9. LINEAR PERTURBATION

In this section we study the effect on the spectrum of an operator $T_0 \in BL(X)$ when it is subjected to a perturbation $V_0 \in BL(X)$. Thus, if we denote the perturbed operator $T_0 + V_0$ by T, we wish to obtain information about $\sigma(T)$ when $\sigma(T_0)$ is known. In this process we shall attempt to allow as 'large' a perturbation V_0 as possible.

We start our investigation by considering the invertibility of an operator which is close to an invertible operator.

We first note that if A and B are both invertible operators in $BL(X)$, then

$$B^{-1} - A^{-1} = B^{-1}(A-B)A^{-1} = A^{-1}(A-B)B^{-1} . \tag{9.1}$$

More generally, if $z \in \rho(A) \cap \rho(B)$, then

$$\begin{aligned} R(B,z) - R(A,z) &= R(B,z)(A-B)R(A,z) \\ &= R(A,z)(A-B)R(B,z) . \end{aligned} \tag{9.2}$$

This follows on replacing A by $A - zI$ and B by $B - zI$ in (9.1). The relation (9.2) is known as the <u>second resolvent identity</u>.

THEOREM 9.1 Let $A, B \in BL(X)$ and A be invertible. Let

$$r_\sigma((A-B)A^{-1}) < 1 . \tag{9.3}$$

Then B is invertible, and

$$B^{-1} = A^{-1} \sum_{k=0}^{\infty} [(A-B)A^{-1}]^k = \sum_{k=0}^{\infty} [A^{-1}(A-B)]^k A^{-1} . \tag{9.4}$$

If , in fact,

$$\|[(A-B)A^{-1}]^2\| < 1 , \tag{9.5}$$

then

$$\|B^{-1}\| \le \frac{\|A^{-1}\| \ \|I+(A-B)A^{-1}\|}{1 - \|[(A-B)A^{-1}]^2\|} , \tag{9.6}$$

$$\|B^{-1} - A^{-1}\| \le \frac{\|A^{-1}\| \ \|(A-B)A^{-1}\| \ \|I+(A-B)A^{-1}\|}{1 - \|[(A-B)A^{-1}]^2\|} . \tag{9.7}$$

Proof Let $C = (A-B)A^{-1}$. Then $r_\sigma(C) < 1$, and it follows by putting $z = 1$ in (5.8) that $I - C = BA^{-1}$ is invertible and

$$(I-C)^{-1} = \sum_{k=0}^{\infty} C^k .$$

We claim that $A^{-1}(BA^{-1})^{-1}$ is the inverse of B . For,

$$B[A^{-1}(BA^{-1})^{-1}] = (BA^{-1})(BA^{-1})^{-1} = I ,$$

and since $(BA^{-1})^{-1}(BA^{-1}) = I$, we also have

$$\begin{aligned} I &= A^{-1}(BA^{-1})^{-1}(BA^{-1})A \\ &= [A^{-1}(BA^{-1})^{-1}]B . \end{aligned}$$

Thus, B is invertible, and

$$B^{-1} = A^{-1}(I-C)^{-1} = A^{-1}\sum_{k=0}^{\infty} [(A-B)A^{-1}]^k = \sum_{k=0}^{\infty} [A^{-1}(A-B)]^k A^{-1} ,$$

which proves (9.4). Now, let (9.5) hold, i.e., $\|C^2\| < 1$. Then $r_\sigma(C^2) < 1$, $(I-C^2)$ is invertible, and since $(I-C)^{-1} = (I+C)(I-C^2)^{-1}$,

$$B^{-1} = A^{-1}(I+C)(I-C^2)^{-1} .$$

Also, by (9.1),

$$B^{-1} - A^{-1} = B^{-1}(A-B)A^{-1} = B^{-1}C = A^{-1}(I+C)(I-C^2)^{-1}C .$$

The inequalities (9.6) and (9.7) now follow easily since by (5.9), we have $\|(I-C^2)^{-1}\| \le 1/(1-\|C^2\|)$. //

COROLLARY 9.2 Let $A, B \in BL(X)$, A invertible and $\|(A-B)A^{-1}\| < 1$. Then B is invertible, and

$$\|B^{-1}\| \le \frac{\|A^{-1}\|}{1 - \|(A-B)A^{-1}\|}, \tag{9.8}$$

$$\|B^{-1} - A^{-1}\| \le \frac{\|A^{-1}\| \, \|(A-B)A^{-1}\|}{1 - \|(A-B)A^{-1}\|}. \tag{9.9}$$

Proof Let $C = (A-B)A^{-1}$. Then $\|C\| < 1$ implies $\|C^2\| < 1$, and

$$1 - \|C^2\| \ge 1 - \|C\|^2 = (1-\|C\|)(1+\|C\|).$$

Hence the results follow directly from (9.6) and (9.7) //

If we replace A by $A - zI$ and B by $B - zI$ in (9.3) and (9.4), we obtain the following result, known as the <u>second Neumann expansion</u>: If $z \in \rho(A)$ and $r_\sigma((A-B)R(A,z)) < 1$, then $z \in \rho(B)$, and

$$R(B,z) = R(A,z) \sum_{k=0}^{\infty} [(A-B)R(A,z)]^k. \tag{9.10}$$

In this case, bounds similar to (9.6), (9.7), (9.8), and (9.9) can be easily written down.

Let, now, E be a closed subset of $\rho(A)$. Since by (5.9), $\|R(A,z)\| \to 0$ as $z \to \infty$ and $\|R(A,z)\|$ assumes its maximum when z lies in a compact set, we see that

$$\alpha = \max_{z \in E} \|R(A,z)\| < \infty.$$

It follows by (9.10) that if $E \subset \rho(A)$ and $\|A-B\| < 1/\alpha$, then $E \subset \rho(B)$. In other words, if G is an open set in $\mathbb{C}$, $\sigma(A) \subset G$, and

$$\|A-B\| < 1 / \max\{\|R(A,z)\| : z \notin G\}, \tag{9.11}$$

then $\sigma(B) \subset G$. This property is known as the upper semicontinuity of the spectrum. Let $\epsilon > 0$. By letting

$$G = \{z \in \mathbb{C} : \text{dist}(z,\sigma(A)) < \epsilon\}$$

and δ to be the right hand side of (9.11), we see that whenever $\|A-B\| < \delta$, we have $\text{dist}(\mu,\sigma(A)) < \epsilon$ for every $\mu \in \sigma(B)$, i.e., if $\mu \in \sigma(B)$, then there is $\lambda \in \sigma(A)$ with $|\mu-\lambda| < \epsilon$. This says that if the operator A is perturbed to the operator B by the addition of $B - A$ and if $\|B - A\|$ is small enough, then the spectrum cannot suddenly get enlarged. On the other hand, the spectrum can suddenly shrink, as the following example shows.

Let $X = \ell^2(\mathbb{Z})$, the space of all doubly infinite square-summable complex sequences. For $x = [\ldots,x(-2),x(-1),x(0),x(1),x(2),\ldots]^t \in X$, consider the left shift operator

$$Ax(i) = \begin{cases} x(i+1) , & \text{if } i \neq -1 \\ 0 , & \text{if } i = -1 \end{cases} ,$$

and let

$$A_0x(i) = \begin{cases} 0 , & \text{if } i \neq -1 \\ x(0) , & \text{if } i = -1 \end{cases} ,$$

Then for $t \in \mathbb{C}$,

$$(A+tA_0)x(i) = \begin{cases} x(i+1) , & \text{if } i \neq -1 \\ tx(0) , & \text{if } i = -1 \end{cases} .$$

It can be seen easily that

$$r_\sigma(A) \leq \|A\| = 1 ,$$

and every $\lambda \in \mathbb{C}$ with $|\lambda| < 1$ is an eigenvalue of A with $[\ldots,0,0,1,\lambda,\lambda^2,\ldots]^t$ as a corresponding eigenvector. Since $\sigma(A)$ is closed, we have

$$\sigma(A) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\} .$$

On the other hand, if $0 < |t| \leq 1$, we show that

$$\sigma(A+tA_0) = \{\lambda \in \mathbb{C} : |\lambda| = 1\} .$$

First note that $r_\sigma(A+tA_0) \leq \|A+tA_0\| = 1$. Also,

$$(A+tA_0)^{-1}x = [\ldots,x(-3),x(-2),\frac{x(-1)}{t},x(0),x(1),\ldots]^t .$$

We can similarly write down $[(A+tA_0)^{-1}]^k x$, $k = 2,3,\ldots,$ to find that

$$\|[(A+tA_0)^{-1}]^k\| = \frac{1}{|t|} .$$

Hence by the spectral radius formula (5.10),

$$r_\sigma((A+tA_0)^{-1}) = \lim_{k\to\infty} [\frac{1}{|t|}]^{1/k} = 1 ,$$

so that $\{w \in \mathbb{C} : |w| > 1\} \subset \rho((A+tA_0)^{-1})$, or $\{z \in \mathbb{C} : |z| < 1\}$ is contained in $\rho(A+tA_0)$. It can be seen that if $|\lambda| = 1$ and then $A + tA_0 - \lambda I$ is not onto since the vector y defined by $y(-1) = 1$, $y(i) = 0$, if $i \neq -1$, is not in its range. Thus, because of the perturbation tA_0 (which is arbitrarily small when $|t|$ is so), the spectrum of A has shrunk from the closed unit disk to the unit circle. We note that $(A+tA_0)$ has no eigenvalues if $t \neq 0$.

The above example points out the lack of *lower semicontinuity* of the spectrum, i.e., an open set containing a point of $\sigma(A)$ may not contain a point of $\sigma(B)$ even when $\|B - A\|$ is arbitrarily small. If, however, A commutes with B, or if A and B are self-adjoint, we do have a kind of continuity of the spectrum. See Problem 9.4 and Proposition 13.1.

Let us now consider an (unperturbed) operator T_0, a (perturbation) operator V_0, and let $T = T_0 + V_0$ be called the perturbed operator.

This kind of situation often occurs in quantum mechanics, although the operators T_0 and V_0 are usually unbounded; T_0 is the Hamiltonian of an unperturbed system and V_0 is a potential energy operator, so that $T_0 + V_0$ is the Hamiltonian of the perturbed system.

For $t \in \mathbb{C}$, we study the family of operators

$$T(t) = T_0 + tV_0 . \qquad (9.12)$$

Observe that $T(0) = T_0$ and $T(1) = T_0 + V_0$. Since the function $t \mapsto tV_0$ is linear in t, we say that $T(t) = T_0 + tV_0$ is obtained from T_0 by a linear perturbation. One can consider quadratic or higher order perturbations. In fact, a comprehensive treatment of the analytic perturbation theory when

$$T(t) = T_0 + tV_0 + t^2V_1 + \dots$$

is an 'analytic family of operators' can be found in [K], Chapters II and VII.

The perturbation analysis given here for a family $T(t)$ of bounded operators can be carried out if T_0 is a densely defined closed (linear) operator in X (i.e., the domain D_{T_0} of T_0 is a dense subspace of X and the graph $\{(x, T_0x) : x \in D_{T_0}\}$ of T_0 is a closed subset of $X \times X$) and if for all small $|t|$, $T(t)$ is a closed operator with the same domain as T_0. On the other hand, the analysis breaks down if the domains of $T(t)$ are different from D_{T_0}. For example, let $X = L^2(\mathbb{R})$, and

$$T(t)x(s) = x''(s) + s^2x(s) + ts^4x(s) , \quad s \in \mathbb{R} ,$$

with

$$D_{T_0} = \left\{x \in X : \int_{\mathbb{R}} s^4 |x(s)|^2 ds < \infty , \quad \int_{\mathbb{R}} p^4 |\hat{x}(p)|^2 dp < \infty\right\}$$

and for $t \neq 0$,

$$D_{T(t)} = \left\{x \in X : \int_{\mathbb{R}} s^8 |x(s)|^2 ds < \infty , \quad \int_{\mathbb{R}} p^4 |\hat{x}(p)|^2 dp < \infty\right\} ,$$

where $\hat{x}$ denotes the Fourier transform of x . In this situation, $V_0 x(s) = s^4 x(s)$, $s \in \mathbb{R}$, is called a singular perturbation of $T_0 x(s) = x''(s) + s^2 x(s)$, $s \in \mathbb{R}$. Analytic properties of a singular perturbation are difficult to establish.

For notational ease, we denote $R(T(t),z)$ by $R(t,z)$ when $z \in \rho(T(t))$, and if $t = 0$, we denote $R(0,z)$ by $R_0(z)$. We now prove that for a fixed z , the map $t \mapsto R(t,z)$ is analytic.

THEOREM 9.3 Let $t_0 \in \mathbb{C}$ and fix $z \in \rho(T(t_0))$. If

$$|t-t_0| < 1 / r_\sigma(V_0 R(t_0,z)) ,$$

then $z \in \rho(T(t))$ and

$$(9.13) \qquad R(t,z) = R(t_0,z) \sum_{k=0}^{\infty} [-V_0 R(t_0,z)]^k (t-t_0)^k .$$

The function $t \mapsto R(t,z)$ is thus analytic on a neighbourhood of t_0 , for every fixed $z \in \rho(T(t_0))$.

Further, let E be a closed subset of $\rho(T(t_0))$. Then the series (9.13) converges absolutely and uniformly for $z \in E$ and t in any closed subset of the disk

$$\{t \in \mathbb{C} : |t-t_0| < 1/\max_{z \in E} \|V_0 R(t_0,z)\|\} .$$

Proof Consider $t \in \mathbb{C}$ such that $|t-t_0| < 1/r_\sigma(V_0 R(t_0,z))$. Letting $A = T(t_0) - zI$ and $B = T(t) - zI$, we have $A - B = -(t-t_0)V_0$, and

$$r_\sigma((A-B)A^{-1}) = |t-t_0| r_\sigma(V_0 R(t_0,z)) < 1 .$$

By Theorem 9.1, B is invertible, i.e., $z \in \rho(T(t))$, and

$$R(t,z) = B^{-1} = A^{-1} \sum_{k=0}^{\infty} [(A-B)A^{-1}]^k$$
$$= R(t_0,z) \sum_{k=0}^{\infty} [-V_0R(t_0,z)]^k (t-t_0)^k ,$$

which proves (9.13), and also shows that the function $t \mapsto R(t,z) \in BL(X)$ is analytic on a neighbourhood of t_0 by Theorem 4.8.

Next, for a closed subset E of $\rho(T(t_0))$, let

$$\beta = \max_{z\in E} \|V_0R(t_0,z)\| < \infty .$$

If D is any closed subset of the disk

$$\{t \in \mathbb{C} : |t-t_0| < 1/\beta\} ,$$

then for all $t \in D$, we have $|t-t_0| \leq \delta$ for some $\delta < 1/\beta$. Now, in Proposition 4.6, let $S = E \times D$, and for $(z,t) \in E \times D$, let

$$c_k(z,t) = [-V_0R(t_0,z)]^k (t-t_0)^k , \quad k = 0,1,\dots .$$

Then

$$\sup_{(z,t)\in E\times D} \|c_k(z,t)\|^{1/k} \leq \sup_{(z,t)\in E\times D} \|-V_0R(t_0,z)\| \; |t-t_0| \leq \beta\delta .$$

Since $\beta\delta < 1$, it follows that the series (9.13) converges absolutely and uniformly for $z \in E$ and $t \in D$. //

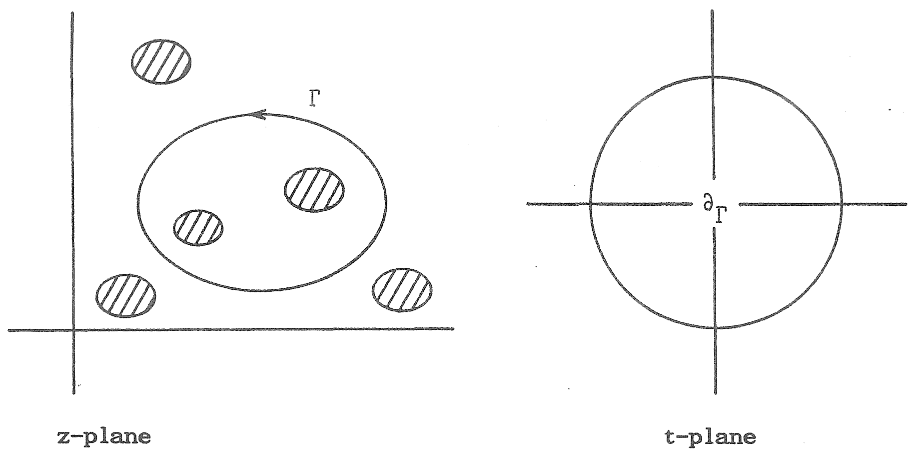
We move on to prove the analyticity of the spectral projection associated with T(t) and a curve Γ in $\rho(T_0)$. Since Γ is a compact set and the function $z \mapsto r_\sigma(V_0R_0(z))$ is upper semicontinuous for $z \in \Gamma$, we see by Corollary 5.5 that

$$\sup_{z\in\Gamma} r_\sigma(V_0R_0(z)) < \infty ,$$

and that there is $z_0 \in \Gamma$ such that $r_\sigma(V_0R_0(z_0)) = \max_{z\in\Gamma} r_\sigma(V_0R_0(z))$.

The following open disk about 0 in the t-plane, which depends on the curve Γ in the z-plane, will be of special interest to us. It was first studied extensively in [C]. Let

(9.14) $$\partial_\Gamma = \{t \in \mathbb{C} : |t| < 1/\max_{z\in\Gamma} r_\sigma(V_0R_0(z))\} .$$



/// : spectrum of T_0

Figure 9.1

Let us denote the spectral projection $P_\Gamma(T_0)$ by P_0 .

THEOREM 9.4 Let $\Gamma \subset \rho(T_0)$. For $t \in \partial_\Gamma$, we have $\Gamma \subset \rho(T(t))$. The spectral projection $P(t) \in BL(X)$ associated with $T(t)$ and Γ is an analytic function of t . In fact, for $t \in \partial_\Gamma$, we have the <u>Kato-Rellich perturbation series</u>

(9.15) $$P(t) = P_0 + \sum_{k=0}^{\infty} P_{(k)} t^k ,$$

where

(9.16) $$P_{(k)} = \frac{(-1)^{k+1}}{2\pi i} \int_\Gamma R_0(z)[V_0R_0(z)]^k dz .$$

Proof Let $z \in \Gamma$, so that $z \in \rho(T_0)$. Letting $t_0 = 0$ in Theorem 9.3, we see that $z \in \rho(T(t))$ for every $t \in \partial_\Gamma$, since

$$|t| < 1/\max_{z\in\Gamma} r_\sigma(V_0R_0(z)) \leq 1/r_\sigma(V_0R_0(z)) .$$

Thus, $\Gamma \subset \rho(T(t))$ for every $t \in \partial_\Gamma$.

Now, fix $t_0 \in \partial_\Gamma$. Letting $E = \Gamma$, in Theorem 9.3, we see that for t in some neighbourhood of t_0, the series

$$R(t,z) = R(t_0,z)\sum_{k=0}^{\infty} [-V_0R(t_0,z)]^k(t-t_0)^k$$

converges uniformly for $z \in \Gamma$. This allows us to integrate the series term by term on Γ (cf. (4.8)), and obtain

$$\begin{aligned} P(t) &= -\frac{1}{2\pi i}\int_\Gamma R(t,z)dz \\ &= \sum_{k=0}^{\infty} \left[-\frac{1}{2\pi i}\int_\Gamma R(t_0,z)[-V_0R(t_0,z)]^k dz\right](t-t_0)^k , \end{aligned}$$

for t near enough to t_0. Thus, $t \mapsto P(t)$ is analytic for t in a neighbourhood of t_0. But since t_0 is an arbitrary point of ∂_Γ, we see that $P(t)$ is analytic on ∂_Γ. The Taylor expansion of $P(t)$ around $t = 0$ is given by the series (9.15). The converse part of Theorem 4.8 shows that this expansion is valid for all t in ∂_Γ. //

The analyticity of the spectral projection $P(t)$ implies, in particular, that $P(t)$ depends continuously on t : if t_1 and t_2 are close then so are $P(t_1)$ and $P(t_2)$ as elements of $BL(X)$. We wish to show that in this case, the ranks of $P(t_1)$ and $P(t_2)$ are equal. For this purpose, we prove some preliminary results which are important in their own right.

LEMMA 9.5 Let P and Q be projections in $BL(X)$ such that

$$r_\sigma(P(P-Q)) < 1 .$$

Then the map $B : P(X) \to P(X)$ given by $Bx = PQx$, $x \in P(X)$, is invertible. In particular,

$$\text{rank } P \leq \text{rank } Q .$$

Proof Let $A = I|_{P(X)}$, which is invertible in $BL(P(X))$. For $x \in P(X)$,

$$(A-B)A^{-1}x = (A-B)x = x - PQx = P(P-Q)x .$$

Thus, $(A-B)A^{-1} = P(P-Q)|_{P(X)}$. But

$$P(P-Q)P = P(P-Q)P|_{P(X)} \oplus P(P-Q)P|_{(I-P)(X)} .$$

Since $P(P-Q)P|_{P(X)} = P(P-Q)|_{P(X)}$, and $P(P-Q)P|_{(I-P)(X)} = 0$, we have by (6.2) and (5.12),

$$r_\sigma((A-B)A^{-1}) = r_\sigma(P(P-Q)P) = r_\sigma(P(P-Q)) < 1 .$$

Now Theorem 9.1 shows that $B : P(X) \to P(X)$ is invertible. In particular, B is onto. Hence

$$\text{rank } P = \dim B(P(X)) = \dim PQ(P(X)) \leq \dim P(Q(X)) \leq \text{rank } Q . \qquad //$$

PROPOSITION 9.6 Let P and Q be projections in $BL(X)$ such that

$$(9.17) \qquad r_\sigma(P(P-Q)) < 1 \text{ and } r_\sigma(Q(Q-P)) < 1 .$$

Then the map $J : P(X) \to Q(X)$ given by $Jx = Qx$, $x \in P(X)$, is a linear homeomorphism onto. In particular,

$$\text{rank } P = \text{rank } Q .$$

These conclusions hold if

$$r_\sigma(P-Q) < 1 .$$

Proof The map J is clearly linear and continuous. It is one to one since if $Jx = Qx = 0$ for some $x \in P(X)$, then $PQx = 0$, and this implies that $x = 0$, as the map $B : P(X) \to P(X)$ given by $Bx = PQx$ is one to one by Lemma 9.5. Next, we show that J is onto. Let $y \in Q(X)$. Then by interchanging P and Q in Lemma 9.5, we see that the map $\tilde{B} : Q(X) \to Q(X)$ given by $\tilde{B}x = QPx$, $x \in Q(X)$, is onto. Hence there is $x \in Q(X)$ such that $QPx = y$, i.e., $J(Px) = y$. As $P(X)$ and $Q(X)$ are closed subspaces of X, they are Banach spaces. The *open mapping theorem* now shows that J^{-1} is continuous, i.e., J is a homeomorphism.

Finally, let $r_\sigma(P-Q) < 1$. Since $P^2 = P$ and $P(P-Q)P = P(P-Q)^2P$, we see by (5.12),

$$r_\sigma(P(P-Q)) = r_\sigma(P(P-Q)P) - r_\sigma((P-Q)^2P) .$$

Now, $(P-Q)^2$ maps $P(X)$ into $P(X)$. Hence by (6.2) and (5.11),

$$r_\sigma((P-Q)^2P) = r_\sigma((P-Q)^2|_{P(X)}) \le r_\sigma((P-Q)^2) = [r_\sigma(P-Q)]^2 < 1 .$$

But we have seen in the proof of Lemma 9.5 that

$$r_\sigma(P(P-Q)|_{P(X)}) = r_\sigma(P(P-Q)).$$

Thus, $r_\sigma(P(P-Q)) < 1$. Also, $r_\sigma(Q(P-Q)) < 1$ by interchanging P and Q. Hence the desired conclusions hold if $r_\sigma(P-Q) < 1$. //

COROLLARY 9.7 Let $\Gamma \subset \rho(T_0)$. Then for every $t \in \partial_\Gamma$,

$$\operatorname{rank} P(t) = \operatorname{rank} P_0 ,$$
$$\operatorname{rank}[I-P(t)] = \operatorname{rank}[I-P_0] .$$

If $\{x_i\}$ is a basis of $P_0(X)$, then $\{P(t)x_i\}$ is a basis of $P(t)(X)$ when $|t|$ is sufficiently small.

Proof By Theorem 9.4, the map $t \mapsto P(t)$ is analytic on ∂_Γ , and hence it is continuous. Thus, for every $t_0 \in \partial_\Gamma$, there is $\epsilon(t_0) > 0$ such that $|t-t_0| < \epsilon(t_0)$ implies

$$r_\sigma(P(t)-P(t_0)) \leq \|P(t)-P(t_0)\| < 1 .$$

Hence rank $P(t)$ = rank $P(t_0)$, by letting $P = P(t_0)$ and $Q = P(t)$ in Proposition 9.6. Now, the nonempty set

$$\{t \in \partial_\Gamma : \dim P(t)(X) = \dim P_0(X)\}$$

is open as well as closed in ∂_Γ , and as such it coincides with ∂_Γ since the disk ∂_Γ is connected. Thus, for all $t \in \partial_\Gamma$,

$$\text{rank } P(t) = \text{rank } P_0 .$$

The statement about rank$[I-P(t)]$ follows similarly by considering the continuity of the map $t \mapsto I - P(t) \in BL(X)$.

Lastly, let $\{x_i\}$ be a basis of $P_0(X)$. For t near 0 , consider the map $J : P_0(X) \to P(t)(X)$, given by

$$Jx = P(t)x , \quad x \in P_0(X) .$$

By Proposition 9.6, J is linear, one to one and onto, and hence sends a basis of $P_0(X)$ to a basis of $P(t)(X)$, showing that $\{P(t)x_i\}$ is a basis of $P(t)(X)$. //

Theorem 9.4 and Corollary 9.7 point out the following interesting facts. If $\Gamma \subset \rho(T_0)$ and the operator T_0 is perturbed to $T(t) = T_0 + tV_0$, then as long as $t \in \partial_\Gamma$, the curve Γ continues to lie in $\rho(T(t))$ and the spectral projection $P(t)$ associated with

$T(t)$ and Γ changes analytically with t ; more importantly, the dimension of $P(t)$ equals the dimension of $P_0 = P(0)$ for all $t \in \partial_\Gamma$.

Since the spectrum of $T(t)$ lying inside Γ is the spectrum of $T(t)|_{P(t)(X)}$, we may expect the spectral values of $T(t)$ inside Γ to depend analytically on t . However, this is not the case for individual spectral values. As an example, let $X = \mathbb{C}^2$, and

$$T_0 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad V_0 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Let Γ denote the unit circle, which encloses the double eigenvalue $\lambda_0 = 0$ of T_0 . For $z \neq 0$,

$$R_0(z) = - \begin{bmatrix} 1/z & 1/z^2 \\ 0 & 1/z \end{bmatrix}, \quad \text{and} \quad V_0R_0(z) = - \begin{bmatrix} 0 & 0 \\ 1/z & 1/z^2 \end{bmatrix}.$$

Hence

$$\sigma(V_0R_0(z)) = \{0, -1/z^2\},$$
$$r_\sigma(V_0R_0(z)) = 1/|z|^2, \quad \max_{z\in\Gamma} r_\sigma(V_0R_0(z)) = 1,$$
$$\partial_\Gamma = \{t \in \mathbb{C} : |t| < 1\}.$$

Now, $T(t) = \begin{bmatrix} 0 & 1 \\ t & 0 \end{bmatrix}$, and for $t \in \partial_\Gamma$, the spectral values of $T(t)$ lying inside Γ are $\pm\sqrt{t}$. However, there is no analytic function $t \mapsto \lambda(t) \in \sigma(T(t)) \cap \text{Int}\,\Gamma = \{\pm\sqrt{t}\}$ for $t \in \partial_\Gamma$.

All the same, we prove that if P_0 is of finite rank, then the arithmetic mean of the spectral points of $T(t)$ inside Γ is indeed an analytic function of $t \in \partial_\Gamma$.

THEOREM 9.8 Let rank $P_0 = m$, $1 \leq m < \infty$. Then for every $t \in \partial_\Gamma$, the only spectral points of $T(t)$ inside Γ are m eigenvalues, say, $\lambda_1(t), \dots, \lambda_m(t)$, counted according to their algebraic multiplicities. The function $\hat{\lambda}$ is analytic on ∂_Γ , where

(9.18) $$\hat{\lambda}(t) = \frac{1}{m}[\lambda_1(t) + \dots + \lambda_m(t)] = \operatorname{tr}(T(t)P(t)).$$

Let $x_i \in X$ and $x_j^* \in X^*$ be such that the matrix $[\langle P_0 x_i, x_j^* \rangle]$, $1 \leq i,j \leq m$ is invertible, and for $t \in \partial_\Gamma$, let $a_{i,j}(t) = \langle P(t)x_i, x_j^* \rangle$, $1 \leq i,j \leq m$. If $A(t)$ denotes the matrix $[a_{i,j}(t)]$, then for $|t|$ sufficiently small, $A(t)$ is invertible; if

$$[A(t)]^{-1} = [b_{i,j}(t)] , \quad c_{i,j}(t) = \langle T(t)P(t)x_i, x_j^* \rangle , \quad i,j = 1,\dots,m ,$$

then

$$\hat{\lambda}(t) = \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{m} b_{i,j}(t) c_{j,i}(t) . \tag{9.19}$$

Proof By Corollary 9.7, rank $P(t) = m < \infty$ for all $t \in \partial_\Gamma$. Hence by Theorem 7.9, the spectrum of $T(t)$ inside Γ consists of a finite number of eigenvalues with finite algebraic multiplicities.

Since $T(t)$ and $P(t)$ commute, we have

$$R(T(t)P(t)) \subset R(P(t)) ,$$

which is of dimension m. Thus, the operator $T(t)P(t)$ is of finite rank and Proposition 3.6 shows that

$$\begin{aligned} \operatorname{tr}(T(t)P(t)) &= \operatorname{tr}(T(t)P(t)|_{P(t)(X)}) \\ &= \operatorname{tr}(T(t)|_{P(t)(X)}) \\ &= \text{the sum of the eigenvalues} \\ &\qquad \text{of } T(t)|_{P(t)(X)} , \text{ by (7.18)} \\ &= \lambda_1(t) + \dots + \lambda_m(t) \\ &= m\, \hat{\lambda}(t) . \end{aligned}$$

This proves (9.18).

For $t \in \partial_\Gamma$, let

$$x_i(t) = P(t)x_i , \quad 1 \leq i \leq m .$$

Then $A(t) = [\langle x_i(t), x_j^*\rangle]$, $1 \le i,j \le m$. Since $A(0) = [\langle P_0 x_i, x_j^*\rangle]$ is invertible and the function $t \mapsto x_i(t) = P(t)x_i \in X$ is analytic (and hence continuous) for each $i = 1,\ldots,m$, we see by Theorem 9.1 that $A(t)$ is invertible if $|t|$ is small enough.

It follows by Remark 3.4 that the set $\{x_1(t),\ldots,x_m(t)\}$ is linearly independent and forms a basis of $P(t)(X)$. Also, if we let

$$y_j^*(t) = \sum_{k=1}^{m} \overline{b_{k,j}(t)} x_k^* , \quad j = 1,\ldots,m ,$$

then

$$\langle x_i(t), y_j^*(t)\rangle = \delta_{i,j} , \quad i,j = 1,\ldots,m .$$

(cf. (3.6).) Now, Proposition 3.6 shows that for $|t|$ small enough,

$$\begin{aligned} m\,\hat{\lambda}(t) &= \mathrm{tr}(T(t)P(t)) \\ &= \sum_{j=1}^{m} \langle T(t)P(t)x_j(t), y_j^*(t)\rangle \\ &= \sum_{j=1}^{m} \langle T(t)P(t)x_j, \sum_{i=1}^{m} \overline{b_{i,j}(t)} x_i^*\rangle \\ &= \sum_{i=1}^{m}\sum_{j=1}^{m} b_{i,j}(t) c_{j,i}(t) . \end{aligned}$$

Since the functions $t \mapsto T(t) \in BL(X)$ and $t \mapsto P(t)x_j \in X$ are analytic, we see that the functions $t \mapsto b_{i,j}(t) \in \mathbb{C}$ and $t \mapsto c_{i,j}(t) \in \mathbb{C}$ are analytic. (See Problem 4.1.) We conclude that the function $t \mapsto \hat{\lambda}(t)$ is analytic on a neighbourhood of 0 . A very similar argument establishes the analyticity of this function in a neighbourhood of an arbitrary point $t_0 \in \partial_\Gamma$. //

Let the spectrum of T_0 inside Γ consist of a single eigenvalue λ_0 of finite algebraic multiplicity. Then by (7.8),

$$R_0(z) = \sum_{k=0}^{\infty} S_0^{k+1}(z-\lambda_0)^k - \frac{P_0}{z-\lambda_0} - \sum_{k=1}^{\ell-1} \frac{D_0^k}{(z-\lambda_0)^{k+1}} .$$

We can use this Laurent expansion of $R_0(z)$ to calculate the coefficients

$$P_{(k)} = \frac{(-1)^{k+1}}{2\pi i} \int_{\Gamma} R_0(z)[V_0 R_0(z)]^k dz$$

in the perturbation series (9.15) for $P(t)$ in terms of P_0, S_0, D_0 and V_0. These can then be used to obtain a series expansion of the arithmetic mean $\hat{\lambda}(t) = \mathrm{tr}(T(t)P(t))$ of the eigenvalues of $T(t)$ inside Γ. These series are considered in [K], p.76 and p.379. We shall not pursue their study here because the coefficients of these series cannot be calculated in an iterative manner. Let λ_0 be a simple eigenvalue of T_0. In the next section, we shall consider series expansions for the simple eigenvalue $\lambda(t)$ of $T(t)$ and for a suitably normalized eigenvector of $T(t)$ corresponding to $\lambda(t)$ which can be calculated in an iterative manner. With this in view, let us study the important special case of a simple eigenvalue.

COROLLARY 9.9 Let the only spectral value of T_0 inside Γ be a simple eigenvalue λ_0. Then for every $t \in \partial_\Gamma$, Γ encloses only one spectral value $\lambda(t)$ of $T(t)$ and it is also a simple eigenvalue. The function $t \mapsto \lambda(t)$ is analytic on ∂_Γ.

Let $x_0 \in X$ and $x_0^* \in X^*$ be such that $\langle P_0 x_0, x_0^* \rangle \neq 0$. If $|t|$ is small enough, we have

$$\lambda(t) = \frac{\langle T(t)P(t)x_0, x_0^* \rangle}{\langle P(t)x_0, x_0^* \rangle}; \tag{9.20}$$

also,

$$x(t) = \frac{P(t)x_0}{\langle P(t)x_0, x_0^* \rangle} \tag{9.21}$$

is an eigenvector of $T(t)$ corresponding to $\lambda(t)$ such that

$\langle x(t), x_0^* \rangle = 1$; $x(t)$ is an analytic function of t in a neighbourhood of 0.

Proof We have $m = \dim P_0(X) = 1$. Hence by Theorem 9.8, $t \mapsto \hat{\lambda}(t) = \lambda(t)$ is analytic on ∂_Γ. Also, let $x_1 = x_0$ and $x_1^* = x_0^*$. Then for $|t|$ small, we have

$$a_{1,1}(t) = \langle P(t)x_0, x_0^* \rangle ,$$
$$b_{1,1}(t) = 1 / \langle P(t)x_0, x_0^* \rangle ,$$
$$c_{1,1}(t) = \langle T(t)P(t)x_0, x_0^* \rangle .$$

Thus, (9.20) follows directly from (9.19). Also, since $\langle P(0)x_0, x_0^* \rangle = \langle P_0x_0, x_0^* \rangle \neq 0$, we see that for $|t|$ small, $\langle P(t)x_0, x_0^* \rangle \neq 0$, so that $P(t)x_0 \neq 0$. Now, since $\lambda(t)$ is simple, we have $P(t)x_0 \in P(t)(X) = Z(T(t) - \lambda(t)I)$. This shows that $x(t)$ is an eigenvector of $T(t)$ corresponding to $\lambda(t)$. The relation $\langle x(t), x_0^* \rangle = 1$ is immediate. Since both the numerator and the denominator of $x(t)$ are analytic and the denominator does not vanish, we see that $x(t)$ is analytic on a neighbourhood of 0. //

One can give a direct proof of the analyticity of the function $t \mapsto \lambda(t)$ of Corollary 9.9 without invoking Theorem 9.8. Since $\lambda(t)$ is a simple eigenvalue of $T(t)$ for $t \in \partial_\Gamma$, we have $T(t)P(t) = \lambda(t)P(t)$, so that

$$\langle T(t)P(t)x_0, x^* \rangle = \lambda(t)\langle P(t)x_0, x_0^* \rangle .$$

As $\langle P(0)x_0, x_0^* \rangle = \langle P_0x_0, x_0^* \rangle \neq 0$, we see that $\langle P(t)x_0, x_0^* \rangle \neq 0$ if $|t|$ is sufficiently small. Hence the relation (9.20) holds. In particular, $t \mapsto \lambda(t)$ is an analytic function on a neighbourhood of 0.

Problems

9.1 Let $A \in BL(X)$ be invertible and $B \in BL(X)$ satisfy $\|A^{-1}(A-B)\| < 1$. If $Ax = a$ and $By = b$, then

$$\|y-x\| \leq \frac{\|A^{-1}(b-a)\| + \|A^{-1}(A-B)\| \, \|x\|}{1 - \|A^{-1}(A-B)\|} .$$

(Hint: (9.4))

9.2 (Iterative refinement of the solution of an operator equation) Let $A \in BL(X)$ and $y \in X$. Consider an invertible $A_0 \in BL(X)$ such that $r_\sigma((A-A_0)A_0^{-1}) < 1$ and $A_0 x_0 = y$. For $j = 1,2,\ldots,$ let

$$r_{j-1} = y - Ax_{j-1} , \quad A_0 u_j = r_{j-1} , \quad x_{j+1} = x_j + u_j .$$

Then A is invertible and (x_j) converges to the unique $x \in X$ such that $Ax = y$.

9.3 (General Neumann expansion) Let $z \in \rho(A)$. If

$$r_\sigma([(w-z)I+(A-B)]R(A,z)) < 1 ,$$

then $w \in \rho(B)$ and

$$R(B,w) = R(A,z) \sum_{k=0}^{\infty} [[(w-z)I+(A-B)]R(A,z)]^k .$$

($A = B$ gives (5.7) and $w = z$ gives (9.10).) In particular, if $\epsilon_1 + \epsilon_2 \leq \|R(A,z)\|^{-1}$, $|w-z| < \epsilon_1$ and $\|A-B\| < \epsilon_2$, then $w \in \rho(B)$,

$$\|R(B,w)\| \leq \|R(A,x)\|/(1-r) ,$$

$$\|R(B,w) - R(A,z)\| \leq r\|R(A,z)\|/(1-r) ,$$

where $r = (|w-z|+\|A-B\|)\|R(A,z)\|$. The function $(A,z) \mapsto R(A,z) \in BL(X)$ is jointly continuous on $\{(A,z) : A \in BL(X) , z \in \rho(A)\} \subset BL(X) \times \mathbb{C}$.

9.4 Let A, $B \in BL(X)$. Assume either that A and B commute, or that A and B are self-adjoint. Then

$$\max\left\{\max_{\lambda\in\sigma(A)} \mathrm{dist}(\lambda,\sigma(B)) , \max_{\lambda\in\sigma(B)} \mathrm{dist}(\lambda,\sigma(A))\right\} \le r_\sigma(A-B) \le \|A-B\| .$$

9.5 Let Γ and $\tilde{\Gamma}$ be simple closed curves in $\rho(T_0)$ such that $\Gamma \subset \mathrm{Int}\,\tilde{\Gamma}$. Assume that $P_\Gamma(T_0)$ is of finite rank and that T_0 has no spectral values between Γ and $\tilde{\Gamma}$. Then for all $t \in \partial_\Gamma \cap \partial_{\tilde{\Gamma}}$, $T(t)$ has no spectral values between Γ and $\tilde{\Gamma}$.

9.6 Let P and Q be projections such that $r_\sigma(P-Q) < 1$. Then the operator $QP + (I-Q)(I-P)$ is invertible. It maps $R(P)$ onto $R(Q)$ and $Z(P)$ onto $Z(Q)$. Hence $\mathrm{rank}\, P = \mathrm{rank}\, Q$.

9.7 Let D be a connected metric space and for $s \in D$, let $Q(s)$ be a projection in $BL(X)$. If $s \mapsto Q(s)$ is continuous, then the rank of $Q(s)$ is constant (finite or infinite) for $s \in D$.

9.8 Let $m = 2$ in Theorem 9.8. Then for $|t|$ small enough,

$$\hat{\lambda}(t) = \frac{(a_{2,2}c_{1,1} - a_{1,2}c_{2,1} - a_{2,1}c_{1,2} + a_{1,1}c_{2,2})(t)}{2(a_{1,1}a_{2,2} - a_{1,2}a_{2,1})(t)} .$$

9.9 Under the hypothesis of Corollary 9.9, let for $|t| < r$, with r small enough,

$$y(t) = \frac{P(t)x_0}{\sqrt{\langle P(t)x_0, x_0^*\rangle}} , \quad y^*(t) = \frac{[P(t)]^* x^*_0}{\sqrt{\langle [P(t)]^* x_0^*, x_0\rangle}} ,$$

where $\sqrt{\ }$ denotes the principal branch of the square root. Then the function $t \mapsto y(t) \in X$ is analytic, the function $t \mapsto y^*(t)$ is antianalytic (i.e., $t \mapsto y^*(\bar{t})$ is analytic) and $\langle y(t), y^*(t)\rangle = 1$. In particular, if T_0 and V_0 are self-adjoint operators on a Hilbert space X, t is real, and we choose $x_0^* = x_0$, then $\|y(t)\| = 1$.