7. Banach algebras

Definition 7.1. A is called a *Banach algebra* (with unit) if: (1) A is a Banach space;
(2) There is a multiplication $A \times A \to A$ that has the following properties:

$$(xy)z = x(yz), \quad (x+y)z = xz + yz, \quad x(y+z) = xy + xz,$$
$$c(xy) = (cx)y = x(cy)$$

for all $x, y, z \in A$, $c \in \mathbb{C}$. Moreover, there is a *unit element* e: $ex = xe = x$ for all $x \in A$;
(3) $\|e\| = 1$;
(4) $\|xy\| \le \|x\|\,\|y\|$ for all $x, y \in A$.

So a Banach algebra is an algebra, that is, it is a vector space that is also a ring, with a compatibility condition between the two structures, or we can say more informally that an algebra is a vector space with a multiplication, obeying the usual algebraic rules. Moreover, a Banach algebra is also a Banach space, and the norm is compatible with the algebraic structure (conditions (3), (4)).

At the end of the last chapter, we decided to try to analyze normal operators on a Hilbert space H. Banach algebras will prove useful here, because of the following:

Example 7.1. If X is a Banach space, then $A = B(X)$ is a Banach algebra, with the composition of operators as multiplication and the operator norm. Indeed, we know from Theorem 2.12(b) that A is a Banach space, and composition of operators has the properties from (2) of Definition 7.1. The identity operator 1 is the unit element; of course $\|1\| = \sup_{\|x\|=1} \|x\| = 1$, as required, and (4) was discussed in Exercise 2.25.

Of course, there are more examples:

Example 7.2. $A = \mathbb{C}$ with the usual multiplication and the absolute value as norm is a Banach algebra.

Example 7.3. $A = C(K)$ with the pointwise multiplication $(fg)(x) = f(x)g(x)$ is a Banach algebra. Most properties are obvious. The unit element is the function $e(x) \equiv 1$; clearly, this has norm 1, as required. To verify (4), notice that

$$\|fg\| = \max_{x\in K} |f(x)g(x)| \le \max_{x\in K} |f(x)| \max_{x\in K} |g(x)| = \|f\|\,\|g\|.$$

Example 7.4. Similarly, $A = L^\infty$ and $A = \ell^\infty$ with the pointwise multiplication are Banach algebras.

Notice that the last three examples are in fact *commutative* Banach algebras, that is, $xy = yx$ for all $x, y \in A$. On the other hand, $B(X)$ is not commutative if $\dim X > 1$.

Example 7.5. $A = L^1(\mathbb{R})$ with the convolution product

$$(fg)(x) = (f * g)(x) = \int_{-\infty}^{\infty} f(t)g(x-t)\, dt$$

satisfies most of the properties from Definition 7.1, but does not have a unit element, so this would provide an example of a Banach algebra without a unit.

On the other hand, the discrete analog $A = \ell^1(\mathbb{Z})$ with the convolution product

$$(xy)_n = \sum_{j=-\infty}^{\infty} x_j y_{n-j}$$

is a Banach algebra with unit. Both L^1 and ℓ^1 are commutative.

Exercise 7.1. Prove the claims about the unit elements: Show that there is no function $f \in L^1(\mathbb{R})$ such that $f * g = g * f = g$ for all $g \in L^1(\mathbb{R})$. Also, find the unit element e of $\ell^1(\mathbb{Z})$.

We now start to develop the general theory of Banach algebras.

Theorem 7.2. *Multiplication is continuous in Banach algebras: If $x_n \to x$, $y_n \to y$, then $x_n y_n \to xy$.*

Proof.

$$\begin{aligned} \|x_n y_n - xy\| &\le \|(x_n - x)y_n\| + \|x(y_n - y)\| \\ &\le \|x_n - x\|\,\|y_n\| + \|x\|\,\|y_n - y\| \to 0 \end{aligned}$$

□

We call $x \in A$ *invertible* if there exists $y \in A$ such that $xy = yx = e$. Note that on the Banach algebra $B(H)$, this reproduces the definition of invertibility in $B(H)$ that was given earlier, in Chapter 6. Returning to the general situation, we observe that if $x \in A$ is invertible, then y with these properties is unique. We write $y = x^{-1}$ and call x^{-1} the *inverse* of x. We denote the set of invertible elements by $G(A)$. Here, the choice of symbol is motivated by the fact that $G(A)$ is a group, with multiplication as the group operation. Indeed, if $x, y \in G(A)$, then also $xy \in G(A)$ and $x^{-1} \in G(A)$: this can be verified by just writing down the inverses: $(xy)^{-1} = y^{-1}x^{-1}$, $(x^{-1})^{-1} = x$. Moreover, $e \in G(A)$ ($e^{-1} = e$), and of course multiplication is associative.

If A, B are algebras, then a map $\phi : A \to B$ is called a *homomorphism* if it preserves the algebraic structure. More precisely, we demand that

ϕ is linear (as a map between vector spaces) and $\phi(xy) = \phi(x)\phi(y)$ for all $x, y \in A$.

We also could have included the condition that $\phi(e) = e'$ (but didn't); if we want to emphasize the distinction, then such a map could be called a *unital homomorphism.* See also Exercise 7.3 below.

By a *complex homomorphism* we mean a homomorphism $\phi : A \to \mathbb{C}$, $\phi \not\equiv 0$.

Proposition 7.3. *Let ϕ be a complex homomorphism. Then $\phi(e) = 1$ and $\phi(x) \neq 0$ for all $x \in G(A)$.*

Proof. Since $\phi \not\equiv 0$, there is a $y \in A$ with $\phi(y) \neq 0$. Since $\phi(y) = \phi(ey) = \phi(e)\phi(y)$, we have $\phi(e) = 1$. If $x \in G(A)$, then $\phi(x)\phi(x^{-1}) = \phi(e) = 1$, so $\phi(x) \neq 0$. $\square$

Exercise 7.2. Let A be an algebra, with unit e. True or false:
(a) $fx = x$ for all $x \in A \implies f = e$;
(b) $0x = 0$ for all $x \in A$;
(c) $xy = 0 \implies x = 0$ or $y = 0$;
(d) $xy = zx = e \implies x \in G(A)$ and $y = z = x^{-1}$;
(e) $xy, yx \in G(A) \implies x, y \in G(A)$;
(f) $xy = e \implies x \in G(A)$ or $y \in G(A)$.

Exercise 7.3. (a) Give an example of a homomorphism $\phi : A \to B$, $\phi \not\equiv 0$, that is not unital, that is, $\phi(e_A) \neq e_B$.

(b) However, show that if $\phi : A \to B$ is a *surjective* homomorphism, then ϕ is unital.

Theorem 7.4. *Let A be a Banach algebra. If $x \in A$, $\|x\| < 1$, then $e - x \in G(A)$ and*

$$(e - x)^{-1} = \sum_{n=0}^{\infty} x^n. \tag{7.1}$$

Moreover, if ϕ is a complex homomorphism, then $|\phi(x)| < 1$.

Here, we define $x^n = xx\cdots x$ as the n-fold product of x with itself, and $x^0 := e$. The series from (7.1) is then defined, as usual, as the norm limit of the partial sums (existence of this limit is part of the statement, of course). It generalizes the geometric series to the Banach algebra setting and is called the *Neumann series.*

Proof. Property (4) from Definition 7.1 implies that $\|x^n\| \leq \|x\|^n$. Since $\|x\| < 1$, we now see that $\sum \|x^n\|$ converges. It follows that the Neumann series converges, too (see Exercise 2.22). By the continuity of the multiplication in A,

$$(e-x)\sum_{n=0}^{\infty} x^n = (e-x)\lim_{N\to\infty}\sum_{n=0}^{N} x^n = \lim_{N\to\infty}(e-x)\sum_{n=0}^{N} x^n$$
$$= \lim_{N\to\infty}\left(\sum_{n=0}^{N} x^n - \sum_{n=0}^{N} x^{n+1}\right) = \lim_{N\to\infty}\left(e - x^{N+1}\right) = e.$$

A similar calculation shows that $\left(\sum_{n=0}^{\infty} x^n\right)(e-x) = e$, so indeed $e-x \in G(A)$ and the inverse is given by (7.1).

If $c \in \mathbb{C}$, $|c| \geq 1$, then, by what has just been shown, $e-(1/c)x \in G(A)$, so $\phi(e-(1/c)x) = 1-(1/c)\phi(x) \neq 0$ by Proposition 7.3, that is, $\phi(x) \neq c$. □

Corollary 7.5. *(a) $G(A)$ is open. More precisely, if $x \in G(A)$ and $\|h\| < \frac{1}{\|x^{-1}\|}$, then $x+h \in G(A)$ also.*
(b) If ϕ is a complex homomorphism, then $\phi \in A^*$ and $\|\phi\| = 1$.

Proof. (a) Write $x+h = x(e+x^{-1}h)$. Since $\|x^{-1}h\| \leq \|x^{-1}\|\,\|h\| < 1$, Theorem 7.4 shows that $e+x^{-1}h \in G(A)$. Since also $x \in G(A)$ and $G(A)$ is a group, it follows that $x+h \in G(A)$, too.

(b) The last part of Theorem 7.4 says that ϕ is bounded and $\|\phi\| \leq 1$. Since $\phi(e) = 1$ and $\|e\| = 1$, it follows that $\|\phi\| = 1$. □

Exercise 7.4. We can also run a more quantitative version of the argument from (a) to obtain the following: Inversion in Banach algebras is a continuous operation. More precisely, if $x \in G(A)$ and $\epsilon > 0$, then there exists $\delta > 0$ such that if $\|y-x\| < \delta$, then $y \in G(A)$ and $\|y^{-1}-x^{-1}\| < \epsilon$. Prove this.

We now introduce the Banach algebra version of Definition 6.7.

Definition 7.6. Let $x \in A$. Then we define

$$\rho(x) = \{z \in \mathbb{C} : x - ze \in G(A)\},$$
$$\sigma(x) = \mathbb{C} \setminus \rho(x),$$
$$r(x) = \sup\{|z| : z \in \sigma(x)\}.$$

We call $\rho(x)$ the *resolvent set,* $\sigma(x)$ the *spectrum,* and $r(x)$ the *spectral radius* of x. Also, $(x-ze)^{-1}$, which is defined for $z \in \rho(x)$, is called the *resolvent* of x.

Theorem 7.7. *(a) $\rho(x)$ is an open subset of $\mathbb{C}$.*
(b) The resolvent $R(z) = (x-ze)^{-1}$ admits power series representations about every point $z_0 \in \rho(x)$. More specifically, if $z_0 \in \rho(x)$, then there

exists $r > 0$ *with* $\{z : |z - z_0| < r\} \subseteq \rho(x)$ *and*

$$(x - ze)^{-1} = \sum_{n=0}^{\infty} (x - z_0 e)^{-n-1} (z - z_0)^n$$

for all z *with* $|z - z_0| < r$.

Here we define y^{-n}, for $n \geq 0$ and invertible y, as $y^{-n} = (y^{-1})^n$. More succinctly, we can say that the resolvent $R(z)$ is a *holomorphic function* (which takes values in a Banach algebra) on $\rho(x)$; we then simply define this notion by the property from Theorem 7.7(b).

Proof. (a) This is an immediate consequence of Corollary 7.5 because $\|x - ze - (x - z_0)e\| = |z - z_0|$.

(b) As in (a) and the proof of Corollary 7.5(a), we see that $B_r(z_0) \subseteq \rho(x)$ if we take $r = 1/\|(x - z_0 e)^{-1}\|$. Moreover, we can use the Neumann series to expand $R(z)$, as follows:

$$\begin{aligned}
(x - ze)^{-1} &= \left[(e - (z - z_0)(x - z_0 e)^{-1})(x - z_0 e) \right]^{-1} \\
&= (x - z_0 e)^{-1} \left[e - (z - z_0)(x - z_0 e)^{-1} \right]^{-1} \\
&= (x - z_0 e)^{-1} \sum_{n=0}^{\infty} (x - z_0 e)^{-n} (z - z_0)^n \\
&= \sum_{n=0}^{\infty} (x - z_0 e)^{-n-1} (z - z_0)^n
\end{aligned}$$

We have used the continuity of the multiplication in the last step. □

Theorem 7.8. *(a)* $\sigma(x)$ *is a compact, non-empty subset of* $\mathbb{C}$.
(b) $r(x) = \inf_{n \in \mathbb{N}} \|x^n\|^{1/n} = \lim_{n \to \infty} \|x^n\|^{1/n}$

The existence of the limit in part (b) is part of the statement. Note also that $\|x^n\| \leq \|x\|^n$, by using property (4) from Definition 7.1 repeatedly, so we always have $r(x) \leq \|x\|$. Strict inequality is possible here.

The inconspicuous *spectral radius formula* from part (b) has a rather remarkable property: $r(x)$ is a purely algebraic quantity (to work out $r(x)$, find the biggest $|z|$ for which $x - ze$ does not have a multiplicative inverse), but nevertheless $r(x)$ is closely related to the norm on A via the spectral radius formula.

Proof. (a) We know from Theorem 7.7(a) that $\sigma(x) = \mathbb{C} \setminus \rho(x)$ is closed. Moreover, if $|z| > \|x\|$, then $x - ze = (-z)(e - (1/z)x) \in G(A)$ by Theorem 7.4, so $\sigma(x)$ is also bounded and thus a compact subset of $\mathbb{C}$.

We also obtain the representation

$$(x-ze)^{-1} = -\sum_{n=0}^{\infty} z^{-n-1}x^n \tag{7.2}$$

from Theorem 7.4; this is valid for $|z| > \|x\|$. Suppose now that we had $\sigma(x) = \emptyset$. For an arbitrary $F \in A^*$, we can introduce the function $g : \rho(x) \to \mathbb{C}$, $g(z) = F((x-ze)^{-1})$. Since we are assuming that $\sigma(x) = \emptyset$, this function is really defined on all of $\mathbb{C}$. Moreover, by using Theorem 7.7(b) and the continuity of F, we see that g has convergent power series representations about every point and thus is holomorphic (in the traditional sense). If $|z| \geq 2\|x\|$, then (7.2) yields

$$\begin{aligned} |g(z)| &= \left| F\left(\sum_{n=0}^{\infty} z^{-n-1}x^n \right) \right| \leq \|F\| \sum_{n=0}^{\infty} |z|^{-n-1}\|x\|^n \\ &\leq \frac{\|F\|}{|z|} \sum_{n=0}^{\infty} 2^{-n} = \frac{2\|F\|}{|z|}. \end{aligned}$$

So g is a bounded entire function. By Liouville's Theorem, g must be constant. Since $g(z) \to 0$ as $|z| \to \infty$, this constant must be zero. This, however, is not possible, because $F(y) = 0$ for all $F \in A^*$ would imply that $y = 0$, by Corollary 4.2(b), but clearly the inverse $(x-ze)^{-1}$ can not be the zero element of A. The assumption that $\sigma(x) = \emptyset$ must be dropped.

(b) Let $n \in \mathbb{N}$ and let $z \in \mathbb{C}$ be such that $z^n \in \rho(x^n)$. We can write

$$x^n - z^n e = (x-ze)(z^{n-1}e + z^{n-2}x + \cdots + x^{n-1}),$$

and now multiplication from the right by $(x^n - z^n e)^{-1}$ shows that $x - ze$ has a right inverse. A similar calculation provides a left inverse also, so it follows that $z \in \rho(x)$ (we are using Exercise 7.2(d) here!). Put differently, $z^n \in \sigma(x^n)$ if $z \in \sigma(x)$. The proof of part (a) has shown that $|z| \leq \|y\|$ for all $z \in \sigma(y)$, so we now obtain $|z^n| \leq \|x^n\|$ for all $z \in \sigma(x)$. Since the spectral radius $r(x)$ was defined as the maximum of the spectrum (we cautiously worked with the supremum in the original definition, but we now know that $\sigma(x)$ is a compact set), this says that $r(x) \leq \inf \|x^n\|^{1/n}$.

Next, consider again the function $g(z) = F((x-ze)^{-1})$, with $F \in A^*$. This is holomorphic on $\rho(x) \supseteq \{z \in \mathbb{C} : |z| > r(x)\}$. Furthermore, for $|z| > \|x\|$, we have the power series expansion (in z^{-1})

$$g(z) = -\sum_{n=0}^{\infty} F(x^n)\left(z^{-1}\right)^{n+1}.$$

This shows that g is holomorphic near $z = \infty$; more precisely, if we let $\zeta = 1/z$ and $h(\zeta) = g(1/\zeta)$, then h has a convergent power series expansion, $h(\zeta) = -\sum_{n=0}^{\infty} F(x^n)\zeta^{n+1}$, which is valid for small $|\zeta|$. Moreover, by our earlier remarks, h also has a holomorphic extension to the disk $\{\zeta : |\zeta| < 1/r(x)\}$ (the extension is provided by the original definition of g). A power series converges on the biggest disk to which the function can be holomorphically extended; thus the radius of convergence of the series $\sum F(x^n)\zeta^{n+1}$ is at least $1/r(x)$. In particular, if $0 < a < 1/r(x)$, then

$$F(x^n)a^n = F(a^n x^n) \to 0 \qquad (n \to \infty).$$

Since this is true for arbitrary $F \in A^*$, we have in fact shown that $a^n x^n \xrightarrow{w} 0$. Weakly convergent sequences are bounded (Exercise 4.23), so $\|a^n x^n\| \le C$ ($n \in \mathbb{N}$) for suitable $C = C(a) \ge 0$. Hence

$$\|x^n\|^{1/n} \le \frac{1}{a}\, C^{1/n} \to \frac{1}{a},$$

and here $a < 1/r(x)$ was arbitrary and we can take the limit on any subsequence, so $r(x) \ge \limsup_{n\to\infty} \|x^n\|^{1/n}$. On the other hand, we have already proved that

$$r(x) \le \inf_{n\in\mathbb{N}} \|x^n\|^{1/n} \le \liminf_{n\to\infty} \|x^n\|^{1/n},$$

so we now obtain the full claim. □

You should now work out some spectra in concrete examples. The first example is particularly important for us, so I'll state this as a Proposition:

Proposition 7.9. *Consider the Banach algebra $A = C(K)$. Then, for $f \in C(K)$, we have $\sigma(f) = f(K)$, where $f(K) = \{f(x) : x \in K\}$. Moreover, $r(f) = \|f\|$ for all $f \in C(K)$.*

Exercise 7.5. Prove Proposition 7.9.

Exercise 7.6. (a) Show that on $A = \ell^\infty$, we have

$$\sigma(x) = \overline{\{x_n : n \in \mathbb{N}\}}.$$

Also, show that again $r(x) = \|x\|$ for all $x \in \ell^\infty$.
(b) Show that on $A = L^\infty(X, \mu)$, we have

$$\sigma(f) = \{z \in \mathbb{C} : \mu(\{x \in X : |f(x) - z| < \epsilon\}) > 0 \text{ for all } \epsilon > 0\}.$$

(This set is also called the *essential range* of f; roughly speaking, it is the range of f, but we ignore what happens on null sets, in keeping with the usual philosophy. Also, it is again true that $r(f) = \|f\|$.)

Exercise 7.7. Show that on $A = B(\mathbb{C}^n)$, the spectrum $\sigma(T)$ of a matrix $T \in B(\mathbb{C}^n) = \mathbb{C}^{n\times n}$ is the set of eigenvalues of T (this was discussed earlier, in Chapter 6). Now find a matrix $T \in \mathbb{C}^{2\times 2}$ for which $r(T) < \|T\|$.

The fact that spectra are always non-empty has the following consequence:

Theorem 7.10 (Gelfand-Mazur). *If A is a Banach algebra with $G(A) = A \setminus \{0\}$, then $A \cong \mathbb{C}$.*

More specifically, the claim is that there is an identification map between A and $\mathbb{C}$ (thought of as a Banach algebra, with the usual multiplication and the absolute value as the norm) that preserves the complete Banach algebra structure: There is a map $\varphi : A \to \mathbb{C}$ that is bijective (= preserves sets), a homomorphism (= preserves the algebraic structure), and an isometry (= preserves the norm).

Proof. By Theorem 7.8(a), we can pick a number $z(x) \in \sigma(x)$ for each $x \in A$. So $x - z(x)e \notin G(A)$, but the only non-invertible element of A is the zero vector, so $x = z(x)e$ (and we also learn that in fact $\sigma(x) = \{z(x)\}$). The map $\varphi : A \to \mathbb{C}$, $\varphi(x) = z(x)$ has the desired properties. $\square$

In the last part of this chapter, we discuss the problem of how the spectrum of an element changes when we pass to a smaller Banach algebra. Let B be a Banach algebra, and let $A \subseteq B$ be a subalgebra. By this we mean that A with the structure inherited from B is a Banach algebra itself. We also insist that $e \in A$. Note that this latter requirement could be dropped, and in fact that would perhaps be the more common version of the definition of a subalgebra. The following Exercise discusses the difference between the two versions. It may also be helpful to recall Exercise 7.3 here.

Exercise 7.8. Let B be a Banach algebra, and let $C \subseteq B$ be a subset that also is a Banach algebra *with unit element* with the structure (algebraic operations, norm) inherited from B. Give a (simple) example of such a situation where $e \notin C$.
Remark: This is very straightforward. Just make sure you don't get confused. C is required to have a unit (call it f, say), but what exactly is f required to do?

If we now fix an element $x \in A$ of the smaller algebra, we can consider its spectrum with respect to both algebras. From the definition, it is clear that $\sigma_A(x) \supseteq \sigma_B(x)$: everything that is invertible in A remains invertible in B, but we may lose invertibility when going from B to A simply because the required inverse may no longer be part of the algebra.

Furthermore, Theorem 7.8(b) shows that $r_A(x) = r_B(x)$. More can be said about the relation between $\sigma_A(x)$ and $\sigma_B(x)$, but this requires some work. This material will be needed later, but is of a technical character and can be given a light reading at this point.

We need the notion of connected components in a topological space; actually, we only need this for the space $X = \mathbb{C}$. Recall that we call a topological space X *connected* if the only decomposition of X into two disjoint open sets is the trivial one: if $X = U \cup V$, $U \cap V = \emptyset$, and U, V are open, then $U = X$ or $V = X$. A subset $A \subseteq X$ is called *connected* if A with the relative topology is a connected topological space. A *connected component* is a maximal connected set. These connected components always exist and in fact every point lies in a unique connected component, and the whole space can be written as the disjoint union of its connected components.

For a detailed reading of this final section, the following topological warm-up should be helpful. You can either try to solve this directly or do some reading.

Exercise 7.9. (a) Prove these facts. More specifically, show that if $x \in X$, then there exists a unique maximal connected set C_x with $x \in C_x$. So if D is another connected set with $x \in D$, then $D \subseteq C_x$. Also, show that if $x, y \in X$, then either $C_x \cap C_y = \emptyset$ or $C_x = C_y$.
(b) Call $A \subseteq X$ *pathwise connected* if any two points can be joined by a continuous curve: If $x, y \in A$, then there exists a continuous map $\varphi : [0,1] \to A$ with $\varphi(0) = x$, $\varphi(1) = y$. Show that a pathwise connected set is connected.
(c) Show that if $U \subseteq \mathbb{C}$ is open, then all connected components of U are open subsets of $\mathbb{C}$.

We are heading towards the following general result:

Theorem 7.11. *Let $A \subseteq B$ be a subalgebra of the Banach algebra B, and let $x \in A$. Then we have a representation of the following type:*

$$\sigma_A(x) = \sigma_B(x) \cup C,$$

where C is a (necessarily disjoint) union of connected components of $\rho_B(x)$ ($C = \emptyset$ is possible, of course).

This has the following consequences (whose relevance is more obvious):

Corollary 7.12. *(a) If $\rho_B(x)$ is connected, then $\sigma_A(x) = \sigma_B(x)$. In particular, this conclusion holds if $\sigma_B(x) \subseteq \mathbb{R}$.*
(b) If $\overset{\circ}{\sigma}_A(x) = \emptyset$, then $\sigma_A(x) = \sigma_B(x)$.

Here, $\overset{\circ}{C}$ denotes the interior of C, defined as the largest open subset of C.

To prove the Corollary (given the Theorem), note that the hypothesis that $\rho_B(x)$ is connected means that the only connected component of this set is $\rho_B(x)$ itself, but we cannot have $\sigma_A(x) = \sigma_B(x) \cup \rho_B(x)$ because $\rho_B(x)$ is unbounded (being the complement of the compact set $\sigma_B(x)$), and $\sigma_A(x)$ needs to be compact. If $\sigma_B(x)$ is a (compact!) subset of $\mathbb{R}$, then clearly its complement $\rho_B(x)$ is pathwise connected, thus connected. Compare Exercise 7.9(b).

Part (b) follows from the fact that the connected components of the open set $\rho_B(x)$ are open (Exercise 7.9(c)), so if we had $C \neq \emptyset$, then automatically $\sigma_A(x)$ would have non-empty interior.

To prove Theorem 7.11, we need the following topological fact.

Lemma 7.13. *Let $U, V \subseteq X$ be open subsets of the topological space X, and assume that $U \subseteq V$, $(\overline{U} \setminus U) \cap V = \emptyset$. Then $U = \bigcup V_\alpha$, where the V_α are connected components of V (but not necessarily all of these, of course).*

Proof. We must show that if W is a connected component of V with $W \cap U \neq \emptyset$, then $W \subseteq U$ (assuming this, we can then indeed write U as the union of those components of V that intersect U). So let W be such a component. From the assumption of the Lemma, we obtain $W \cap (\overline{U} \setminus U) = \emptyset$. Hence

$$W = (W \cap U) \cup (W \cap \overline{U}^c).$$

This is a decomposition of W into two disjoint relatively (!) open subsets. Since W is connected by assumption, one of these must be all of W, and since $W \cap U \neq \emptyset$, it is the first set: $W \cap U = W$, so $W \subseteq U$, as desired. $\square$

We are now ready for the

Proof of Theorem 7.11. We will verify the hypotheses of Lemma 7.13 for $U = \rho_A(x)$, $V = \rho_B(x)$. The Lemma will then show that $\rho_A(x) = \bigcup_{\alpha \in I_0} V_\alpha$, where the V_α are connected components of $\rho_B(x)$. Also, $\rho_B(x) = \bigcup_{\alpha \in I} V_\alpha$, and $I_0 \subseteq I$, so we indeed obtain

$$\sigma_A(x) = \mathbb{C} \setminus \rho_A(x) = \sigma_B(x) \cup \bigcup_{\alpha \in I \setminus I_0} V_\alpha.$$

Clearly, $\rho_A(x) \subseteq \rho_B(x)$, so we must check that $(\overline{\rho_A(x)} \setminus \rho_A(x)) \cap \rho_B(x) = \emptyset$. Let $z \in \overline{\rho_A(x)} \setminus \rho_A(x)$. Then there are $z_n \in \rho_A(x)$, $z_n \to z$. I now claim that

$$\left\|(x - z_n e)^{-1}\right\| \to \infty \qquad (n \to \infty). \tag{7.3}$$

Suppose this were wrong. Then $|z - z_n| \, \|(x - z_n e)^{-1}\| < 1$ for some (large) n, and hence

$$(x - z_n e)^{-1}(x - ze) = e - (z - z_n)(x - z_n e)^{-1}$$

would be in $G(A)$ by Theorem 7.4, but then also $x - ze \in G(A)$, and this contradicts $z \notin \rho_A(x)$. Thus (7.3) holds. Now (7.3) also prevents $x - ze$ from being invertible in B, because inversion is a continuous operation in Banach algebras (Exercise 7.4). More explicitly, if we had $x - ze \in G(B)$, then, since $x - z_n e \to x - ze$, it would follow that $(x - z_n e)^{-1} \to (x - ze)^{-1}$, but this convergence is ruled out by (7.3). So $x - ze \notin G(B)$, or, put differently, $z \notin \rho_B(x)$. $\square$

Exercise 7.10. Show that $r(xy) = r(yx)$. *Hint:* Use the formula $(xy)^n = x(yx)^{n-1}y$.

Exercise 7.11. Prove that $\sigma(xy)$ and $\sigma(yx)$ can at most differ by the point 0. (In particular, this again implies the result from Exercise 7.10, but of course the direct proof suggested there was much easier.)
Suggested strategy: This essentially amounts to showing that $e - xy$ is invertible if and only if $e - yx$ is invertible. So assume that $e - xy \in G(A)$. Assume also that $\|x\|, \|y\| < 1$ and write $(e-xy)^{-1}$, $(e-yx)^{-1}$ as Neumann series. Use the formula from the previous problem to obtain one inverse in terms of the other. Then show that this formula actually works in complete generality, without the assumptions on x, y.