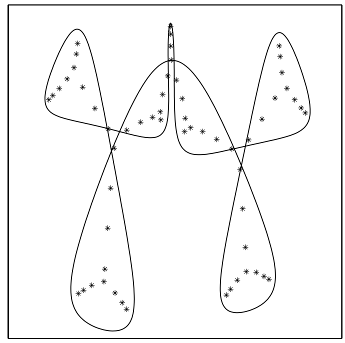

Chapter 7

Substitutes for the Spectrum

As will be seen in Chapter 11, the spectrum $\operatorname{sp} T_n(b)$ need not mimic $\operatorname{sp} T(b)$ as n goes to infinity. In contrast to this, pseudospectra and numerical ranges behave as nicely as we could ever expect. These sets are the concern of the present chapter.

7.1 Pseudospectra

For $\varepsilon > 0$, the *ε-pseudospectrum* of an operator $A \in \mathcal{B}(X)$ on a Banach space X is defined by

$$\operatorname{sp}_\varepsilon A = \left\{\lambda \in \mathbf{C} : \|(A - \lambda I)^{-1}\|_{\mathcal{B}(X)} \geq 1/\varepsilon\right\}. \tag{7.1}$$

Here we put $\|(A - \lambda I)^{-1}\|_{\mathcal{B}(X)} = \infty$ if $A - \lambda I$ is not invertible. Thus, the usual spectrum $\operatorname{sp} A$ is always a subset of $\operatorname{sp}_\varepsilon A$. Clearly, $\operatorname{sp}_\varepsilon A$ depends on X. If A acts on ℓ^p or ℓ^p_n, we denote the ε-pseudospectrum of A by $\operatorname{sp}^{(p)}_\varepsilon A$.

Definition (7.1) admits several modifications and generalizations. An important generalization, which is motivated by plenty of applications, is the so-called *structured pseudospectrum*. In this context we are given three operators $A, B, C \in \mathcal{B}(X)$ and we define

$$\operatorname{sp}^{B,C}_\varepsilon A = \operatorname{sp} A \cup \left\{\lambda \notin \operatorname{sp} A : \|C(A - \lambda I)^{-1} B\|_{\mathcal{B}(X)} \geq 1/\varepsilon\right\}. \tag{7.2}$$

Evidently, $\operatorname{sp}_\varepsilon A$ is nothing but $\operatorname{sp}^{I,I}_\varepsilon A$. Furthermore, some authors prefer (7.1) and (7.2) with "$\geq$" replaced by "$>$". Theorem 7.2 and its Corollary 7.3 will give alternative descriptions of the sets (7.1) and (7.2) in the Hilbert space case.

Lemma 7.1. *If M and N are linear operators, then $I + MN$ is invertible if and only if $I + NM$ is invertible, in which case*

$$(I + MN)^{-1} = I - M(I + NM)^{-1}N.$$

Proof. Simply check that

$$(I + MN)(I - M(I + NM)^{-1}N) = (I - M(I + NM)^{-1}N)(I + MN) = I. \qquad \square$$

Theorem 7.2. *If H is a Hilbert space, $A, B, C \in \mathcal{B}(H)$, and $\varepsilon > 0$, then*

$$\bigcup_{\|K\|\le\varepsilon} \mathrm{sp}\,(A + BKC) = \mathrm{sp}\,A \cup \left\{\lambda \notin \mathrm{sp}\,A : \|C(A-\lambda I)^{-1}B\| \ge 1/\varepsilon\right\}, \tag{7.3}$$

$$\bigcup_{\|K\|<\varepsilon} \mathrm{sp}\,(A + BKC) = \mathrm{sp}\,A \cup \left\{\lambda \notin \mathrm{sp}\,A : \|C(A-\lambda I)^{-1}B\| > 1/\varepsilon\right\}, \tag{7.4}$$

the union taken over all $K \in \mathcal{B}(H)$ with the given norm constraint.

Proof. Suppose A is invertible and $\|CA^{-1}B\| \le 1/\varepsilon$ (respectively, $\|CA^{-1}B\| < 1/\varepsilon$) and $\|K\| < \varepsilon$ (respectively, $\|K\| \le \varepsilon$). In either case, $\|CA^{-1}BK\| < 1$, which implies that $I + CA^{-1}BK$ is invertible. Using Lemma 7.1 with $N = C$ and $M = A^{-1}BK$, we see that $I + A^{-1}BKC$ and thus also $A + BKC = A(I + A^{-1}BKC)$ are invertible. This proves that the left-hand sides of (7.3) and (7.4) are contained in the corresponding right-hand sides.

To show that the right-hand sides of (7.3) and (7.4) are subsets of the corresponding left-hand sides, it suffices to prove that if A is invertible and $\|CA^{-1}B\| = 1/\delta$, then there exists an operator K such that $\|K\| = \delta$ and $A + BKC$ is not invertible. So assume that A is invertible and $\|CA^{-1}B\| = 1/\delta$. Clearly, $A + BKC = A(I + A^{-1}BKC)$ is invertible if and only if so is $I + A^{-1}BKC$, which, by Lemma 7.1 with $M = A^{-1}B$ and $N = KC$, is equivalent to the invertibility of $I + KCA^{-1}B$. Abbreviate $CA^{-1}B$ to S and put $K = -\delta^2 S^*$. Then $\|K\| = \delta$ and $I + KCA^{-1}B = I - \delta^2 S^* S$. The spectral radius of the positive semi-definite selfadjoint operator S^*S coincides with its norm, that is, with $\|S^*S\| = \|S\|^2 = 1/\delta^2$. It follows that $1/\delta^2 \in \mathrm{sp}\,S^*S$. The spectral mapping theorem therefore implies that $0 \in \mathrm{sp}\,(I - \delta^2 S^* S)$. □

Corollary 7.3. *Let H be a Hilbert space and $A \in \mathcal{B}(H)$. Then for every $\varepsilon > 0$,*

$$\mathrm{sp}_\varepsilon A = \bigcup_{\|K\|\le\varepsilon} \mathrm{sp}\,(A + K),$$

the union taken over all $K \in \mathcal{B}(H)$ of norm at most ε.

Proof. This is (7.3) with $B = C = I$. □

The following result sharpens (7.4).

Theorem 7.4. *Let X be a Banach space, let A, B, C be operators in $\mathcal{B}(X)$, and let $\varepsilon > 0$. Then*

$$\mathrm{sp}\,A \cup \{\lambda \notin \mathrm{sp}\,A : \|C(A-\lambda I)^{-1}B\| > 1/\varepsilon\} \tag{7.5}$$

$$= \bigcup_{\|K\|<\varepsilon} \mathrm{sp}\,(A + BKC) \tag{7.6}$$

$$= \bigcup_{\|K\|<\varepsilon,\ \mathrm{rank}\,K=1} \mathrm{sp}\,(A + BKC). \tag{7.7}$$

Proof. The first part of the proof of Theorem 7.2 can literally be used to show that (7.6) is a subset of (7.5).

It remains to prove that (7.5) is contained in (7.7). This will follow as soon as we have shown that if A is invertible and $\|CA^{-1}B\| > 1/\varepsilon$, then there exists a rank-one operator K such that $\|K\| < \varepsilon$ and $A + BKC$ is not invertible. So assume A is invertible and $\|CA^{-1}B\| > 1/\varepsilon$. Then we can find a $u \in X$ such that $\|u\| = 1$ and $\|CA^{-1}Bu\| > 1/\varepsilon$. Thus, $\|CA^{-1}Bu\| = 1/\delta$ with $\delta < \varepsilon$. By the Hahn-Banach theorem, there is a functional $\varphi \in X^*$ such that $\|\varphi\| = 1$ and $\varphi(CA^{-1}Bu) = \|CA^{-1}Bu\| = 1/\delta$. Let $K \in \mathcal{B}(X)$ be the rank-one operator defined by $Kx = -\delta\varphi(x)u$. Clearly, $\|K\| \le \delta < \varepsilon$. Furthermore,

$$\begin{aligned} BKCA^{-1}Bu &= B(-\delta\varphi(CA^{-1}Bu)u) \\ &= -\delta\varphi(CA^{-1}Bu)Bu = -\delta(1/\delta)Bu = -Bu. \end{aligned} \tag{7.8}$$

Put $y = A^{-1}Bu$. If $y = 0$, then $CA^{-1}Bu = Cy = 0$, which contradicts the assumption $\|CA^{-1}Bu\| = 1/\delta > 0$. Consequently, $y \neq 0$. From (7.8) we see that $BKCy = -Bu = -Ay$, whence $(A + BKC)y = 0$. This implies that $A + BKC$ is not invertible. □

We have not been able to prove Theorem 7.4 with strict inequalities replaced by nonstrict inequalities. However, one can show that if X is a Banach space, $A, B, C \in \mathcal{B}(X)$, $\varepsilon > 0$, and at least one of the operators B or C is compact, then

$$\begin{aligned} &\operatorname{sp} A \cup \{\lambda \notin \operatorname{sp} A : \|C(A - \lambda I)^{-1}B\| \ge 1/\varepsilon\} \\ &= \bigcup_{\|K\| \le \varepsilon} \operatorname{sp}(A + BKC) = \bigcup_{\|K\| \le \varepsilon,\ \operatorname{rank} K = 1} \operatorname{sp}(A + BKC). \end{aligned}$$

To see this, assume that A is invertible and $\|CA^{-1}B\| = 1/\delta$. Since $CA^{-1}B$ is compact, there exists a $u \in X$ such that $\|u\| = 1$ and $\|CA^{-1}Bu\| = 1/\delta$. The rest of the proof is as in the proof of Theorem 7.4.

7.2 Norm of the Resolvent

In this section we show that the norm of the resolvent of a bounded operator on ℓ^p $(1 < p < \infty)$ cannot be locally constant. It should be noted that such a result is not true for arbitrary analytic operator-valued functions. To see this, consider the function

$$A : \mathbf{C} \to \mathcal{B}(\ell_2^p),\ \lambda \mapsto \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}.$$

Obviously, $\|A(\lambda)\|_p = \max(|\lambda|, 1)$ and thus $\|A(\lambda)\|_p = 1$ for all λ in the unit disk.

Theorem 7.5 (Daniluk). *Let H be a Hilbert space and $A \in \mathcal{B}(H)$. Suppose that $A - \lambda I$ is invertible for all λ in some open subset U of $\mathbf{C}$. If there is an $M < \infty$ such that $\|(A - \lambda I)^{-1}\| \le M < \infty$ for all $\lambda \in U$, then $\|(A - \lambda I)^{-1}\| < M$ for all $\lambda \in U$.*

Proof. A little thought reveals that what we must show is the following: If U is an open subset of $\mathbf{C}$ containing the origin and $\|(A - \lambda I)^{-1}\| \le M$ for all $\lambda \in U$, then $\|A^{-1}\| < M$.

To prove this, assume the contrary, i.e., let $\|A^{-1}\| = M$. There is a sufficiently small $r > 0$ such that

$$(A - \lambda I)^{-1} = \sum_{j=0}^{\infty} \lambda^j A^{-j-1} \quad \text{for} \quad |\lambda| = r.$$

Given $f \in H$, we therefore get

$$\|(A - \lambda I)^{-1} f\|^2 = \sum_{j,k \geq 0} \lambda^j \overline{\lambda}^k (A^{-j-1} f, A^{-k-1} f)$$

whenever $\lambda = re^{i\varphi}$. Integrating the last equality we obtain

$$\frac{1}{2\pi} \int_0^{2\pi} \|(A - re^{i\varphi} I)^{-1} f\|^2 d\varphi = \sum_{j=0}^{\infty} r^{2j} \|A^{-j-1} f\|^2,$$

and since $\|(A - re^{i\varphi} I)^{-1} f\| \leq M\|f\|$, it follows that $\|A^{-1} f\|^2 + r^2 \|A^{-2} f\|^2 \leq M^2 \|f\|^2$. Now pick an arbitrary $\varepsilon > 0$ and choose an $f_\varepsilon \in H$ such that $\|f_\varepsilon\| = 1$ and $\|A^{-1} f_\varepsilon\|^2 > M^2 - \varepsilon$. Then $M^2 - \varepsilon + r^2 \|A^{-2} f_\varepsilon\|^2 < M^2$, whence $1 = \|f_\varepsilon\|^2 \leq \|A^2\|^2 \|A^{-2} f_\varepsilon\|^2 < \varepsilon r^{-2} \|A^2\|^2$, which is impossible if $\varepsilon > 0$ is small enough. □

The following result extends Theorem 7.5 to operators on ℓ^p for $1 < p < \infty$.

Theorem 7.6. *Let $1 < p < \infty$ and let $A \in \mathcal{B}(\ell^p)$. If $A - \lambda I$ is invertible for all λ in some open set $U \subset \mathbf{C}$ and $\|(A - \lambda I)^{-1}\|_p \leq M < \infty$ for $\lambda \in U$, then $\|(A - \lambda I)^{-1}\|_p < M$ for $\lambda \in U$.*

Proof. We may without loss of generality assume that $p \geq 2$; otherwise we can pass to adjoint operators. Again it suffices to show that $\|A^{-1}\|_p < M$ provided U contains the origin. Assume the contrary, that is, let $\|A^{-1}\|_p = M$. There is an $r > 0$ such that

$$(A - \lambda I)^{-1} = \sum_{j=0}^{\infty} \lambda^j A^{-j-1}$$

for all $\lambda = re^{i\varphi}$. Hence, for every $x \in \ell^p$,

$$\begin{aligned} \|(A - \lambda I)^{-1} x\|_p^p &= \sum_{n=0}^{\infty} \left| \sum_{j=0}^{\infty} \lambda^j (A^{-j-1} x)_n \right|^p = \sum_{n=0}^{\infty} \left| \sum_{j=0}^{\infty} r^j e^{ij\varphi} (A^{-j-1} x)_n \right|^{2p/2} \\ &= \sum_{n=0}^{\infty} \left| \left(\sum_{j=0}^{\infty} r^j e^{ij\varphi} (A^{-j-1} x)_n \right) \left(\sum_{k=0}^{\infty} r^k e^{-ik\varphi} \overline{(A^{-k-1} x)_n} \right) \right|^{p/2} \\ &= \sum_{n=0}^{\infty} \left| C(r, n) + \sum_{\ell=1}^{\infty} B_\ell(r, \varphi, n) \right|^{p/2}, \end{aligned} \tag{7.9}$$

where

$$C(r,n) = \sum_{j=0}^{\infty} r^{2j} |(A^{-j-1}x)_n|^2,$$

$$B_\ell(r,\varphi,n) = 2\sum_{k=0}^{\infty} r^{\ell+2k} \operatorname{Re}\left(e^{i\ell\varphi}(A^{-\ell-k-1}x)_n \overline{(A^{-k-1}x)_n}\right).$$

For $m = 0, 1, 2, \dots$, put

$$I_m(r,\varphi) = \sum_{n=0}^{\infty} \left| C(r,n) + \sum_{\ell=1}^{\infty} B_{2^m\ell}(r,\varphi,n) \right|^{p/2}.$$

Clearly,

$$\lim_{m\to\infty} I_m(r,\varphi) = \sum_{n=0}^{\infty} |C(r,n)|^{p/2}. \tag{7.10}$$

We now apply the inequality

$$|a|^{p/2} \le \frac{1}{2}\left(|a+b|^{p/2} + |a-b|^{p/2}\right) \tag{7.11}$$

to

$$a = C(r,n) + \sum_{\ell=1}^{\infty} B_{2\ell}(r,\varphi,n), \quad b = \sum_{\ell=1}^{\infty} B_{2\ell-1}(r,\varphi,n)$$

and sum up the results for $n = 0, 1, 2, \dots$. Taking into account that $\sum_{n=0}^{\infty} |a-b|^{p/2}$ is nothing but $I_0(r,\varphi+\pi)$, we get

$$I_1(r,\varphi) \le \frac{1}{2}\Big(I_0(r,\varphi) + I_0(r,\varphi+\pi)\Big). \tag{7.12}$$

Letting

$$a = C(r,n) + \sum_{\ell=1}^{\infty} B_{4\ell}(r,\varphi,n), \quad b = \sum_{\ell=1}^{\infty} B_{4\ell-2}(r,\varphi,n)$$

in (7.11), we analogously obtain that

$$I_2(r,\varphi) \le \frac{1}{2}\Big(I_1(r,\varphi) + I_1(r,\varphi+\pi/2)\Big). \tag{7.13}$$

Combining (7.12) and (7.13) we arrive at the inequality

$$\begin{aligned} I_2(r,\varphi) &\le \frac{1}{2}\Big(I_1(r,\varphi) + I_1(r,\varphi+\pi/2)\Big) \\ &\le \frac{1}{4}\Big(I_0(r,\varphi) + I_0(r,\varphi+\pi/2) + I_0(r,\varphi+\pi) + I_0(r,\varphi+3\pi/2)\Big). \end{aligned} \tag{7.14}$$

In the same way we see that

$$I_3(r,\varphi) \le \frac{1}{2}\Big(I_2(r,\varphi) + I_2(r,\varphi+\pi/4)\Big),$$

which together with (7.14) gives

$$I_3(r,\varphi) \le \frac{1}{8}\sum_{k=0}^{7} I_0\left(r,\varphi+\frac{k\pi}{4}\right).$$

Continuing this procedure we get

$$I_m(r,\varphi) \le \frac{1}{2^m}\sum_{k=0}^{2^m-1} I_0\left(r,\varphi+\frac{k\pi}{2^{m-1}}\right) \tag{7.15}$$

for every $m \ge 0$.

Now put $\varphi = 0$ in (7.15) and pass to the limit $m \to \infty$. The limit of the left-hand side is given by (7.10). The right-hand side is an integral sum and hence

$$\lim_{m\to\infty} \frac{1}{2^m}\sum_{k=0}^{2^m-1} I_0\left(r,\frac{k\pi}{2^{m-1}}\right) = \lim_{m\to\infty}\frac{1}{2\pi}\frac{2\pi}{2^m}\sum_{k=0}^{\infty} I_0\left(r,\frac{2k\pi}{2^m}\right) = \frac{1}{2\pi}\int_0^{2\pi} I_0(r,\varphi)d\varphi.$$

Thus,

$$\sum_{n=0}^{\infty} |\,C(r,n)|^{p/2} \le \frac{1}{2\pi}\int_0^{2\pi} I_0(r,\varphi)d\varphi.$$

Since $\|(A-\lambda I)^{-1}\|_p \le M$, we have $I_0(r,\varphi) = \|(A-re^{i\varphi}I)^{-1}x\|_p^p \le M^p\|x\|_p^p$. Consequently,

$$\sum_{n=0}^{\infty} |C(r,n)|^{p/2} \le M^p\|x\|_p^p. \tag{7.16}$$

Because

$$\sum_{n=0}^{\infty}|C(r,n)|^{p/2} \ge \sum_{n=0}^{\infty}\left|\,\left|(A^{-1}x)_n\right|^2 + r^2\left|(A^{-2}x)_n\right|^2\,\right|^{p/2}$$
$$\ge \sum_{n=0}^{\infty}\left|(A^{-1}x)_n\right|^p + r^p\sum_{n=0}^{\infty}\left|(A^{-2}x)_n\right|^p = \|A^{-1}x\|_p^p + r^p\|A^{-2}x\|_p^p$$

(here we used the inequality $(|a|+|b|)^{p/2} \ge |a|^{p/2}+|b|^{p/2}$), we deduce from (7.16) that

$$\|A^{-1}x\|_p^p + r^p\|A^{-2}x\|_p^p \le M^p\|x\|_p^p. \tag{7.17}$$

Finally, let $\varepsilon > 0$ and choose $x_\varepsilon \in \ell^p$ so that $\|x_\varepsilon\|_p = 1$ and $\|A^{-1}x\|_p^p > M^p - \varepsilon$. Then (7.17) yields $M^p - \varepsilon + r^p\|A^{-2}x_\varepsilon\|_p^p < M^p$, and this implies that

$$1 = \|x_\varepsilon\|_p^p \le \|A^2\|_p^p\|A^{-2}x_\varepsilon\|_p^p < \varepsilon r^{-p}\|A^2\|_p^p.$$

This inequality is impossible if $\varepsilon > 0$ is sufficiently small. Thus, our assumption $\|A^{-1}\|_p = M$ must be false. $\square$

7.3 Limits of Pseudospectra

Let $\{M_n\}_{n=1}^{\infty}$ be a sequence of sets $M_n \subset \mathbf{C}$. We define

$$\liminf_{n\to\infty} M_n$$

as the set of all $\lambda \in \mathbf{C}$ for which there are $\lambda_1 \in M_1, \lambda_2 \in M_2, \ldots$ such that $\lambda_n \to \lambda$, and we let

$$\limsup_{n\to\infty} M_n$$

denote the set of all $\lambda \in \mathbf{C}$ for which there exist $n_1 < n_2 < \cdots$ and $\lambda_{n_k} \in M_{n_k}$ such that $\lambda_{n_k} \to \lambda$. In other words, $\lambda \in \liminf M_n$ if and only if λ is the limit of some sequence $\{\lambda_n\}_{n=1}^{\infty}$ with $\lambda_n \in M_n$, while $\lambda \in \limsup M_n$ if and only if λ is a partial limit of such a sequence. We remark that if M and the members of the sequence $\{M_n\}$ are nonempty compact subsets of $\mathbf{C}$, then

$$\liminf_{n\to\infty} M_n = \limsup_{n\to\infty} M_n = M$$

if and only if M_n converges to M in the Hausdorff metric, which means that $d(M_n, M) \to 0$ with

$$d(A, B) := \max\left(\max_{a\in A} \operatorname{dist}(a, B), \max_{b\in B} \operatorname{dist}(b, A)\right).$$

This result is due to Hausdorff. Proofs can be found in [149, Sections 3.1.1 and 3.1.2] and [153, Section 2.8].

Theorem 7.7. *Let b be a Laurent polynomial. Then for every $\varepsilon > 0$ and every $p \in (1, \infty)$,*

$$\liminf_{n\to\infty} \operatorname{sp}_\varepsilon^{(p)} T_n(b) = \limsup_{n\to\infty} \operatorname{sp}_\varepsilon^{(p)} T_n(b) = \operatorname{sp}_\varepsilon^{(p)} T(b) \cup \operatorname{sp}_\varepsilon^{(p)} T(\widetilde{b}). \tag{7.18}$$

Proof. We first show that

$$\operatorname{sp}_\varepsilon^{(p)} T(b) \subset \liminf_{n\to\infty} \operatorname{sp}_\varepsilon^{(p)} T_n(b). \tag{7.19}$$

If $\lambda \in \operatorname{sp} T(b)$, then $\|T_n^{-1}(b-\lambda)\|_p \to \infty$ by virtue of Lemma 3.4. Thus, we have $\|T_n^{-1}(b-\lambda)\|_p \geq 1/\varepsilon$ for all $n \geq n_0$, which implies that $\lambda \in \operatorname{sp}_\varepsilon^{(p)} T_n(b)$ for all $n \geq n_0$. Consequently, λ belongs to $\liminf \operatorname{sp}_\varepsilon^{(p)} T_n(b)$.

Now suppose that $\lambda \in \operatorname{sp}_\varepsilon^{(p)} T(b) \setminus \operatorname{sp} T(b)$. Then $\|T^{-1}(b-\lambda)\|_p \geq 1/\varepsilon$. Let $U \subset \mathbf{C}$ be any open neighborhood of λ. From Theorem 7.6 we deduce that there is a point $\mu \in U$ such that $\|T^{-1}(b-\mu)\|_p > 1/\varepsilon$. Hence, we can find a natural number k_0 such that

$$\|T^{-1}(b-\mu)\|_p \geq \frac{1}{\varepsilon - 1/k} \quad \text{for all } \ k \geq k_0.$$

As U was arbitrary, it follows that there exists a sequence $\mu_1, \mu_2, \ldots$ such that $\mu_k \in \operatorname{sp}_{\varepsilon-1/k}^{(p)} T(b)$ and $\mu_k \to \lambda$. For every invertible operator $A \in \mathcal{B}(\ell^p)$,

$$\|A^{-1}\|_p = \sup_{x\neq 0} \frac{\|A^{-1}x\|_p}{\|x\|_p} = \sup_{y\neq 0} \frac{\|y\|_p}{\|Ay\|_p} = \left(\inf_{y\neq 0} \frac{\|Ay\|_p}{\|y\|_p}\right)^{-1}. \tag{7.20}$$

Since $\|T^{-1}(b-\mu_k)\|_p \ge 1/(\varepsilon - 1/k)$, it results that

$$\inf_{\|y\|_p=1} \|T(b-\mu_k)y\|_p \le \varepsilon - 1/k.$$

Thus, there are $y_k \in \ell^p$ such that $\|y_k\|_p = 1$ and $\|T(b-\mu_k)y_k\|_p < \varepsilon - 1/(2k)$. Clearly, $\|T_n(b-\mu_k)P_n y_k\|_p \to \|T(b-\mu_k)y_k\|_p$ and $\|P_n y_k\|_p \to \|y_k\|_p = 1$ as $n \to \infty$. Hence,

$$\frac{\|T_n(b-\mu_k)P_n y_k\|_p}{\|P_n y_k\|_p} < \varepsilon - 1/(3k)$$

for all $n > n_0(k)$. Again invoking (7.20) we see that

$$\|T_n^{-1}(b-\mu_k)\|_p > (\varepsilon - 1/(3k))^{-1} > 1/\varepsilon$$

and thus $\mu_k \in \mathrm{sp}_\varepsilon^{(p)} T_n(b)$ for all $n > n_0(k)$. This implies that $\lambda = \lim \mu_k$ belongs to $\liminf \mathrm{sp}_\varepsilon^{(p)} T_n(b)$. At this point the proof of (7.19) is complete.

Repeating the above reasoning with $\widetilde{b}$ in place of b we get the inclusion

$$\mathrm{sp}_\varepsilon^{(p)} T(\widetilde{b}) \subset \liminf_{n\to\infty} \mathrm{sp}_\varepsilon^{(p)} T_n(\widetilde{b}). \tag{7.21}$$

As $T_n(\widetilde{b}-\lambda) = W_n T_n(b-\lambda) W_n$ and W_n is an isometry of ℓ_n^p, it is clear that $\mathrm{sp}_\varepsilon^{(p)} T_n(\widetilde{b}) = \mathrm{sp}_\varepsilon^{(p)} T_n(b)$. Thus, in (7.21) we may replace the $T_n(\widetilde{b})$ on the right by $T_n(b)$, which in conjunction with (7.19) proves that $\mathrm{sp}_\varepsilon^{(p)} T(b) \cup \mathrm{sp}_\varepsilon^{(p)} T(\widetilde{b})$ is contained in $\liminf \mathrm{sp}_\varepsilon^{(p)} T_n(b)$.

We are left to prove the inclusion

$$\limsup_{n\to\infty} \mathrm{sp}_\varepsilon^{(p)} T_n(b) \subset \mathrm{sp}_\varepsilon^{(p)} T(b) \cup \mathrm{sp}_\varepsilon^{(p)} T(\widetilde{b}). \tag{7.22}$$

So let $\lambda \notin \mathrm{sp}_\varepsilon^{(p)} T(b) \cup \mathrm{sp}_\varepsilon^{(p)} T(\widetilde{b})$. Then $\|T^{-1}(b-\lambda)\|_p < 1/\varepsilon$ and $\|T^{-1}(\widetilde{b}-\lambda)\|_p < 1/\varepsilon$, whence, by Theorem 6.3,

$$\|T_n^{-1}(b-\lambda)\|_p < 1/\varepsilon - \delta < 1/\varepsilon \quad \text{for all } n \ge n_0 \tag{7.23}$$

with some $\delta > 0$. If $|\mu - \lambda|$ is sufficiently small, then $T_n(b-\mu)$ is invertible together with $T_n(b-\lambda)$, and we have, from the first resolvent identity,

$$\|T_n^{-1}(b-\mu)\|_p \le \frac{\|T_n^{-1}(b-\lambda)\|_p}{1 - |\mu-\lambda|\,\|T_n^{-1}(b-\lambda)\|_p}. \tag{7.24}$$

Let $|\mu - \lambda| < \varepsilon\delta(1/\varepsilon - \delta)^{-1}$. In this case (7.23) and (7.24) give

$$\|T_n^{-1}(b-\mu)\|_p < \frac{1/\varepsilon - \delta}{1 - \varepsilon\delta(1/\varepsilon-\delta)^{-1}(1/\varepsilon-\delta)} = \frac{1}{\varepsilon}.$$

Thus, $\mu \notin \mathrm{sp}_\varepsilon^{(p)} T_n(b)$ for $n \ge n_0$. This shows that λ cannot belong to the left-hand side of (7.22). □

Corollary 7.8. *If b is a Laurent polynomial and $\varepsilon > 0$, then*

$$\liminf_{n\to\infty} \mathrm{sp}_\varepsilon^{(2)} T_n(b) = \limsup_{n\to\infty} \mathrm{sp}_\varepsilon^{(2)} T_n(b) = \mathrm{sp}_\varepsilon^{(2)} T(b).$$

Proof. Since $T(\widetilde{b})$ is simply the transpose of $T(b)$, the two norms $\|T^{-1}(b-\lambda)\|_2$ and $\|T^{-1}(\widetilde{b}-\lambda)\|_2$ coincide. Therefore $\mathrm{sp}_\varepsilon^{(2)} T(b) = \mathrm{sp}_\varepsilon^{(2)} T(\widetilde{b})$. The assertion is now immediate from Theorem 7.7. □

Figures 7.1 and 7.2 show an example with the symbol $b(t) = -(6-13i)t + (5-4i)t^2 + 3t^3 - (4+3i)t^{-2} + 3t^{-3}$.

7.4 Pseudospectra of Infinite Toeplitz Matrices

For every operator $A \in \mathcal{B}(\ell^2)$ the inequality

$$1/\mathrm{dist}\,(\lambda, \mathrm{sp}\,A) \le \|(A-\lambda I)^{-1}\|_2 \tag{7.25}$$

holds. This implies that $\|(A-\lambda I)^{-1}\|_2 \ge 1/\varepsilon$ whenever $\mathrm{dist}\,(\lambda, \mathrm{sp}\,A) \le \varepsilon$ and hence yields the universal lower estimate

$$\mathrm{sp}\,A + \varepsilon\,\overline{\mathbf{D}} \subset \mathrm{sp}_\varepsilon^{(2)} A.$$

For Toeplitz operators, Theorem 4.29 gives

$$\|(T(a)-\lambda I)^{-1}\|_2 = \|T^{-1}(a-\lambda)\|_2 \le 1/\mathrm{dist}\,(\lambda, \mathrm{conv}\,\mathcal{R}(a)). \tag{7.26}$$

Consequently, if $\mathrm{dist}\,(\lambda, \mathrm{conv}\,\mathcal{R}(a)) > \varepsilon$ then $\|T^{-1}(a-\lambda)\|_2 < 1/\varepsilon$ and λ cannot belong to $\mathrm{sp}_\varepsilon^{(2)} T(a)$. We therefore arrive at the upper estimate

$$\mathrm{sp}_\varepsilon^{(2)} T(a) \subset \mathrm{conv}\,\mathcal{R}(a) + \varepsilon\,\overline{\mathbf{D}}.$$

Given $a \in W$, let $V(a)$ be the set of all $\lambda \in \mathbf{C} \setminus \mathrm{sp}\,T(a)$ for which

$$\mathrm{dist}\,(\lambda, \mathrm{sp}\,T(a)) = \mathrm{dist}\,(\lambda, \mathrm{conv}\,\mathcal{R}(a)).$$

From (7.25) and (7.26) we infer that if $\lambda \in V(a)$, then

$$\|T^{-1}(a-\lambda)\|_2 = 1/\mathrm{dist}\,(\lambda, \mathrm{conv}\,\mathcal{R}(a))$$

and hence in $V(a)$ the level curves $\|T^{-1}(a-\lambda)\|_2 = 1/\varepsilon$ coincide with the curves $\mathrm{dist}\,(\lambda, \mathrm{conv}\,\mathcal{R}(a)) = \varepsilon$. If $V(a) = \mathbf{C} \setminus \mathrm{sp}\,T(a)$, or equivalently, if $\mathrm{sp}\,T(a)$ is a convex set, then

$$\mathrm{sp}_\varepsilon^{(2)} T(a) = \mathrm{conv}\,\mathcal{R}(a) + \varepsilon\,\overline{\mathbf{D}}. \tag{7.27}$$

Equality (7.27) is particularly true for tridiagonal Toeplitz matrices, in which case $\mathrm{sp}\,T(a) = \mathrm{conv}\,\mathcal{R}(a)$ is an ellipse.

7.5 Numerical Range

Let X be a Banach space and put

$$\Pi(X) = \{(f, x) \in X^* \times X : \|f\| = 1, \|x\| = 1, f(x) = 1\}.$$

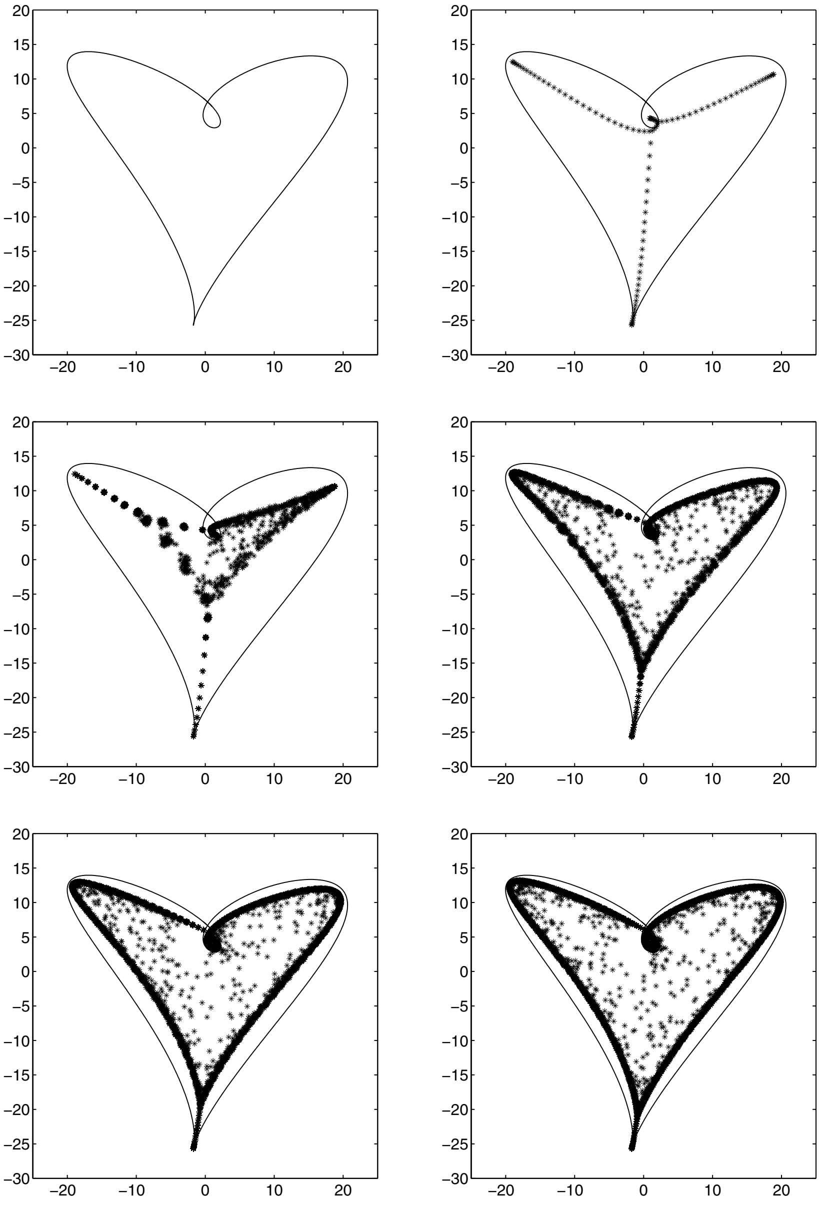

Figure 7.1. *In the two top pictures we see* $b(\mathbf{T})$ *(left) and* $b(\mathbf{T})$ *together with the* 100 *eigenvalues of* $T_{100}(b)$ *(right). The other four pictures indicate* $\mathrm{sp}_\varepsilon^{(2)} T_n(b)$ *for* $\varepsilon = 1/100$ *and* $n = 50, 100, 150, 200$*. Each picture shows the superposition of the spectra* $\mathrm{sp}\,(T_n(b) + E)$ *for* 50 *randomly chosen matrices* E *with* $\|E\|_2 = \varepsilon$.

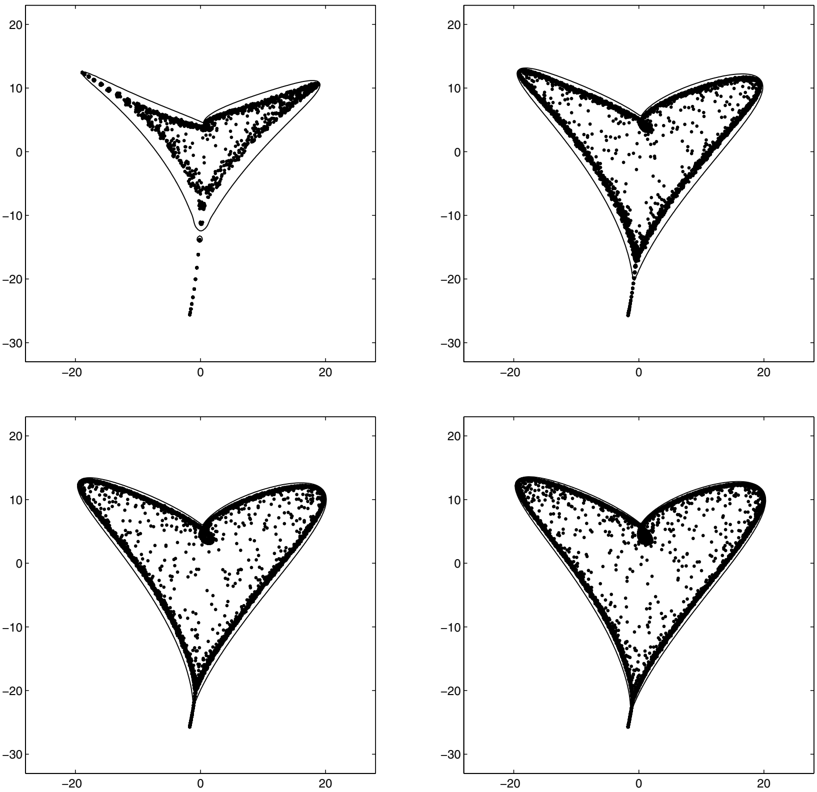

Figure 7.2. *These pictures were done by Mark Embree. In contrast to the lower four pictures of Figure* 7.1*, the solid curves are the boundaries of the pseudospectra* $\mathrm{sp}_\varepsilon^{(2)} T_n(b)$ *for* $\varepsilon = 1/100$ *and* $n = 50, 100, 150, 200$*. These curves were determined with the help of Tom Wright's package* [299].

The (spatial) *numerical range* (=*field of values*) $\mathcal{H}_X(A)$ of an operator $A \in \mathcal{B}(X)$ is defined as

$$\mathcal{H}_X(A) = \{f(Ax) : (f, x) \in \Pi(X)\}.$$

If $X = H$ is a Hilbert space, we can identify the dual space X^* with H and, accordingly,

$$\Pi(H) = \{(y, x) \in H \times H : \|y\| = 1, \|x\| = 1, (y, x) = 1\}.$$

Since equality holds in the Cauchy-Schwarz inequality $|(y, x)| \le \|y\| \, \|x\|$ if and only if y and x are linearly dependent, we see that actually

$$\Pi(H) = \{(x, x) \in H \times H : \|x\| = 1\}.$$

This implies that in the Hilbert space case the numerical range may also be defined by

$$\mathcal{H}_H(A) = \{(Ax, x) : \|x\| = 1\}.$$

It is well known that $\mathcal{H}_X(A)$ is always a bounded and connected set whose closure contains the spectrum of A: $\operatorname{sp} A \subset \operatorname{clos} \mathcal{H}_X(A)$. The Toeplitz-Hausdorff-Stone theorem says that if $X = H$ is a Hilbert space, then $\mathcal{H}_H(A)$ is necessarily convex. For finite-dimensional Banach spaces X, the numerical range $\mathcal{H}_X(A)$ is obviously closed. This shows that if $X = H = \ell_n^2$, then $\mathcal{H}_H(A)$ contains the convex hull of the eigenvalues of A.

If $X = \ell^p$ or $X = \ell_n^p$, we denote $\mathcal{H}_X(A)$ by $\mathcal{H}_p(A)$. If $A \in \mathcal{B}(\ell^p)$, then $P_n A P_n$ may be thought of as an operator on ℓ_n^p. The purpose of this section is to show that if $1 < p < \infty$, then always

$$\liminf_{n\to\infty} \mathcal{H}_p(P_n A P_n) = \limsup_{n\to\infty} \mathcal{H}_p(P_n A P_n) = \operatorname{clos} \mathcal{H}_p(A).$$

In particular,

$$\liminf_{n\to\infty} \mathcal{H}_p(T_n(b)) = \limsup_{n\to\infty} \mathcal{H}_p(T_n(b)) = \operatorname{clos} \mathcal{H}_p(T(b)).$$

The dual space of ℓ^p may be identified with ℓ^q $(1/p + 1/q = 1)$. Thus, for f in $(\ell^p)^* = \ell^q$, $P_n f$ is a well-defined element of $(\ell^p)^* = \ell^q$.

Lemma 7.9. *Let $1 < p < \infty$. If $(f, x) \in \Pi(\ell^p)$, then $(P_n f/\|P_n f\|_q, P_n x/\|P_n x\|_p)$ is in $\Pi(\ell_n^p)$ for all sufficiently large n.*

Proof. Since $\|P_n f\|_q \to \|f\|_q = 1$ and $\|P_n x\|_p \to \|x\|_p = 1$, it follows that $\|P_n f\|_q \neq 0$ and $\|P_n x\|_p \neq 0$ for all n large enough. Put $f_n = P_n f/\|P_n f\|_q$ and $x_n = P_n x/\|P_n x\|_p$. We are left with showing that $f_n(x_n) = 1$.

Let $x = \{r_0 e^{i\varphi_0}, r_1 e^{i\varphi_1}, \dots\}$ with $0 \le r_j < \infty$ and $0 \le \varphi_j < 2\pi$. By assumption,

$$\|x\|_p = (r_0^p + r_1^p + \cdots)^{1/p} = 1.$$

Set $g = \{r_0^{p/q} e^{-i\varphi_0}, r_1^{p/q} e^{-i\varphi_1}, \dots\}$. Then

$$\|g\|_q = (r_0^p + r_1^p + \cdots)^{1/q} = 1$$

and

$$g(x) = r_0^{p/q} r_0 + r_1^{p/q} r_1 + \cdots = r_0^p + r_1^p + \cdots = 1.$$

Thus, $(g, x) \in \Pi(\ell^p)$.

The space ℓ^q is uniformly convex. This means that if $\|h_1\|_q = \|h_2\|_q = 1$ and $\|h_1 + h_2\|_q = 2$, then $h_1 = h_2$. Since $\|f\|_q = \|g\|_q = 1$ and

$$\|f + g\|_q \ge |f(x) + g(x)| = |1 + 1| = 2,$$

we arrive at the conclusion that $f = g$. Consequently,

$$\begin{aligned}
&\|P_n f\|_q \|P_n x\|_p = \|P_n g\|_q \|P_n x\|_p \\
&= \left(r_0^p + r_1^p + \cdots + r_{n-1}^p\right)^{1/q} \left(r_0^p + r_1^p + \cdots + r_{n-1}^p\right)^{1/p} \\
&= r_0^p + r_1^p + \cdots + r_{n-1}^p \\
&= r_0^{p/q} r_0 + r_1^{p/q} r_1 + \cdots + r_{n-1}^{p/q} r_{n-1} = (P_n g)(P_n x) = (P_n f)(P_n x),
\end{aligned}$$

which is equivalent to the desired equality $f_n(x_n) = 1$. □

Theorem 7.10 (Roch). *Let* $1 < p < \infty$ *and* $A \in \mathcal{B}(\ell^p)$. *Then*

$$\liminf_{n\to\infty} \mathcal{H}_p(P_n A P_n) = \limsup_{n\to\infty} \mathcal{H}_p(P_n A P_n) = \operatorname{clos} \mathcal{H}_p(A).$$

Proof. On regarding ℓ_n^p as a subspace of ℓ_m^p for $n \le m$, we have

$$\Pi(\ell_n^p) \subset \Pi(\ell_m^p) \subset \Pi(\ell^p),$$

whence

$$\mathcal{H}_p(P_n A P_n) \subset \mathcal{H}_p(P_m A P_m) \subset \mathcal{H}_p(A).$$

This shows that

$$\limsup_{n\to\infty} \mathcal{H}_p(P_n A P_n) \subset \operatorname{clos} \mathcal{H}_p(A).$$

To prove the reverse inclusion, let $(f, x) \in \Pi(\ell^p)$ and define $(f_n, x_n) \in \Pi(\ell_n^p)$ as in (the proof of) Lemma 7.9. Since $\|f_n - f\|_q \to 0$ and $\|x_n - x\|_p \to 0$, we obtain that

$$f(Ax) = \lim_{n\to\infty} f_n(Ax_n) = \lim_{n\to\infty} f_n(P_n A P_n x_n).$$

From Lemma 7.9 we infer that $(f_n, x_n) \in \Pi(\ell_n^p)$. Thus,

$$\mathcal{H}_p(A) \subset \liminf_{n\to\infty} \mathcal{H}_p(P_n A P_n),$$

and since limiting sets are always closed, it results that

$$\operatorname{clos} \mathcal{H}_p(A) \subset \liminf_{n\to\infty} \mathcal{H}_p(P_n A P_n). \qquad \square$$

In the case $p = 2$ and $A = T(a)$, the limit in Theorem 7.10 is known.

Theorem 7.11. *If* $a \in W$, *then*

$$\operatorname{clos} \mathcal{H}_2(T(a)) = \operatorname{conv} \operatorname{sp} T(a) = \operatorname{conv} a(\mathbf{T}).$$

Proof. Let $M(a) : L^2 \to L^2$ be the operator of multiplication by a. From Section 1.6 we know that $\Phi^{-1} T(a) \Phi = PM(a)|H^2$, where P is the orthogonal projection of L^2 onto H^2. This implies that

$$\begin{aligned}
\mathcal{H}_2(T(a)) &= \mathcal{H}_{H^2}(\Phi^{-1} T(a) \Phi) = \mathcal{H}_{H^2}(PM(a)|H^2) \\
&= \{(PM(a)f, f) : f \in H^2, \|f\|_2 = 1\} \\
&= \{(M(a)f, f) : f \in H^2, \|f\|_2 = 1\} \quad (\text{since } P^* f = Pf = f) \\
&\subset \{(M(a)g, g) : g \in L^2, \|g\|_2 = 1\} = \mathcal{H}_{L^2}(M(a)).
\end{aligned} \tag{7.28}$$

The closure of the numerical range of a normal operator is the convex hull of its spectrum (see, e.g., [150, Problem 171]). As $M(a)$ is normal, we deduce that

$$\operatorname{clos} \mathcal{H}_{L^2}(M(a)) = \operatorname{conv} \operatorname{sp} M(a) = \operatorname{conv} a(\mathbf{T}). \tag{7.29}$$

Consequently,

$$\begin{aligned}\operatorname{clos} \mathcal{H}_2(T(a)) &\subset \operatorname{clos} \mathcal{H}_{L^2}(M(a)) \quad \text{(by (7.28))}\\ &= \operatorname{conv} a(\mathbf{T}) \quad \text{(by (7.29))}\\ &\subset \operatorname{conv} \operatorname{sp} T(a) \quad \text{(by Corollary 1.12)}\\ &\subset \operatorname{clos} \mathcal{H}_2(T(a)) \quad \text{(since always } \operatorname{sp} A \subset \operatorname{clos} \mathcal{H}_X(A)),\end{aligned}$$

which gives the assertion. □

7.6 Collective Perturbations

Let $\mathbf{G}$ be the collection of all sequences $\{G_n\}_{n=1}^{\infty}$ of complex $n \times n$ matrices G_n such that $\|G_n\|_2 \to 0$ as $n \to \infty$.

Theorem 7.12 (Roch). *If $a \in W$ then*

$$\bigcup_{\{G_n\} \in \mathbf{G}} \limsup_{n \to \infty} \operatorname{sp} (T_n(a) + G_n) = \operatorname{sp} T(a).$$

Proof. Let $\lambda \notin \operatorname{sp} T(a)$. Then, by Theorem 3.7, $\{T_n(a - \lambda)\}$ is a stable sequence:

$$\limsup_{n \to \infty} \|T_n^{-1}(a - \lambda)\|_2 < \infty.$$

It follows that if $\|G_n\|_2 \to 0$ and μ is in some sufficiently small open neighborhood U of λ, then

$$\limsup_{n \to \infty} \| \, (T_n(a - \mu) + G_n)^{-1} \, \|_2 < \infty.$$

This implies that $U \cap \operatorname{sp} (T_n(a) + G_n) = \emptyset$ for all sufficiently large n, whence

$$\lambda \notin \limsup_{n \to \infty} \operatorname{sp} (T_n(a) + G_n).$$

Now take $\lambda \in \operatorname{sp} T(a)$. By virtue of Theorem 3.7, $\{T_n(a - \lambda)\}$ is not stable. If $T_{n_k}(a - \lambda)$ is not invertible for infinitely many n_k, then

$$\lambda \in \limsup_{n \to \infty} \operatorname{sp} T_n(a).$$

So assume $T_n(a - \lambda)$ is invertible for all $n \geq n_0$ but

$$\limsup_{n \to \infty} \|T_n^{-1}(a - \lambda)\|_2 = \infty.$$

There are $n_1 < n_2 < n_3 < \cdots$ and $x_{n_k} \in \ell^2_{n_k}$ such that

$$\|x_{n_k}\|_2 = 1 \ \text{ and } \ \|T_{n_k}^{-1}(a-\lambda)x_{n_k}\|_2 \geq k.$$

Put $y_{n_k} = T_{n_k}^{-1}(a-\lambda)x_{n_k}$ and let G_{n_k} be the matrix of the linear operator

$$G_{n_k} : \ell^2_{n_k} \to \ell^2_{n_k}, \ z \mapsto -\frac{(z, y_{n_k})x_{n_k}}{\|y_{n_k}\|_2^2}.$$

Obviously,

$$\|G_{n_k}\|_2 = \sup_{\|z\|_2=1} \frac{|(z, y_{n_k})| \, \|x_{n_k}\|_2}{\|y_{n_k}\|_2^2} \leq \frac{\|x_{n_k}\|_2}{\|y_{n_k}\|_2} \leq \frac{1}{k}.$$

Let $G_n = 0$ for $n \in \mathbf{N} \setminus \{n_1, n_2, \dots\}$. Then $\{G_n\} \in \mathbf{G}$. Since

$$\big(T_{n_k}(a) + G_{n_k} - \lambda I\big)\, y_{n_k} = T_{n_k}(a-\lambda)y_{n_k} + G_{n_k}y_{n_k} = x_{n_k} - x_{n_k} = 0,$$

it follows that $T_{n_k} + G_{n_k} - \lambda I$ is not invertible. Hence

$$\lambda \in \limsup_{n\to\infty} \operatorname{sp}(T_n(a) + G_n). \qquad \square$$

Exercises

1. Let K be a compact operator on ℓ^2. Prove that

$$\liminf_{n\to\infty} \operatorname{sp}(I + P_nKP_n) = \limsup_{n\to\infty} \operatorname{sp}(I + P_nKP_n) = \operatorname{sp}(I + K).$$

2. Let a and b satisfy $0 < a \leq b \leq \infty$. Show that there exists a selfadjoint operator $A \in \mathcal{B}(\ell^2)$ such that

$$\|A^{-1}\|_2 = a \quad \text{and} \quad \limsup_{n\to\infty} \|(P_nAP_n)^{-1}P_n\|_2 = b.$$

3. Let A be a selfadjoint operator on ℓ^2 and suppose that $\operatorname{sp}_{\text{ess}} A$ is a connected set. Put $A_n = P_nAP_n|\operatorname{Im} P_n$. Prove that

$$\liminf_{n\to\infty} \operatorname{sp} A_n = \limsup_{n\to\infty} \operatorname{sp} A_n = \operatorname{sp} A,$$
$$\lim_{n\to\infty} \|(A_n - \lambda I)^{-1}\|_2 = \|(A - \lambda I)^{-1}\|_2 \quad (\lambda \in \mathbf{C} \setminus \operatorname{sp} A).$$

4. For $m \geq 0$, let $A : \ell^2 \to \ell^2$ be the operator

$$A : \{x_0, x_1, x_2, \dots\} \mapsto \{x_m, x_{m+1}, \dots\}.$$

Prove that $\operatorname{sp}_\varepsilon^{(2)} A = (1+\varepsilon)\overline{\mathbf{D}}$.

5. Let $m \geq 2$ and $b(t) = t + t^{-m}/m$. Find the set $V(b)$ and show that

$$\{\lambda \in V(b) : \|T^{-1}(b - \lambda)\|_2 = 1/\varepsilon\}$$

is the union of $n + 1$ pure circular arcs of curvature $1/\varepsilon$.

6. Let A_n be an $n \times n$ matrix and put

$$R_i'(A_n) = \left(\sum_{j=1}^{n} |a_{ij}| \right) - |a_{ii}|.$$

Show that

$$\mathrm{sp}_\varepsilon^{(p)}(A_n) \subset \bigcup_{i=1}^{n} \{\lambda \in \mathbf{C} : |\lambda - a_{ii}| \leq R_i'(A_n) + \varepsilon n\}.$$

7. Let $\varepsilon > 0$. Show that there exist Laurent polynomials b and c such that

$$\mathrm{sp}\, T(b) + \varepsilon \overline{\mathbf{D}} \neq \mathrm{sp}_\varepsilon^{(2)} T(b), \quad \mathrm{sp}_\varepsilon^{(2)} T(c) \neq \mathrm{conv}\, \mathcal{R}(c) + \varepsilon \overline{\mathbf{D}}.$$

8. Show that there exist finite matrices A and B such that

$$\|(\lambda I - A)^{-1}\|_2 = \|(\lambda I - B)^{-1}\|_2 \ \text{ for all } \ \lambda \in \mathbf{C}$$

but $\|p(A)\|_2 \neq \|p(B)\|_2$ for some polynomial p.

9. Show that the numerical range is robust in the following sense: If $E \in \mathcal{B}(X)$ and $\|E\| \leq \varepsilon$, then $\mathcal{H}_X(A + E) \subset \mathcal{H}_X(A) + \varepsilon \overline{\mathbf{D}}$.

10. Let A be an $n \times n$ matrix. The numbers

$$\alpha_{\mathcal{H}}(A) = \max_{\lambda \in \mathcal{H}_2(A)} \mathrm{Re}\, \lambda, \quad \alpha(A) = \max_{\lambda \in \mathrm{sp}\, A} \mathrm{Re}\, \lambda$$

are called the numerical and spectral abscissas of A, respectively. Prove that

$$\lim_{t \to 0+0} \frac{d}{dt} \log \|e^{tA}\|_2 = \alpha_{\mathcal{H}}(A), \quad \lim_{t \to +\infty} \frac{d}{dt} \log \|e^{tA}\|_2 = \alpha(A).$$

11. Let $a, b \in \mathcal{P} \backslash \{0\}$ and suppose that $a_0 = b_0 = 0$. Prove that the equality $T_n(a)T_n(b) = 0$ is impossible for $n \leq 3$ but possible for $n \geq 4$.

Notes

Embree and Trefethen's Web page [110] and book [275] are inexhaustible sources on all aspects of pseudospectra. The first chapter of [275] is on eigenvalues and it ends as follows: "In the highly nonnormal case, vivid though the image may be, the location of the eigenvalues may be as fragile an indicator of underlying character as the hair color of a Hollywood actor.

We shall see that pseudospectra provide equally compelling images that may capture the spirit underneath more robustly."

"In summary, eigenvalues and eigenfunctions have a distinguished history of application throughout the mathematical sciences; we could not get around without them. Their clearest successes, however, are associated with problems that involve well-behaved systems of eigenvectors, which in most contexts means matrices or operators that are normal or nearly so. This class encompasses the majority of applications, but not all of them. For nonnormal problems, the record is less clear, and even the conceptual significance of eigenvalues is open to question."

As for the history of pseudospectra, we take the liberty of citing Trefethen and Embree [275] again. "These data suggest that the notion of pseudospectra has been invented at least five times:

J. M. Varah	1967	r-approximate eigenvalue
	1979	ε-spectrum
H. J. Landau	1975	ε-approximate eigenvalues
S. K. Godunov et al.	1975	spectral portrait
L. N. Trefethen	1990	ε-pseudospectrum
D. Hinrichsen and A. J. Pritchard	1992	spectral value set

One should not trust this table too far, however, as even recent history is notoriously hard to pin down. It is entirely possible that Godunov or Wilkinson thought about pseudospectra in the 1960s, and indeed von Neumann may have thought about them in the 1930s. Nor were others such as Dunford and Schwartz, Gohberg, Halmos, Kato, Keldysch, or Kreiss far away."

The infinite Toeplitz matrix $T(b)$ is normal if and only if the essential range of b is a line segment in the complex plane [77]. Consequently, infinite Toeplitz matrices are typically nonnormal and hence pseudospectra are expected to tell us more about them than spectra. The pioneering work on pseudospectra of Toeplitz matrices is the paper [219] by Reichel and Trefethen. This paper was the source of inspiration for one of the authors' paper [34] and thus for investigations that have essentially resulted in large parts of the present book.

For $B = C = I$, that is, in the unstructured case, Theorems 7.2 and 7.4 are in principle already in [269], [270]. In the structured case, these theorems are due to Hinrichsen, Kelb, Pritchard, and Gallestey [124], [125], [162], [163]. Section 7.1 is based on ideas of [125] and follows our paper [50].

The question whether the resolvent norm $\|(A-\lambda I)^{-1}\|$ may be locally constant arose in connection with [34]. One of the authors (A. B.) posed this question as an open problem at a Banach semester in Warsaw in 1994, and a few weeks later, Andrzej Daniluk of Cracow was able to solve the problem. The proof of Theorem 7.5 is due to him. Theorem 7.6 was established in [62].

Corollary 7.8, that is, Theorem 7.7 for $p = 2$, is due to Landau [185], [186], [187] and Reichel and Trefethen [219]. The first clean proof of this result was given in [34]. For general $p \in (1, \infty)$, Theorem 7.7 was proved in [62].

The (Hilbert space) numerical range $\mathcal{H}_H(A)$ was introduced by Otto Toeplitz [268]. For more on numerical ranges, in particular for proofs of the properties quoted in the text, we refer to the books [30], [148], [150], [167], [275]. Theorem 7.10 was established by

Roch [220]. Theorem 7.12 is also Roch's; it appeared first in [149, Theorem 3.19]. Theorem 7.11 and the proof given here are due to Halmos [150]. This theorem gives us the closure of $\mathcal{H}_2(T(b))$. The set $\mathcal{H}_2(T(b))$ itself was determined by Klein [180]. There are two theorems in [180]. Theorem 1 says that if the Toeplitz operator has a nonconstant symbol and is normal so that the spectrum is a closed interval $[\gamma, \delta] \subset \mathbf{C}$, then the numerical range is the corresponding open interval (γ, δ). Theorem 2 states that if the Toeplitz operator is not normal, then its numerical range is the interior of the convex hull of its spectrum. We remark that Halmos and Klein's results are actually true for arbitrary $b \in L^\infty$.

We will say more on the numerical range of finite Toeplitz matrices in the notes to Chapter 8.

Exercise 6 is from [111]. For Exercise 7 see [71]. Exercise 8 is a result by Greenbaum and Trefethen [144] (and can also be found in [275]). A solution to Exercise 10, which shows that $\alpha_{\mathcal{H}}(A)$ and $\alpha(A)$ give the initial and final slope of the curve $t \mapsto \log \|e^{tA}\|_2$, is in [275], for example. Exercise 11 is from [147].

Further results: convergence speed. Let b be a Laurent polynomial. If λ is in $\mathbf{C} \setminus b(\mathbf{T})$ and $\operatorname{wind}(b, \lambda) \neq 0$, then $\|T_n^{-1}(b-\lambda)\|_2$ goes to infinity at least exponentially due to Theorem 4.1. This implies that the inequality $\|T_n^{-1}(b-\lambda)\|_2 \geq 1/\varepsilon$ is already satisfied for n's of moderate size, and consequently, the convergence of $\mathrm{sp}_\varepsilon^{(2)} T_n(b)$ to $\mathrm{sp}_\varepsilon^{(2)} T(b)$, which is ensured by Corollary 7.8, is very fast. In [34], it is shown that Corollary 7.8 remains true for dense Toeplitz matrices provided the symbol b is piecewise continuous. It was observed in [45] that in the case of piecewise continuous symbols the convergence of $\mathrm{sp}_\varepsilon^{(2)} T_n(b)$ to $\mathrm{sp}_\varepsilon^{(2)} T(b)$ may be spectacularly slow, which has its reason in the fact that $\|T_n^{-1}(b-\lambda)\|_2$ may grow only polynomially. The main result of our paper [54] says that such a slow convergence of pseudospectra is generic even within the class of continuous symbols. In [54], we proved the following. Let $b \in C^2$ and let $\lambda \in \mathbf{C}$ be a point whose winding number with respect to $b(\mathbf{T})$ is -1 (respectively, 1). Then $\|T_n^{-1}(b-\lambda)\|_2$ increases faster than every polynomial,

$$\lim_{n\to\infty} \|T_n^{-1}(b-\lambda)\|_2\, n^{-\beta} = \infty \quad \text{for each} \quad \beta > 0,$$

if and only if Pb (respectively, Qb) is in C^∞. Here $(Pb)(t) := \sum_{j=0}^{\infty} b_j t^j$ and $(Qb)(t) := \sum_{j=-\infty}^{-1} b_j t^j$ for $b(t) = \sum_{j=-\infty}^{\infty} b_j t^j$.

Further results: operator polynomials. Roch [221] considered the polynomials

$$\begin{aligned} L_n(\lambda) &= T_n(b_0) + T_n(b_1)\lambda + \cdots + T_n(b_k)\lambda^k, \\ L_\infty(\lambda) &= T(b_0) + T(b_1)\lambda + \cdots + T(b_k)\lambda^k, \end{aligned}$$

thought of as acting on ℓ_n^2 and ℓ^2, respectively, and proved that if $T(b_k)$ is invertible, then

$$\begin{aligned} \liminf_{n\to\infty} \{\lambda \in \mathbf{C} : \|L_n^{-1}(\lambda)\|_2 \geq 1/\varepsilon\} &= \limsup_{n\to\infty} \{\lambda \in \mathbf{C} : \|L_n^{-1}(\lambda)\|_2 \geq 1/\varepsilon\} \\ &= \{\lambda \in \mathbf{C} : \|L_\infty^{-1}(\lambda)\|_2 \geq 1/\varepsilon\} \end{aligned}$$

for each $\varepsilon > 0$.

Further results: higher order relative spectra. Let $A \in \mathcal{B}(\ell^2)$ and let L be a closed subspace of ℓ^2. We denote by P_L the orthogonal projection of ℓ^2 onto L. For a natural number k, the kth order spectrum $\mathrm{sp}_k(A, L)$ of A relative to L is defined as the set of all $\lambda \in \mathbf{C}$ for which the compression $P_L(A - \lambda I)^k P_L | L$ is not invertible on L. This definition is due to Brian Davies [94], who suggested that second order spectra might be useful for the approximate computation of spectra of self-adjoint operators.

Shargorodsky's paper [252] is devoted to the geometry of $\mathrm{sp}_k(A, L)$ for fixed L and to the limiting behavior of $\mathrm{sp}_k(A, L_n)$ as P_{L_n} converges strongly to the identity operator. One main result is a purely geometric description of the minimal set $\mathcal{Q}_k(K)$ with the property that $\mathrm{sp}_k(A, L) \subset \mathcal{Q}_k(K)$ whenever A is a normal operator with $\mathrm{sp}\, A \subset K$. Let, for example, K be a compact subset of $\mathbf{R}$. Put $a = \min K$ and $b = \max K$. The set $(a, b) \setminus K$ is an at most countable union of open intervals, that is, of the form $\cup_j (a_j, b_j)$. Let $B(c_1, c_2)$ denote the closed disk with diameter $[c_1, c_2]$. Then

$$\mathcal{Q}_2(K) = B(a, b) \setminus \bigcup_j \mathrm{Int}\, B(a_j, b_j),$$

where Int stands for the interior points.

Another remarkable result of [252] states that if A is normal, then

$$\bigcup_{\{L_n\}} \limsup_{n\to\infty} \mathrm{sp}_k(A, L_n) = \mathrm{sp}\, A \cup \mathcal{Q}_k(\mathrm{sp}_{\mathrm{ess}} A),$$

where $\mathrm{sp}_{\mathrm{ess}} A$ is the essential spectrum of A and the union is over all sequences $\{L_n\}$ for which P_{L_n} converges strongly to the identity operator. As a consequence, Shargorodsky obtains that if k is even, then

$$\bigcup_{\{L_n\}} \limsup_{n\to\infty} \mathrm{sp}_k(A, L_n) \cap \mathbf{R} = \mathrm{sp}\, A$$

for every selfadjoint operator A. (In the last two equalities, lim sup may be replaced by lim inf.) This shows that, in contrast to usual spectra, even order relative spectra do not deliver spurious points in the gaps of $\mathrm{sp}_{\mathrm{ess}} A$ when employing a projection method for the approximate computation of $\mathrm{sp}\, A$. For an arbitrary bounded operator A, the estimate

$$\bigcup_{\{L_n\}} \limsup_{n\to\infty} \mathrm{sp}_k(A, L_n) \subset \mathrm{sp}\, A \;\cup\; \frac{\|A\|_{\mathrm{ess}}}{\sin(\pi/(2k))}\, \overline{\mathbf{D}}$$

is proved in [252].

Further results: normal finite Toeplitz matrices. The characterization of finite normal Toeplitz matrices is discussed in [114], [126], [147], [169], [170], [171], [172]. The approach of Gu and Patton [147] is especially elegant and is applicable to the more general problem of determining all $n \times n$ Toeplitz matrices A, B, C, D such that $AB - CD$ is again Toeplitz or zero. For example, [147] contains a simple proof of the following result: The matrix $T_n(a) = (a_{j-k})_{j,k=1}^n$ is normal if and only if there is a $\lambda \in \mathbf{T}$ such that either $a_j = \lambda a_{-(n-j)}$ for $1 \le j \le n-1$ or $a_j = \lambda\, \overline{a_{-j}}$ for $1 \le j \le n-1$. Note that if $a_j = \lambda\, \overline{a_{-j}}$

with $\lambda \in \mathbf{T}$ for $1 \le j \le n-1$, then $a(\mathbf{T})$ is a line segment. Indeed, choose $\mu \in \mathbf{T}$ so that $\mu^2 = \lambda$. Then, for $t \in \mathbf{T}$,

$$\begin{aligned} a(t) &= a_0 + \sum_{j=1}^{n} \left(\lambda\, \overline{a_{-j}}\, t^j + a_{-j} t^{-j}\right) \\ &= a_0 + \mu \sum_{j=1}^{n} \left(\mu\, \overline{a_{-j}}\, t^j + \overline{\mu} a_{-j} t^{-j}\right) = a_0 + 2\mu \sum_{j=1}^{n} \mathrm{Re}\, \left(\mu\, \overline{a_{-j}}\, t^j\right). \end{aligned}$$

In the other case, $a_j = \lambda\, a_{-(n-j)}$ with $\lambda \in \mathbf{T}$ for $1 \le j \le n-1$, the range $a(\mathbf{T})$ need not to be a line segment (consider, for instance, $n = 3$ and $a(t) = t^{-1} + i t^2$).