi_0^*$, and at the j-th step consider the generalized Rayleigh quotient of T based at (φ_j, φ_j^*) :

$$q_j = \langle T\varphi_j, \varphi_j^* \rangle \ / \ \langle \varphi_j, \varphi_j^* \rangle \ ,$$

provided $\langle \varphi_j, \varphi_j^* \rangle \ne 0$. If $\gamma_0 < 1/4$ and $\gamma_0^* < 1/4$, then $\varphi_j \to \varphi$ and $\varphi_j^* \to \varphi^*$, where $\langle \varphi, \varphi^* \rangle \ne 0$ since φ and φ^* are eigenvectors of T and T^* corresponding to the *simple* eigenvalues λ and $\bar{\lambda}$, respectively. Hence $\langle \varphi_j, \varphi_j^* \rangle \ne 0$ for all large j. We then have by (8.11),

$$(11.28) \qquad \begin{aligned} |\lambda-q_j| &\le \|T-\lambda I\| \ \|\varphi-\varphi_j\| \ \|\varphi^*-\varphi_j^*\| \ / \ |\langle \varphi_j, \varphi_j^* \rangle| \\ &\le \|T-\lambda I\|[3\eta_0 s_0(4\gamma_0)^j][3\eta_0^* s_0(4\gamma_0^*)^j] \ / \ |\langle \varphi_j, \varphi_j^* \rangle| \\ &= 9\eta_0\eta_0^* s_0^2(16\gamma_0\gamma_0^*)^j \|T-\lambda I\| \ / \ |\langle \varphi_j, \varphi_j^* \rangle| \ . \end{aligned}$$

In case T_0 and V_0 are self-adjoint operators, there is only one procedure on λ_0, φ_0 to be carried out, and since $\gamma_0^* = \gamma_0$, we see that q_j is an approximation of λ with guaranteed double accuracy as compared with λ_j. Moreover, it is available without any extra work!

(v) All the above procedures are useful if $\gamma_0 < 1/4$. If this is not the case, then one has to look for a sharper error analysis. We merely mention that error bounds in terms of the following quantity

$$\epsilon_0 = \max\left\{\|(V_0S_0)^2\| \ , \ \beta_0^2 \ , \ \alpha_0^{3/2}\beta_0^{1/2}\right\} , \tag{11.29}$$

where $\beta_0 = \eta_0 p_0 s_0$, can be given for the Rayleigh-Schrödinger iteration scheme (11.18) as well as for the fixed point iteration scheme (11.19): Let $\sqrt{\epsilon_0} < 1/4$. Then for $j = 0,1,\dots$,

$$\begin{aligned} \|\varphi-\varphi_{2j}\| \ , \ |\lambda-\lambda_{2j}| &= O(\eta_0(4\sqrt{\epsilon_0})^{2j} \ , \\ \|\varphi-\varphi_{2j+1}\| \ , \ |\lambda-\lambda_{2j+1}| &= O(\eta_0\gamma_0(4\sqrt{\epsilon_0})^{2j}) \ . \end{aligned} \tag{11.30}$$

Note that $\epsilon_0 \le \gamma_0^2$. If $\sqrt{\epsilon_0} < 1/4$, but $\gamma_0 \ge 1/4$, then we have better bounds for the successive iterates at *every other* step. See Problems 11.1 and 11.4. (See Table 19.4, Rayleigh-Schrödinger and fixed point schemes.)

We have seen in Lemma 11.3 that $\varphi \in X$ is an eigenvector of T and satisfies $\langle\varphi,\varphi_0^*\rangle = 1$ if and only if

$$\varphi = \varphi_0 + S_0[-V_0\varphi + \langle V_0\varphi,\varphi_0^*\rangle\varphi] \ .$$

Let us assume now that φ is an eigenvector of T satisfying $\langle\varphi,\varphi_0^*\rangle = 1$, and that the corresponding eigenvalue $\lambda = \langle T\varphi,\varphi_0^*\rangle$ of T is *not zero*. Then $\varphi = T\varphi/\langle T\varphi,\varphi_0^*\rangle$, and the above equation becomes

$$\varphi = \varphi_0 + \frac{S_0}{\langle T\varphi,\varphi_0^*\rangle}\left[-V_0T\varphi + \frac{\langle V_0T\varphi,\varphi_0^*\rangle T\varphi}{\langle T\varphi,\varphi_0^*\rangle}\right] .$$

This leads us to consider the following fixed point iteration scheme:

$$\varphi_j = \varphi_0 + \frac{S_0}{\lambda_j}\left[-V_0T\varphi_{j-1} + \frac{\langle V_0T\varphi_{j-1},\varphi_0^*\rangle T\varphi_{j-1}}{\lambda_j}\right] ,$$

where $\lambda_j = \langle T\varphi_{j-1},\varphi_0^*\rangle$ for $j = 1,2,\dots$, provided $\lambda_j \ne 0$.
Substituting $V_0 = T - T_0$ and noting that

$$\frac{S_0}{\lambda_j}\left[T_0T\varphi_{j-1} - \frac{\langle T_0T\varphi_{j-1},\varphi_0^*\rangle T\varphi_{j-1}}{\lambda_j}\right] = \frac{S_0}{\lambda_j}[(T_0-\lambda_0 I)T\varphi_{j-1}] = \frac{T\varphi_{j-1}}{\lambda_j} - \varphi_0 ,$$

we have

$$\text{(11.31)} \qquad \begin{aligned} \varphi_j &= \frac{T\varphi_{j-1}}{\lambda_j} + \frac{S_0T}{\lambda_j}\left[-T\varphi_{j-1} + \frac{\langle T^2\varphi_{j-1},\varphi_0^*\rangle\varphi_{j-1}}{\lambda_j}\right] , \\ \lambda_j &= \langle T\varphi_{j-1},\varphi_0^*\rangle , \end{aligned}$$

provided $\lambda_j \neq 0$ for $j = 1,2,\ldots$. We now prove the convergence of the above <u>modified fixed point iteration scheme</u>, and give error bounds for $\|\varphi-\varphi_j\|$ and $|\lambda-\lambda_j|$ in terms of the quantity $\|(T-T_0)T\|$. Notice that the iterate φ_j of (11.31) can be obtained from the iterate φ_j of (11.19) if we replace φ_{j-1} by $T\varphi_{j-1}/\lambda_j$.

THEOREM 11.10 Let φ be an eigenvector of T corresponding to a simple eigenvalue $\lambda \neq 0$, which satisfies $\langle\varphi,\varphi_0^*\rangle = 1$. Let P denote the spectral projection associated with T and λ . Suppose that there is a constant c such that

$$\text{(11.32)} \qquad |\lambda-\lambda_0| ,\ \|\varphi-\varphi_0\| \le c\|(T-T_0)P\| \le \frac{|\lambda|}{2p_0\|T\|} ,$$

$$\text{(11.33)} \qquad \|(T-T_0)T\| \le \frac{1}{d_0} ,$$

where

$$d_0 = \frac{2s_0}{|\lambda|} + \frac{2c\|T\|\ \|P\|}{|\lambda|^2}\left[p_0 + 2s_0 + 2p_0s_0\left[\frac{\|T\|^2}{|\lambda|} + \frac{|\lambda_0|\,\|T\|}{|\lambda|} + \|T\| + |\lambda|\right]\right] .$$

Then φ_j and λ_j are well defined by (11.31), and $\lambda_j \neq 0$ for $j = 1,2,\ldots$; also

$$\text{(11.34)} \qquad \begin{aligned} \|\varphi-\varphi_j\| &\le c\|(T-T_0)P\|\left[d_0\|(T-T_0)T\|\right]^j , \\ |\lambda-\lambda_j| &\le cp_0\|T-T_0\|\ \|(T-T_0)P\|\left[d_0\|(T-T_0)T\|\right]^{j-1} . \end{aligned}$$

In particular, if $\|(T-T_0)T\| < 1/d_0$, then $\varphi_j \to \varphi$ and $\lambda_j \to \lambda$.

Proof We prove by induction that for $j = 0,1,2,\ldots$

(i) $\|\varphi-\varphi_j\| \le c d_0^j \|(T-T_0)P\| \, \|(T-T_0)T\|^j$,

(ii) $|\lambda-\lambda_{j+1}| \le c d_0^j p_0 \|T\| \, \|(T-T_0)P\| \, \|(T-T_0)T\|^j$,

(iii) $|\lambda_{j+1}| \ge |\lambda|/2$.

Let $j = 0$. Then (i) holds by the initial assumption, and (ii) follows since

$$|\lambda-\lambda_1| = |\langle T(\varphi-\varphi_0),\varphi_0^*\rangle| \le c p_0 \|T\| \, \|(T-T_0)P\| .$$

Since $c p_0 \|T\| \, \|(T-T_0)P\| \le |\lambda|/2$, we see that $|\lambda-\lambda_1| \le |\lambda|/2$ and hence $|\lambda_1| \ge |\lambda|/2$. Now assuming (i), (ii) and (iii) for j , we prove these statements for $j+1$.

Noting that $(T_0-\lambda_0 I)S_0 = I - P_0$, $\lambda_{j+1} \ne 0$, and $P_0 T(-\lambda_{j+1}T\varphi_j + \langle T^2\varphi_j,\varphi_0^*\rangle\varphi_j) = 0$, we have

$$\begin{aligned}
(T_0-\lambda_0 I)\varphi_{j+1} &= (T_0-\lambda_0 I)\frac{T\varphi_j}{\lambda_{j+1}} + \frac{T}{\lambda_{j+1}}\left[-T\varphi_j + \frac{\langle T^2\varphi_j,\varphi_0^*\rangle\varphi_j}{\lambda_{j+1}}\right] \\
&= \frac{1}{\lambda_{j+1}}\left[(T_0-T)T\varphi_j - \lambda_0 T\varphi_j\right] + \frac{\langle T^2\varphi_j,\varphi_0^*\rangle T\varphi_j}{\lambda_{j+1}^2} \\
&= \frac{1}{\lambda_{j+1}}\left[(T_0-T)T(\varphi_j-\varphi) + (T_0-T)T\varphi - \lambda_0 T\varphi_j\right] \\
&\quad + \frac{1}{\lambda_{j+1}^2}\left[\langle T^2(\varphi_j-\varphi),\varphi_0^*\rangle T(\varphi_j-\varphi) + \langle T^2\varphi,\varphi_0^*\rangle T\varphi_j + \langle T^2(\varphi_j-\varphi),\varphi_0^*\rangle T\varphi\right] .
\end{aligned}$$

Now, since $T\varphi = \lambda\varphi$, we can verify that

$$\begin{aligned}
&\frac{(T_0-T)T\varphi}{\lambda_{j+1}} + \frac{\langle T^2\varphi,\varphi_0^*\rangle T\varphi_j}{\lambda_{j+1}^2} - \frac{\lambda_0 T\varphi_j}{\lambda_{j+1}} - (T_0-\lambda_0 I)\varphi \\
&\quad = \frac{(\lambda-\lambda_{j+1})(T_0-\lambda_0 I)\varphi}{\lambda_{j+1}} + \frac{(\lambda^2-\lambda_0\lambda_{j+1})T(\varphi_j-\varphi)+\lambda^2(\lambda-\lambda_{j+1})\varphi}{\lambda_{j+1}^2} .
\end{aligned}$$

Hence

$$(T_0-\lambda_0 I)(\varphi_{j+1}-\varphi) = \frac{1}{\lambda_{j+1}}\left[(T_0-T)T(\varphi_j-\varphi) + (\lambda-\lambda_{j+1})(T_0-\lambda_0 I)\varphi\right]$$
$$+ \frac{1}{\lambda_{j+1}^2}\Big[[\langle T^2(\varphi_j-\varphi),\varphi_0^*\rangle + \lambda^2-\lambda_0\lambda_{j+1}]T(\varphi_j-\varphi)$$
$$+ \lambda[\langle T^2(\varphi_j-\varphi),\varphi_0^*\rangle + \lambda(\lambda-\lambda_{j+1})]\varphi\Big] .$$

Applying S_0 on both sides and noting that $P_0(\varphi_{j+1}-\varphi) = 0$ and $S_0\varphi_0 = 0$, we have

$$\varphi_{j+1} - \varphi = S_0(T_0-\lambda_0 I)(\varphi_{j+1}-\varphi)$$
$$= \frac{1}{\lambda_{j+1}}\left[S_0(T_0-T)T(\varphi_j-\varphi) + (\lambda-\lambda_{j+1})(\varphi-\varphi_0)\right]$$
$$+ \frac{1}{\lambda_{j+1}^2}\Big[[\langle T^2(\varphi_j-\varphi),\varphi_0^*\rangle + \lambda(\lambda-\lambda_0) + \lambda_0(\lambda-\lambda_{j+1})]S_0T(\varphi_j-\varphi)$$
$$+ \lambda[\langle T^2(\varphi_j-\varphi),\varphi_0^*\rangle + \lambda(\lambda-\lambda_{j+1})]S_0(\varphi-\varphi_0)\Big] .$$

We see that a bound for each term has the common factor

$$c d_0^j \|(T-T_0)P\| \; \|(T-T_0)T\|^{j+1} ,$$

and the sum of the other factors is

$$\frac{2s_0}{|\lambda|} + \frac{c\|P\|}{|\lambda|}\Big[\frac{2}{|\lambda|} p_0\|T\| + \frac{4}{|\lambda|^2} p_0 s_0\|T\|^3 + \frac{4}{|\lambda|} s_0\|T\| + \frac{4|\lambda_0|}{|\lambda|^2} p_0 s_0\|T\|^2$$
$$+ \frac{4}{|\lambda|} p_0 s_0\|T\|^2 + 4p_0 s_0\|T\|\Big] ,$$

which equals d_0. This can be proved by using the induction hypothesis, (11.32), (11.33), and noting that since λ is semisimple, we have $TP = \lambda P$, so that $\|(T-T_0)P\| = \|(T-T_0)TP\|/|\lambda| \le \|(T-T_0)T\|$ $\|P\|/|\lambda|$. Thus, we have

$$\|\varphi-\varphi_{j+1}\| \le c d_0^{j+1} \|(T-T_0)P\| \; \|(T-T_0)T\|^{j+1} ,$$

which proves (i) with j replaced by $j + 1$. As a consequence,

$$|\lambda-\lambda_{j+2}| = |\langle T(\varphi-\varphi_{j+1}),\varphi_0^*\rangle| \le c d_0^{j+1} p_0\|T\| \; \|(T-T_0)P\| \; \|(T-T_0)T\|^{j+1} .$$

Also, since $\|(T-T_0)T\| \leq 1/d_0$, and $2cp_0\|T\| \, \|(T-T_0)P\| \leq |\lambda|$ we see that $|\lambda-\lambda_{j+2}| \leq |\lambda|/2$, so that $|\lambda_{j+2}| \geq |\lambda|/2$. This completes the induction proof of (i), (ii), (iii). The proof of (11.34) is complete if we note that

$$\lambda - \lambda_j = \langle T(\varphi-\varphi_{j-1}), \varphi_0^*\rangle = \langle (T-T_0)(\varphi-\varphi_{j-1}), \varphi_0^*\rangle \ .$$

It is easy to see that if $d_0\|(T-T_0)T\| < 1$, then $\lambda_j \to \lambda$ and $\varphi_j \to \varphi$. //

REMARK 11.11 It follows from Theorem 11.10 that the estimates for the eigenvalue approximation λ_j and the eigenvector approximation φ_j given by the scheme (11.31) are of the same order if $\|T-T_0\|$ and $\|(T-T_0)T\|$ are of the same order of magnitude. If, on the other hand, $\|T-T_0\|$ is not small but $\|(T-T_0)T\|$ is small, we have a better guaranteed accuracy for φ_j than for λ_j ; in particular the Rayleigh quotient $\lambda_1 = \langle T\varphi_0, \varphi_0^*\rangle$ may not improve upon λ_0 , while φ_1 may very well improve upon φ_0 .

We shall point out in Section 16 some practical situations where the conditions (11.32) and (11.33) are satisfied for $T_0 = T_n$, when (T_n) is an approximation of a *compact* operator T .

We remark that the iteration scheme (11.31) is considered in [DL] and is only slightly different from the <u>Ahués iteration scheme</u>

$$\varphi_j = \frac{T\varphi_{j-1}}{\lambda_j} + \frac{S_0T}{\lambda_j}\,[-T\varphi_{j-1} + \lambda_j\varphi_{j-1}]. \tag{11.35}$$

(See [C], (5.26) on p.26, and [A], DCB2 on p.149.) However, (11.31) sometimes gives much better numerical results. (See Tables 19.3, 19.4 and 19.5.)

We conclude this longish section by considering what is perhaps the most simple-minded and the most well-known iteration scheme for finding eigenelements of T .

Let φ be an eigenvector of T corresponding to a *nonzero* eigenvalue λ . If $x_0^* \in X^*$ is such that $\langle \varphi, x_0^* \rangle = 1$, then

$$\varphi = T\varphi \,/\, \langle T\varphi, x_0^* \rangle ,$$

i.e., φ is a fixed point of the function $G(x) = Tx/\langle Tx, \varphi_0^* \rangle$, assuming that the denominator $\langle Tx, \varphi_0^* \rangle$ does not vanish for x in an appropriate set.

Starting with some $x_0 \in X$, we can define the iteration scheme

$$x_j = Tx_{j-1} \,/\, \langle Tx_{j-1}, x_0^* \rangle , \; j = 1,2,\ldots ,$$

provided $\langle Tx_{j-1}, x_0^* \rangle \neq 0$; in that case it follows by induction on j that

$$x_j = T^j x_0 \,/\, \langle T^j x_0, x_0^* \rangle , \; j = 1,2,\ldots .$$

This is why the above iteration is known as the <u>power method</u>. The main limitation of this method is that for most starting vectors x_0 and x_0^* , the sequence (x_j) converges to an eigenvector φ of T corresponding to an eigenvalue of largest absolute value, whereas the iteration schemes developed earlier can approximate an intermediate eigenvalue of T , namely the one which is closest to the starting eigenvalue λ_0 .

We say that an isolated spectral value λ of T is <u>dominant</u> if it is the only spectral value of T satisfying $|\lambda| = r_\sigma(T)$. It follows by (8.1) that λ is the dominant spectral value of T if and only if $\bar{\lambda}$ is the dominant spectral value of T^* .

THEOREM 11.12 Assume that $0 \neq T \in BL(X)$ has the dominant spectral value λ, which is a pole of order ℓ of the resolvent operator $R(T,z)$. Let P and D be respectively the spectral projection and the quasi-nilpotent operator associated with T and λ.

Let $x_0 \in X$ and $x_0^* \in X^*$ be such that $\langle D^{\ell-1}x_0, x_0^*\rangle \neq 0$, where $D^0 = P$. Then for all large j, $\langle T^j x_0, x_0^*\rangle \neq 0$, so that

$$(11.36) \qquad x_j = Tx_{j-1} / \langle Tx_{j-1}, x_0^*\rangle = T^j x_0 / \langle T^j x_0, x_0^*\rangle$$

converges to the eigenvector $D^{\ell-1}x_0/\langle D^{\ell-1}x_0, x_0^*\rangle$ of T, and

$$(11.37) \qquad \lambda_j = \langle Tx_{j-1}, x_0^*\rangle = \langle T^j x_0, x_0^*\rangle / \langle T^{j-1}x_0, x_0^*\rangle$$

converges to λ.

Proof Since $D = (T-\lambda)P$, we have

$$T = TP + T(I-P) = \lambda P + D + T(I-P) .$$

Also, since P commutes with T,

$$T^j = [\lambda P + D]^j + [T(I-P)]^j , \quad j = 1,2,\ldots .$$

Let $Y = R(P)$ and $Z = Z(P)$, so that $X = Y \oplus Z$ and

$$T(I-P) = 0 \oplus T|_Z .$$

By the spectral decomposition theorem (Theorem 6.3), we see that

$$\sigma(T(I-P)) = \{0\} \cup \sigma(T|_Z)$$
$$= \{0\} \cup \{\mu \in \sigma(T) : \mu \neq \lambda\} .$$

Since $\lambda \neq 0$ is the dominant spectral value of T, we have

$$0 \leq r_\sigma(T(I-P)) < |\lambda| .$$

If we let

$$A = T(I-P) / \lambda ,$$

we see that $r_\sigma(A) < 1$ and hence $\|A^j\| \to 0$ as $j \to \infty$ by (5.8) or (5.10). Now, $T^j = \lambda^j \left[P + \binom{j}{1} \frac{D}{\lambda} + \dots + \binom{j}{\ell-1} \frac{D^{\ell-1}}{\lambda^{\ell-1}} + A^j\right]$, so that

$$T^j x_0 = \lambda^j \left[P x_0 + \binom{j}{1} \frac{D x_0}{\lambda} + \dots + \binom{j}{\ell-1} \frac{D^{\ell-1} x_0}{\lambda^{\ell-1}} + A^j x_0\right] ,$$

$$\langle T^j x_0, x_0^* \rangle = \lambda^j \left[\langle P x_0, x_0^* \rangle + \binom{j}{1} \frac{\langle D x_0, x_0^* \rangle}{\lambda} + \dots + \binom{j}{\ell-1} \frac{\langle D^{\ell-1} x_0, x_0^* \rangle}{\lambda^{\ell-1}} + \langle A^j x_0, x_0^* \rangle\right] .$$

But $A^j x_0 \to 0$, $\langle A^j x_0, x_0^* \rangle \to 0$ as $j \to \infty$, and $\binom{j}{\ell-1} \frac{\langle D^{\ell-1} x_0, x_0^* \rangle}{\lambda^{\ell-1}}$ is the dominating term in the expression for $\langle T^j x_0, x_0^* \rangle / \lambda^j$, as $j \to \infty$. Since $\langle D^{\ell-1} x_0, x_0^* \rangle \neq 0$, we see that $\langle T^j x_0, x_0^* \rangle \neq 0$ for all large j, and $x_j = \frac{T^j x_0}{\langle T^j x_0, x_0^* \rangle} \longrightarrow \frac{D^{\ell-1} x_0}{\langle D^{\ell-1} x_0, x_0^* \rangle}$ as $j \to \infty$. Since

$$\begin{aligned} T(D^{\ell-1} x_0) &= (T-\lambda I) D^{\ell-1} x_0 + \lambda D^{\ell-1} x_0 = (T-\lambda I) P D^{\ell-1} x_0 + \lambda D^{\ell-1} x_0 \\ &= D^\ell x_0 + \lambda D^{\ell-1} x_0 = \lambda D^{\ell-1} x_0 , \end{aligned}$$

we see that $D^{\ell-1} x_0 / \langle D^{\ell-1} x_0, x_0^* \rangle$ is an eigenvector of T corresponding to λ. Also, for all large j, it can be seen that

$$\lambda_j = \langle T x_{j-1}, x_0^* \rangle = \langle T^j x_0, x_0^* \rangle / \langle T^{j-1} x_0, x_0^* \rangle ,$$

where the numerator equals

$$\lambda^j \left[\langle P x_0, x_0^* \rangle + \binom{j}{1} \frac{\langle D x_0, x_0^* \rangle}{\lambda} + \dots + \binom{j}{\ell-1} \frac{\langle D^{\ell-1} x_0, x_0^* \rangle}{\lambda^{\ell-1}} + \langle A^j x_0, x_0^* \rangle\right] .$$

The denominator is obtained by replacing j by $j - 1$. Hence $\lambda_j \to \lambda$ as $j \to \infty$. //

REMARK 11.13 (a) It is significant to note that the dominant eigenvalue λ is assumed to be a pole of the resolvent operator, but it need not be of finite algebraic multiplicity. The only condition required of the starting vectors x_0 and x_0^* is that $\langle D^{\ell-1}x_0, x_0^*\rangle \neq 0$ when λ is a pole of order ℓ of $R(T,z)$. In case λ is a *simple* eigenvalue, we have $\ell = 1$, $D^{\ell-1} = P$, and $Px = \langle x, \varphi^*\rangle\varphi$, where φ (resp., φ^*) is an eigenvector of T (resp., T^*) corresponding to λ (resp., $\bar{\lambda}$) satisfying $\langle \varphi, \varphi^*\rangle = 1$. Thus, the condition $\langle D^{\ell-1}x_0, x_0^*\rangle \neq 0$ is equivalent to $\langle x_0, \varphi^*\rangle \neq 0$ and $\langle \varphi, x_0^*\rangle \neq 0$. Since an arbitrary choice of $x_0 \in X$ and $x_0^* \in X^*$ is most likely to satisfy these conditions, such a random choice is made in practice. A more appropriate procedure for the choice of x_0 and x_0^* is as follows. Let T_0 be a known 'nearby' operator having a simple eigenvalue λ_0 , and such that

$$r_\sigma(P_0(P-P_0)) < 1 ,$$

where P_0 is the spectral projection associated with T_0 and λ_0 . Then $P_0x = \langle x, \varphi_0^*\rangle\varphi_0$, where φ_0 (resp., φ_0^*) is an eigenvector of T_0 (resp., T_0^*) corresponding to λ_0 (resp., $\bar{\lambda}_0$). Now, in Lemma 9.5, letting $P = P_0$ and $Q = P$, we see that $P_0P|_{P_0(X)}$ is invertible. Since $\varphi_0 \in P_0(X)$ and $\varphi_0 \neq 0$, we have

$$0 \neq P_0P\varphi_0 = \langle P\varphi_0, \varphi_0^*\rangle\varphi_0 = \langle \varphi_0, \varphi^*\rangle\langle \varphi, \varphi_0^*\rangle = \langle P\varphi_0, \varphi_0^*\rangle .$$

This shows that we can choose $x_0 = \varphi_0$ and $x_0^* = \varphi_0^*$.

REMARK 11.14 While the power method is relatively simple to implement and the conditions on the starting vectors x_0 and x_0^* are not stringent, the main limitation of the power method is that it approximates only the *dominant* eigenvalue. If we replace the operator T by $T - z_0I$, where z_0 is a scalar, then the power method applied to $T - z_0 I$ will approximate an isolated eigenvalue λ of T which satisfies $|\lambda - z_0| > |\lambda' - z_0|$ for all $\lambda' \in \sigma(T)$, $\lambda' \neq \lambda$, provided such λ exists. However, the choice of such a scalar z_0 is difficult to make unless one has a good knowledge of the entire spectrum of T. Also, if $\sigma(T)$ has real spectrum and one wishes to use only real scalars z_0, then one can hope to approximate only the largest and the smallest eigenvalue of T in this manner.

If T is invertible, has a spectral value λ of the smallest modulus, and if λ is a pole of the resolvent operator, then the power method applied to T^{-1} (known as the <u>inverse power method</u>) will approximate this λ, because $1/\lambda$ is then the dominant eigenvalue of T^{-1}, and it is a pole of the resolvent operator $R(T^{-1},w) = -\frac{1}{w}TR(T,\frac{1}{w})$, $w \in \rho(T^{-1})$. More generally, let λ be an isolated spectral value of T which is a pole of $R(T,z)$, $z \in \rho(T)$. If we can find a scalar $z_0 \in \rho(T)$ such that $|\lambda - z_0| < |\lambda' - z_0|$ for every $\lambda' \in \sigma(T)$, $\lambda' \neq \lambda$, then the inverse power method applied to $T - z_0 I$ will approximate λ. The scalar z_0 is usually found as an initial approximation of λ by some other method: either as an eigenvalue of a nearby operator T_0, or by one of the methods described in Section 12.

We note that in the inverse power method with a shift z_0, we need not calculate $(T-z_0 I)^{-1}x$ for $x \in X$. It is only necessary to solve equations involving the operator T:

Let $x_0 \in X$, $x_0^* \in X^*$ be such that $\langle D^{\ell-1}x_0, x_0^*\rangle \neq 0$, where D is the nilpotent operator associated with T and λ. (See Problem 7.7.) For $j = 1,2,\ldots$, find $\tilde{x}_j \in X$ such that

(11.38) $$(T-z_0 I)\tilde{x}_j = x_{j-1},$$

and put $x_j = \dfrac{\tilde{x}_j}{\langle \tilde{x}_j, x_0^*\rangle}$. Then (x_j) converges to the eigenvector

$D^{\ell-1}x_0 \,/\, \langle D^{\ell-1}x_0, x_0^*\rangle$ of T, and if we let $\lambda_j = \langle Tx_{j-1}, x_0^*\rangle$, then $z_0 + 1/\lambda_j$ converges to the corresponding eigenvalue λ. This is known as the inverse iteration.

Instead of considering a fixed shift z_0, one can vary it at each step. Let X be a Hilbert space, and let $x_0 \neq 0$ be an approximation of an eigenvector x of T. Then the minimum residual property (8.9) of the Rayleigh quotient $q(x_0) = \langle Tx_0, x_0\rangle / \|x_0\|^2$ says that $q(x_0)$ is a judicious choice for an approximation for the corresponding eigenvalue. The inverse iteration principle then says that we should consider a shift by $q(x_0)$. Repetition of this process gives the Rayleigh quotient iteration: For $j = 1,2,\ldots,$ let

$$
\begin{aligned}
z_j &= q(x_{j-1}) , \\
(T - z_j I)\tilde{x}_j &= x_{j-1} \\
x_j &= \frac{\tilde{x}_j}{\langle \tilde{x}_j, x_0^*\rangle} .
\end{aligned}
\tag{11.39}
$$

Problems

11.1 ([LN], Proposition 3.1) Let $\psi_{(k)}$, $k = 1,2,\ldots,$ be defined by (11.16), and ϵ_0 be as in (11.29). Then

$$
\|\psi_{(k)}\| \le \begin{cases} a_k \eta_0 (\sqrt{\epsilon_0})^{k-1} , & k \text{ odd} \\ a_k \eta_0 \gamma_0 (\sqrt{\epsilon_0})^{k-2} , & k \text{ even} , \end{cases}
$$

$$
\|V_0 S_0 \psi_{(k)}\| \le \begin{cases} a_k \eta_0 (\sqrt{\epsilon_0})^{k} , & k \text{ even} \\ a_k \eta_0 \alpha_0 (\sqrt{\epsilon_0})^{k-1} , & k \text{ odd} . \end{cases}
$$

Hence the estimates in (11.30) can be deduced. Also,

$$
\begin{aligned}
|\lambda - \lambda_0| &\le 2\sqrt{\epsilon_0}(1+\sqrt{\epsilon_0})/(1+\alpha_0)s_0 , \\
\|\varphi - \varphi_0\| &\le \sqrt{\epsilon_0}(1+\sqrt{\epsilon_0})/(\alpha_0+\epsilon_0)p_0 .
\end{aligned}
$$

11.2 ([LN], Theorem 3.4.) Let λ_0, φ_0 (resp., λ, φ) be eigenelements of T_0 (resp., T) and let λ_0 be simple. Let $\delta_0 = \|[T-T_0)S_0]^2\|$, and assume that $\delta_0 + p_0(\alpha_0+\delta_0)\|\varphi-\varphi_0\| < 1$. Then the set

$$\tilde{\Lambda}_0 = \left\{ z \notin \mathbb{C} : |z-\lambda_0| < \frac{1-\delta_0-p_0(\alpha_0+\delta_0)\|\varphi-\varphi_0\|}{s_0(1+\alpha_0)} \right\}$$

contains no spectral value of T, except possibly λ. If $\lambda \in \tilde{\Lambda}_0$, then λ is simple. (Note that the set Λ_0 given by (11.22) is contained in $\tilde{\Lambda}_0$). Hence Theorem 11.8 can be improved as follows: Let $\sqrt{\epsilon_0} < 1/4$. Then

$$|\lambda-\lambda_0| \le 2\sqrt{\epsilon_0}\,\frac{1+\sqrt{\epsilon_0}}{s_0(1+\alpha_0)},$$

and there is no other spectral value of T in

$$\tilde{D}_0 = \left\{ z \in \mathbb{C} : |z-\lambda_0| < (1-2\sqrt{\epsilon_0})\,\frac{1+\sqrt{\epsilon_0}}{(1+\alpha_0)s_0} \right\}.$$

11.3 Let $0 \le \gamma_0 < 1/4$ and

$$\eta_0[g(\gamma_0)-1]/\gamma_0 \le r < (1-2\gamma_0)\eta_0/2\gamma_0^2 .$$

Then the map $\tilde{F}$ given by (11.13) is a contraction from $\{x \in X : \|x\| \le r\}$ onto itself, the constant of contraction being $2(\gamma_0 + \gamma_0^2 r/\eta_0)$. Consider the special case $\gamma_0 = (\sqrt{2}-1)/2$ and $r = 2\eta_0$ to obtain error bounds similar to (11.20) and (11.21).

11.4 Let ψ_j, $j = 1,2,\ldots$ be defined by (11.16), and let $\sqrt{\epsilon_0} < 1/4$, where ϵ_0 is defined by (11.29). Then $\psi_j \to \psi$, with

$$\|\psi\| \le 16(1+\alpha_0)\eta_0/5 ,$$

$$\|\psi-\psi_0\| \le c\eta_0\gamma_0 , \quad \|\psi-\psi_1\| \le c\eta_0\epsilon_0 ,$$

(the constant c depends on η_0, γ_0 and ϵ_0), and for $j = 2,3,\ldots$,

$$\|\psi-\psi_j\| \le \begin{cases} \|\psi-\psi_0\|(4\sqrt{\epsilon_0})^j & , \ j \text{ even} \\ \|\psi-\psi_1\|(4\sqrt{\epsilon_0})^{j-1} & , \ j \text{ odd} . \end{cases}$$

Hence the estimates in (11.30) can be deduced.

11.5 For both the iteration schemes (11.18) and (11.19), the first iterate φ_1 is given by $\varphi_0 - S_0T\varphi_0$ (and hence $\lambda_2 = \lambda_1 - \langle TS_0T\varphi_0, \varphi_0^* \rangle$, while $\varphi_2 = \varphi_1 - S_0T\varphi_1 + \lambda_1 S_0\varphi_1$ for (11.18), and $\varphi_2 = \varphi_1 - S_0T\varphi_1 + \lambda_2 S_0\varphi_1$ for (11.19).

Let $\lambda_0 \neq 0$ and $T_0V_0\varphi_0 = 0$. Then $\lambda_1 = \lambda_0$ and $\varphi_1 = T\varphi_0/\lambda_0$.

11.6 Let $P_0V_0P_0 = 0 = S_0V_0S_0$, and $\eta_0 p_0 s_0 \alpha_0 < 1/4$. Then for the iteration scheme (11.18), we have

$$|\lambda-\lambda_0| = |\lambda-\lambda_1| \le 2\eta_0 p_0 \alpha_0 , \ \|\varphi-\varphi_0\| \le 2\eta_0 s_0 ,$$

and for $j = 1,2,\ldots$,

$$|\lambda-\lambda_{2j}| = |\lambda-\lambda_{2j+1}| \le \eta_0 p_0 \alpha_0 (4\eta_0 p_0 s_0 \alpha_0)^j ,$$
$$\|\varphi-\varphi_{2j-1}\| = \|\varphi-\varphi_{2j}\| \le \eta_0 s_0 (4\eta_0 p_0 s_0 \alpha_0)^j .$$

Thus, the iterations converge if $\gamma_0 < 1/2$.

11.7 Let $T \in BL(X)$. Assume that $T_0 \in BL(X)$ has a simple eigenvalue λ_0 and let φ_0, φ_0^*, S_0 have usual meanings. For $j = 1,2,\ldots$, let

$$\lambda_j = \langle T\varphi_{j-1}, \varphi_0^* \rangle , \quad r_{j-1} = \lambda_j\varphi_{j-1} - T\varphi_{j-1} ,$$
$$(T_0-\lambda_0 I)u_j = r_{j-1} , \quad \varphi_j = \varphi_{j-1} + u_j .$$

If $\|(T-T_0)S_0\|$ and $\|(T-T_0)\varphi_0\| \, \|\varphi_0^*\| \, \|S_0\|$ are less than $1/4$, then $\varphi_j \to \varphi$, $\lambda_j \to \lambda$ such that $T\varphi = \lambda\varphi$, $\langle \varphi, \varphi_0^* \rangle = 1$. Morerover, λ is a simple eigenvalue of T and it is the nearest spectral point of T from λ_0. (Compare Problem 9.2.)