Bruce K. Driver

Functional Analysis Tools with Examples

February 13, 2020 *File:241Functional˙2020s.tex*

Contents

Part I

Background

Vector Valued Integration Theory

[The reader interested in integrals of Hilbert valued functions, may go directly to Section 1.5 below and bypass the Bochner integral altogether.]

Let X be a Banach space and $(\Omega, \mathcal{F}, \mu)$ be a measure space. Given a "nice enough" function, $f : \Omega \to X$, we would like to define $\int_\Omega f d\mu$ as an element in X. Whatever integration theory we develop we minimally want to require that

$$\varphi\left(\int_\Omega f d\mu\right) = \int_\Omega \varphi \circ f \; d\mu \text{ for all } \varphi \in X^*. \tag{1.1}$$

Basically, the Pettis Integral developed below makes definitions so that there is an element $\int_\Omega f d\mu \in X$ such that Eq. (1.1) holds. There are some subtleties to this theory in its full generality which we will avoid for the most part. For many more details see [15–18] and especially [48]. Other references are **Pettis Integral** (See Craig Evans PDE book?) also see

http://en.wikipedia.org/wiki/Pettis_integral

and

http://www.math.umn.edu/~garrett/m/fun/Notes/07_vv_integrals.pdf

1.1 Pettis Integral

Remark 1.1 (Wikipedia quote). In mathematics, the Pettis integral or Gelfand–Pettis integral, named after I. M. Gelfand and B.J. Pettis, extends the definition of the Lebesgue integral to functions on a measure space which take values in a Banach space, by the use of duality. The integral was introduced by Gelfand for the case when the measure space is an interval with Lebesgue measure. The integral is also called the weak integral in contrast to the Bochner integral, which is the strong integral.

We start by describing a weak form of measurability and integrability

Definition 1.2. *Let X be a Banach space and $(\Omega, \mathcal{F}, \mu)$ be a measure space. We say a function $u : \Omega \to X$ is **weakly measurable** if $f \circ u : \Omega \to \mathbb{C}$ is measurable for all $f \in X^*$.*

Definition 1.3. *A weakly measurable function $u : \Omega \to X$ is said to be weakly L^1 if there exists $U \in L^1(\Omega, \mathcal{F}, \mu)$ such that $\|u(\omega)\| \le U(\omega)$ for μ-a.e. $\omega \in \Omega$. We denote the weakly L^1 functions by $L^1(\mu : X)$ and for $u \in L^1(\mu : X)$ we define,*

$$\|u\|_1 := \inf\left\{\int_\Omega U(\omega)\, d\mu(\omega) : U \ \ni \ \|u(\cdot)\| \le U(\cdot) \ \text{a.e.}\right\}.$$

Remark 1.4. It is easy to check that $L^1(\Omega, \mathcal{F}, \mu)$ is a vector space and that $\|\cdot\|_1$ satisfies

$$\begin{aligned} \|zu\|_1 &= |z| \, \|u\|_1 \text{ and} \\ \|u+v\|_1 &\le \|u\|_1 + \|v\|_1 \end{aligned}$$

for all $z \in \mathbb{F}$ and $u, v \in L^1(\mu : X)$. As usual $\|u\|_1 = 0$ iff $u(\omega) = 0$ except for ω in a μ-null set. Indeed, if $\|u\|_1 = 0$, there exists U_n such that $\|u(\cdot)\| \le U_n(\cdot)$ *a.e.* and $\int_\Omega U_n d\mu \downarrow 0$ as $n \to \infty$. Let E be the null set, $E = \cup_n E_n$, where E_n is a null set such that $\|u(\omega)\| \le U_n(\omega)$ for $\omega \notin E$. Now by replacing U_n by $\min_{k \le n} U_n$ if necessary we may assume that U_n is a decreasing sequence such that $\|u\| \le U := \lim_{n\to\infty} U_n$ off of E and by DCT $\int_\Omega U d\mu = 0$. This shows $\{U \neq 0\}$ is a null set and therefore $\|u(\omega)\| = 0$ if ω is not in the null set, $E \cup \{U \neq 0\}$.

To each $u \in L^1(\mu : X)$ let

$$\tilde{u}(\varphi) := \int_\Omega \varphi \circ u d\mu \tag{1.2}$$

which is well defined since $\varphi \circ u$ is measurable and $|\varphi \circ u| \le \|\varphi\|_{X^*} \|u(\cdot)\| \le \|\varphi\|_{X^*} U(\cdot)$ a.e. Moreover it follows that

$$|\tilde{u}(\varphi)| \le \|\varphi\|_{X^*} \int_\Omega U d\mu \implies |\tilde{u}(\varphi)| \le \|\varphi\|_{X^*} \|u\|_1$$

which shows $\tilde{u} \in X^{**}$ and

$$\|\tilde{u}\|_{X^{**}} \le \|u\|_1. \tag{1.3}$$

Definition 1.5. *We say* $u \in L^1(\mu : X)$ *is* ***Pettis integrable*** *(and write* $u \in L^1_{Pet}(\mu : X)$*) if there exists (a necessarily unique)* $x_u \in X$ *such that* $\tilde{u}(\varphi) = \varphi(x_u)$ *for all* $\varphi \in X^*$*. We say that* x_u *is the* ***Pettis integral*** *of* u *and denote* x_u *by* $\int_\Omega u d\mu$*. Thus the Pettis integral of* u*, if it exists, is the unique element* $\int_\Omega u d\mu \in X$ *such that*

$$\varphi\left(\int_\Omega u d\mu\right) = \int_\Omega (\varphi \circ u)\, d\mu. \tag{1.4}$$

Let us summarize the easily proved properties of the Pettis integral in the next theorem.

Theorem 1.6 (Pettis Integral Properties). *The space,* $L^1_{Pet}(\mu : X)$*, is a vector space, the map,*

$$L^1_{Pet}(\mu : X) \ni u \to \int_\Omega f d\mu \in X$$

is linear, and

$$\left\|\int_\Omega u d\mu\right\|_X \le \|u\|_1 \text{ for all } u \in L^1_{Pet}(\mu : X). \tag{1.5}$$

Moreover, if X *is reflexive then* $L^1(\mu : X) = L^1_{Pet}(\mu : X)$*.*

Proof. These assertions are straight forward and will be left to the reader with the exception of Eq. (1.5). To verify Eq. (1.5) we recall that the map $X \ni x \to \hat{x} \in X^{**}$ (where $\hat{x}(\varphi) := \varphi(x)$) is an isometry and the Pettis integral, x_u, is defined so that $\hat{x}_u = \tilde{u}$. Therefore,

$$\left\|\int_\Omega u d\mu\right\|_X = \|x_u\|_X = \|\hat{x}_u\|_{X^{**}} = \|\tilde{u}\|_{X^{**}} \le \|u\|_1. \tag{1.6}$$

wherein we have used Eq. (1.3) for the last inequality.

■

Exercise 1.1. Suppose $(\Omega, \mathcal{F}, \mu)$ is a measure space, X and Y are Banach spaces, and $T \in B(X,Y)$. If $u \in L^1_{Pet}(\mu; X)$ then $T \circ u \in L^1_{Pet}(\mu; Y)$ and

$$\int_\Omega T \circ u d\mu = T \int_\Omega u d\mu. \tag{1.7}$$

When X is a separable metric space (or more generally when u takes values in a separable subspace of X), the Pettis integral (now called the Bochner integral) is a fair bit better behaved, see Theorem 1.13 below. As a warm up let us consider Riemann integrals of continuous integrands which is typically all we will need in these notes.

1.2 Riemann Integrals of Continuous Integrands

In this section, suppose that $-\infty < a < b < \infty$ and $f \in C([a,b], X)$ and for $\delta > 0$ let

$$\operatorname{osc}_\delta(f) := \max\{\|f(c) - f(c')\| : c, c' \in [a,b] \text{ with } |c - c'| \le \delta\}.$$

By uniform continuity, we know that $\operatorname{osc}_\delta(f) \to 0$ as $\delta \downarrow 0$. It is easy to check that $f \in L^1(m : X)$ where m is Lebesgue measure on $[a,b]$ and moreover in this case $t \to \|f(t)\|_X$ is continuous and hence measurable.

Theorem 1.7. *If* $f \in C([a,b], X)$*, then* $f \in L^1_{Pet}(m; X)$*. Moreover if*

$$\Pi = \{a = t_0 < t_1 < \cdots < t_n = b\} \subset [a,b],$$

$\{c_i\}_{i=1}^n$ *are arbitrarily chosen so that* $t_{i-1} \le c_i \le t_i$ *for all* i*, and* $|\Pi| := \max_i |t_i - t_{i-1}|$ *denotes the mesh size of let* Π*, then*

$$\left\|\int_a^b f(t)\, dt - \sum_{i=1}^n f(c_i)(t_i - t_{i-1})\right\|_X \le (b-a)\operatorname{osc}_{|\Pi|}(f). \tag{1.8}$$

Proof. Using the notation in the statement of the theorem, let

$$S_\Pi(f) := \sum_{i=1}^n f(c_i)(t_i - t_{i-1}).$$

If $t_{i-1} = s_0 < s_1 < \cdots < s_k = t_i$ and $s_{j-1} \le c'_j \le s_j$ for $1 \le j \le k$, then

$$\begin{aligned}
&\left\|f(c_i)(t_i - t_{i-1}) - \sum_{j=1}^k f\left(c'_j\right)(s_j - s_{j-1})\right\| \\
&= \left\|\sum_{j=1}^k f(c_i) - f\left(c'_j\right)(s_j - s_{j-1})\right\| \\
&\le \sum_{j=1}^k \left\|f(c_i) - f\left(c'_j\right)\right\|(s_j - s_{j-1}) \\
&\le \operatorname{osc}_{|\Pi|}(f)\sum_{j=1}^k (s_j - s_{j-1}) = \operatorname{osc}_{|\Pi|}(f)(t_i - t_{i-1}).
\end{aligned}$$

So if Π' refines Π, then by the above argument applied to each pair, t_{i-1}, t_i, it follows that

$$\|S_\Pi(f) - S_{\Pi'}(f)\| \le \sum_{i=1}^{n} \mathrm{osc}_{|\Pi|}(f)(t_i - t_{i-1}) = \mathrm{osc}_{|\Pi|}(f)\cdot(b-a). \quad (1.9)$$

Now suppose that $\{\Pi_n\}_{n=1}^\infty$ is a sequence of increasing partitions (i.e. $\Pi_n \subset \Pi_{n+1}\ \forall\ n \in \mathbb{N}$) with $|\Pi_n| \to 0$ as $n \to \infty$. Then by the previously displayed equation it follows that

$$\|S_{\Pi_n}(f) - S_{\Pi_m}(f)\| \le \mathrm{osc}_{|\Pi_{m\wedge n}|}(f)\cdot(b-a).$$

As the latter expression goes to zero as $m, n \to \infty$, it follows that $\lim_{n\to\infty} S_{\Pi_n}(f)$ exists and in particular,

$$\varphi\left(\lim_{n\to\infty} S_{\Pi_n}(f)\right) = \lim_{n\to\infty} S_{\Pi_n}(\varphi\circ f) = \int_a^b \varphi(f(t))\,dt\ \forall\ \varphi \in X^*.$$

Since the right member of the previous equation is the standard real variable Riemann or Lebesgue integral, it is independent of the choice of partitions, $\{\Pi_n\}$, and of the corresponding c's and we may conclude $\lim_{n\to\infty} S_{\Pi_n}(f)$ is also independent of any choices we made. We have now shown that $f \in L^1_{Pet}(m;X)$ and that

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} S_{\Pi_n}(f).$$

To prove the estimate in Eq. (1.8), simply choose $\{\Pi_n\}_{n=1}^\infty$ as above so that $\Pi \subset \Pi_1$ and then from Eq. (1.9) it follows that

$$\|S_\Pi(f) - S_{\Pi_n}(f)\| \le \mathrm{osc}_{|\Pi|}(f)\cdot(b-a)\ \ \forall\ n \in \mathbb{N}.$$

Letting $n \to \infty$ in this inequality gives the estimate in Eq. (1.8). ■

Remark 1.8. Let $f \in C(\mathbb{R}, X)$. We leave the proof of the following properties to the reader with the caveat that many of the properties follow directly from their real variable cousins after testing the identities against a $\varphi \in X^*$.

1. For $a < b < c$,
$$\int_a^c f(t)\,dt = \int_a^b f(t)\,dt + \int_b^c f(t)\,dt$$
and moreover this result holds independent of the ordering of $a, b, c \in \mathbb{R}$ provided we define,
$$\int_a^c f(t)\,dt := -\int_c^a f(t)\,dt \text{ when } c < a.$$
2. For all $a \in \mathbb{R}$,
$$\frac{d}{dt}\int_a^t f(s)\,ds = f(t) \ \text{ for all } t \in \mathbb{R}.$$

3. If $f \in C^1(\mathbb{R}, X)$, then
$$f(t) - f(s) = \int_s^t \dot f(\tau)\,d\tau\ \forall\ s, t \in \mathbb{R}$$
where
$$\dot f(t) := \lim_{h\to 0}\frac{f(t+h) - f(t)}{h} \in X.$$
4. Again the triangle inequality holds,
$$\left\|\int_a^b f(t)\,dt\right\|_X \le \left|\int_a^b \|f(t)\|_X\,dt\right|\ \forall\ a, b \in \mathbb{R}.$$

Exercise 1.2. Suppose that $(X, \|\cdot\|)$ is a Banach space, $J = (a,b)$ with $-\infty \le a < b \le \infty$ and $f_n : J \to X$ are continuously differentiable functions such that there exists a summable sequence $\{a_n\}_{n=1}^\infty$ satisfying

$$\|f_n(t)\| + \left\|\dot f_n(t)\right\| \le a_n \text{ for all } t \in J \text{ and } n \in \mathbb{N}. \quad (1.10)$$

Show:

1. $\sup\left\{\left\|\frac{f_n(t+h) - f_n(t)}{h}\right\| : (t,h) \in J \times \mathbb{R}\ \ni\ t+h \in J \text{ and } h \neq 0\right\} \le a_n.$
2. The function $F : \mathbb{R} \to X$ defined by
$$F(t) := \sum_{n=1}^\infty f_n(t) \ \text{ for all } t \in J$$
is differentiable and for $t \in J$,
$$\dot F(t) = \sum_{n=1}^\infty \dot f_n(t).$$

Definition 1.9. *A function f from an open set $\Omega \subset \mathbb{C}$ to a complex Banach space X is* ***analytic on*** *Ω if*

$$f'(z) := \lim_{h\to 0}\frac{f(z+h) - f(z)}{h} \ \textit{exists}\ \forall\ z \in \Omega$$

and is ***weakly analytic*** *on Ω if $\ell \circ f$ is analytic on Ω for every $\ell \in X^*$.*

Analytic functions are trivially weakly analytic and next theorem shows the converse is true as well. In what follows let

$$D(z_0, z) := \{z \in \mathbb{C} : |z - z_0| < \rho\}$$

be the open disk in $\mathbb{C}$ centered at z_0 of radius $\rho > 0$.

Theorem 1.10. *If $f:\Omega\to X$ is a weakly analytic function then f is analytic. Moreover if $z_0\in\Omega$ and $\rho>0$ is such that $\overline{D(z_0,\rho)}\subset\Omega$, then for all $w\in D(z_0,\rho)$,*

$$f(w)=\frac{1}{2\pi i}\oint_{\partial D(z_0,\rho)}\frac{f(z)}{z-w}dz, \tag{1.11}$$

$$f^{(n)}(w)=\frac{n!}{2\pi i}\oint_{\partial D(z_0,\rho)}\frac{f(z)}{(z-w)^{n+1}}dz,\ \ and \tag{1.12}$$

$$f(w)=\sum_{n=0}^{\infty}\frac{f^{(n)}(z_0)}{n!}(w-z_0)^n. \tag{1.13}$$

Proof. Let $K\subset\Omega$ be a compact set and $\varepsilon>0$ such that $z+h\in\Omega$ for all $|h|\le\varepsilon$. Since $\ell\circ f$ is analytic we know that

$$\left|\ell\left(\frac{f(z+h)-f(z)}{h}\right)\right|=\left|\frac{\ell\circ f(z+h)-\ell\circ f(z)}{h}\right|\le M_\ell<\infty$$

for all $z\in K$ and $0<|h|\le\varepsilon$ where

$$M_\ell=\sup_{z\in K\text{ and }|h|\le\varepsilon}\left|(\ell\circ f)'(z+h)\right|.$$

Therefore by the uniform boundedness principle,

$$\sup_{z\in K,0<|h|\le\varepsilon}\left\|\frac{f(z+h)-f(z)}{h}\right\|_X=\sup_{z\in K,0<|h|\le\varepsilon}\left\|\left[\frac{f(z+h)-f(z)}{h}\right]^{\hat{}}\right\|_{X^{**}}<\infty$$

from which it follows that f is necessarily continuous.

If $\overline{D(z_0,\rho)}\subset\Omega$ and $\ell\in X^*$, then for all $w\in D(z_0,\rho)$ we have by the standard theory of analytic functions that

$$\ell\circ f(w)=\frac{1}{2\pi i}\oint_{\partial D(z_0,\rho)}\frac{\ell\circ f(z)}{z-w}dz=\ell\circ\left(\frac{1}{2\pi i}\oint_{\partial D(z_0,\rho)}\frac{f(z)}{z-w}dz\right).$$

As this identity holds for all $\ell\in X^*$ it follows that Eq. (1.11) is valid. Equation (1.12) now follows by repeated differentiation past the integral and in particular it now follows that f is analytic. The power series expansion for f in Eq. (1.13) now follows exactly as in the standard analytic function setting. Namely we write

$$\frac{1}{z-w}=\frac{1}{z-z_0-(w-z_0)}=\frac{1}{z-z_0}\frac{1}{1-\frac{w-z_0}{z-z_0}}$$
$$=\frac{1}{z-z_0}\sum_{n=0}^{\infty}\left(\frac{w-z_0}{z-z_0}\right)^n$$

and plug this identity into Eq. (1.11) to discover,

$$f(w)=\sum_{n=0}^{\infty}a_n(w-z_0)^n$$

where

$$a_n=\frac{1}{2\pi i}\oint_{\partial D(z_0,\rho)}\frac{f(z)}{(z-z_0)^{n+1}}dz=f^{(n)}(z_0).$$

■

Remark 1.11. If X is a complex Banach space, J is an open subset of $\mathbb{C}$, and $f_n:J\to X$ are analytic functions such that Eq. (1.10) holds, then the results of the Exercise 1.2 continues to hold provided $\dot{f}_n(t)$ and $\dot{f}(t)$ is replaced by $f_n'(z)$ and $f'(z)$ everywhere. In particular, if $\{a_n\}\subset X$ and $\rho>0$ are such that

$$f(z):=\sum_{n=0}^{\infty}a_n(z-z_0)^n\text{ is convergent for }|z-z_0|<\rho,$$

then f is analytic in on $D(z_0,z)$ and

$$f'(z)=\sum_{n=1}^{\infty}na_n(z-z_0)^{n-1}.$$

Corollary 1.12 (Liouville's Theorem). *Suppose that $f:\mathbb{C}\to X$ is a bounded analytic function, then $f(z)=x_0$ for some $x_0\in X$.*

Proof. Let $M:=\sup_{z\in\mathbb{C}}\|f(z)\|$ which is finite by assumption. From Eq. (1.12) with $z_0=0$ and simple estimates it follows that

$$\begin{aligned}\|f'(w)\|&=\left\|\frac{1}{2\pi i}\oint_{\partial D(0,\rho)}\frac{f(z)}{(z-w)^2}dz\right\|\\&=\left\|\frac{1}{2\pi i}\int_{-\pi}^{\pi}\frac{f(\rho e^{i\theta})}{(\rho e^{i\theta}-w)^2}i\rho e^{i\theta}d\theta\right\|\\&\le\frac{M}{2\pi}\max_{|\theta|\le\pi}\frac{\rho}{|\rho e^{i\theta}-w|^2}.\end{aligned}$$

Letting $\rho\uparrow\infty$ in this inequality shows $\|f'(w)\|=0$ for all $w\in\mathbb{C}$ and hence f is constant by FTC or by noting the that power series expansion is $f(w)=f(0)=x_0$.

Alternatively: one can simply apply the standard Liouville's theorem to $\xi\circ f$ for $\xi\in X^*$ in order to show $\xi\circ f(z)=\xi\circ f(0)$ for each $z\in\mathbb{C}$. As $\xi\in X^*$ was arbitrary it follows that $f(z)=f(0)=x_0$ for all $z\in\mathbb{C}$. ■

Exercise 1.3 (Conway, Exr. 4, p. 198 cont.). Let H be a separable Hilbert space. Give an example of a discontinuous function, $f:[0,\infty)\to H$, such that $t\to\langle f(t),h\rangle$ is continuous for all $t\ge0$.

1.3 Bochner Integral (integrands with separable range)

The main results of this section are summarized in the following theorem.

Theorem 1.13. *If we suppose that X is a separable Banach space, then;*

1. *The Borel σ – algebra $(\mathcal{B}_X)$ on X is the same as $\sigma(X^*)$ – the σ – algebra generated X^*.*
2. *The $\|\cdot\|_X$ is then of course $\mathcal{B}_X = \sigma(X^*)$ measurable.*
3. *A function, $u : (\Omega, \mathcal{F}) \to X$, is weakly measurable iff if is $\mathcal{F}/\mathcal{B}_X$ measurable and in which case $\|u(\cdot)\|_X$ is measurable.*
4. *The Pettis integrable functions are now easily describe as*

$$\begin{aligned} L^1_{Pet}(\mu; X) &= L^1(\mu; X) \\ &= \left\{ u : \Omega \to X |\ u \text{ is } \mathcal{F}/\mathcal{B}_X \text{ - meas. \& } \int_\Omega \|u(\cdot)\| \, d\mu < \infty \right\}. \end{aligned}$$

5. *$L^1(\mu; X)$ is complete, i.e. $L^1(\mu; X)$ is a Banach space.*
6. *The dominated convergence theorem holds, i.e. if $\{u_n\} \subset L^1(\mu; X)$ is such that $u(\omega) = \lim_{n\to\infty} u_n(\omega)$ exists for μ-a.e. x and there exists $g \in L^1(\mu)$ such that $\|u_n\|_X \le g$ a.e. for all n, then $u \in L^1(\mu; X)$ and $\lim_{n\to\infty} \|u - u_n\|_1 = 0$ and in particular,*

$$\left\| \int_\Omega u d\mu - \int_\Omega u_n d\mu \right\|_X \le \|u - u_n\|_1 \to 0 \text{ as } n \to \infty.$$

For the rest of this section, X will always be a separable Banach space.

Exercise 1.4 (Differentiate past the integral). Suppose that $J = (a, b) \subset \mathbb{R}$ is a non-empty open interval, $f : J \times \Omega \to X$ is a function such that;

1. for each $t \in J$, $f(t, \cdot) \in L^1(\mu; X)$,
2. for each ω, $J \ni t \to f(t, \omega)$ is a C^1-function.
3. There exists $g \in L^1(\mu)$ such that $\left\| \dot f(t, \omega) \right\|_X \le g(\omega)$ for all ω where $\dot f(t, \omega) := \frac{d}{dt} f(t, \omega)$.

Then $F : J \to X$ defined by

$$F(t) := \int_\Omega f(t, \omega) \, d\mu(\omega)$$

is a C^1-function with

$$\dot F(t) = \int_\Omega \dot f(t, \omega) \, d\mu(\omega).$$

The rest of this section is now essentially devoted to the proof of Theorem 1.13.

1.3.1 Proof of Theorem 1.13

Proposition 1.14. *If X is a separable Banach space, there exists $\{\varphi_n\}_{n=1}^\infty \subset X^*$ such that*

$$\|x\| = \sup_n |\varphi_n(x)| \text{ for all } x \in X. \tag{1.14}$$

Proof. If $\varphi \in X^*$, then $\varphi : X \to \mathbb{R}$ is continuous and hence Borel measurable. Therefore $\sigma(X^*) \subset \mathcal{B}$. For the converse. Choose $x_n \in X$ such that $\|x_n\| = 1$ for all n and

$$\overline{\{x_n\}} = S = \{x \in X : \|X\| = 1\}.$$

By the Hahn Banach Theorem **??** (or Corollary **??** with $x = x_n$ and $M = \{0\}$), there exists $\varphi_n \in X^*$ such that i) $\varphi_n(x_n) = 1$ and ii) $\|\varphi_n\|_{X^*} = 1$ for all n.

As $|\varphi_n(x)| \le \|x\|$ for all n we certainly have $\sup_n |\varphi_n(x)| \le \|x\|$. For the converse inequality, let $x \in X \setminus \{0\}$ and choose $\{n_k\}_{k=1}^\infty \subset \mathbb{N}$ such that $x / \|x\| = \lim_{k\to\infty} x_{n_k}$. It then follows that

$$\left| \varphi_{n_k}\left(\frac{x}{\|x\|} \right) - 1 \right| = \left| \varphi_{n_k}\left(\frac{x}{\|x\|} - x_{n_k} \right) \right| \le \left\| \frac{x}{\|x\|} - x_{n_k} \right\| \to 0 \text{ as } k \to \infty,$$

i.e. $\lim_{k\to\infty} |\varphi_{n_k}(x)| = \|x\|$ which shows $\sup_n |\varphi_n(x)| \ge \|x\|$. ■

Corollary 1.15. *If X is a separable Banach space, then Borel σ – algebra of X and the σ – algebra generated by $\varphi \in X^*$ are the same, i.e. $\sigma(X^*) = \mathcal{B}_X$ – the Borel σ-algebra on X.*

Proof. Since every $\varphi \in X^*$ is continuous it $\mathcal{B}_X$ – measurable and hence $\sigma(X^*) \subset \mathcal{B}_X$. For the converse inclusion, let $\{\varphi_n\}_{n=1}^\infty \subset X^*$ be as in Proposition **??**. We then have for any $x_0 \in X$ that

$$\| \cdot - x_0 \| = \sup_n |\varphi_n(\cdot - x_0)| = \sup_n |\varphi_n(\cdot) - \varphi_n(x_0)|.$$

This shows $\| \cdot - x_0 \|$ is $\sigma(X^*)$–measurable for each $x_0 \in X$ and hence

$$\{x : \|x - x_0\| < \delta\} \in \sigma(X^*).$$

Hence $\sigma(X^*)$ contains all open balls in X. As X is separable, every open set may be written as a countable union of open balls and therefore we may conclude $\sigma(X^*)$ contains all open sets and hence $\mathcal{B}_X \subset \sigma(X^*)$. ■

Corollary 1.16. *If X is a separable Banach space, then a function $u : \Omega \to X$ is $\mathcal{F}/\mathcal{B}_X$ – measurable iff $\lambda \circ u : \Omega \to \mathbb{F}$ is measurable for all $\lambda \in X^*$.*

Proof. This follows directly from Corollary 1.15 of the appendix which asserts that $\sigma(X^*) = \mathcal{B}_X$ when X is separable. ■

Corollary 1.17. *If X is separable and $u_n : \Omega \to X$ are measurable functions such that $u(\omega) := \lim_{n\to\infty} u_n(\omega)$ exists in X for all $\omega \in \Omega$, then $u : \Omega \to X$ is measurable as well.*

Proof. We need only observe that for any $\lambda \in X^*$, $\lambda \circ u = \lim_{n\to\infty} \lambda \circ u_n$ is measurable and hence the result follows from Corollary 1.16. ■

Corollary 1.18. *If $(\Omega, \mathcal{F}, \mu)$ is a measure space and X is a separable Banach space, a function $u : \Omega \to X$ is weakly integrable iff $u : \Omega \to X$ is $\mathcal{F}/\mathcal{B}_X$ – measurable and*

$$\int_\Omega \|u(\omega)\| \, d\mu(\omega) < \infty.$$

Corollary 1.19. *Suppose that $(\Omega, \mathcal{F}, \mu)$ is a measure space and $F, G : \Omega \to X$ are $\mathcal{F}/\mathcal{B}_X$ – measurable functions. Then $F(\omega) = G(\omega)$ for μ – a.e. $\omega \in \Omega$ iff $\varphi \circ F(\omega) = \varphi \circ G(\omega)$ for μ – a.e. $\omega \in \Omega$ and every $\varphi \in X^*$.*

Proof. The direction, "$\Longrightarrow$", is clear. For the converse direction let $\{\varphi_n\} \subset X^*$ be as in Proposition 1.14 and for $n \in \mathbb{N}$, let

$$E_n := \{\omega \in \Omega : \varphi_n \circ F(\omega) \neq \varphi_n \circ G(\omega)\}.$$

By assumption $\mu(E_n) = 0$ and therefore $E := \cup_{n=1}^\infty E_n$ is a μ – null set as well. This completes the proof since $\varphi_n(F - G) = 0$ on E^c and therefore, by Eq. (1.14)

$$\|F - G\| = \sup_n |\varphi_n(F - G)| = 0 \text{ on } E^c.$$

■

Recall that we have already seen in this case that the Borel σ – field $\mathcal{B}$ on X is the same as the σ – field $(\sigma(X^*))$ which is generated by X^* – the continuous linear functionals on X. As a consequence $F : \Omega \to X$ is $\mathcal{F}/\mathcal{B}$ measurable iff $\varphi \circ F : \Omega \to \mathbb{R}$ is $\mathcal{F}/\mathcal{B}(\mathbb{R})$ – measurable for all $\varphi \in X^*$. In particular it follows that if $F, G : \Omega \to X$ are measurable functions then so is $F + G$ and λF for all $\lambda \in \mathbb{F}$ and it follows that $\{F \neq G\} = \{F - G \neq 0\}$ is measurable as well. Also note that $\|\cdot\| : X \to [0, \infty)$ is continuous and hence measurable and hence $\omega \to \|F(\omega)\|_X$ is the composition of two measurable functions and therefore measurable.

Definition 1.20. *For $1 \le p < \infty$ let $L^p(\mu; X)$ denote the space of measurable functions $F : \Omega \to X$ such that $\int_\Omega \|F\|^p d\mu < \infty$. For $F \in L^p(\mu; X)$, define*

$$\|F\|_{L^p} = \left(\int_\Omega \|F\|_X^p d\mu \right)^{\frac{1}{p}}.$$

As usual in L^p – spaces we will identify two measurable functions, $F, G : \Omega \to X$, if $F = G$ a.e.

Lemma 1.21. *Suppose $a_n \in X$ and $\|a_{n+1} - a_n\| \le \varepsilon_n$ and $\sum_{n=1}^\infty \varepsilon_n < \infty$. Then $\lim_{n\to\infty} a_n = a \in X$ exists and $\|a - a_n\| \le \delta_n := \sum_{k=n}^\infty \varepsilon_k$.*

Proof. Let $m > n$ then

$$\|a_m - a_n\| = \left\| \sum_{k=n}^{m-1} (a_{k+1} - a_k) \right\| \le \sum_{k=n}^{m-1} \|a_{k+1} - a_k\| \le \sum_{k=n}^\infty \varepsilon_k := \delta_n. \quad (1.15)$$

So $\|a_m - a_n\| \le \delta_{\min(m,n)} \to 0$ as $, m, n \to \infty$, i.e. $\{a_n\}$ is Cauchy. Let $m \to \infty$ in (1.15) to find $\|a - a_n\| \le \delta_n$. ■

Lemma 1.22. *Suppose that $\{F_n\}$ is Cauchy in measure, i.e. $\lim_{m,n\to\infty} \mu(\|F_n - F_m\| \ge \varepsilon) = 0$ for all $\varepsilon > 0$. Then there exists a subsequence $G_j = F_{n_j}$ such that $F := \lim_{j\to\infty} G_j$ exists μ – a.e. and moreover $F_n \xrightarrow{\mu} F$ as $n \to \infty$, i.e. $\lim_{n\to\infty} \mu(\|F_n - F\| \ge \varepsilon) = 0$ for all $\varepsilon > 0$.*

Proof. Let $\varepsilon_n > 0$ such that $\sum_{n=1}^\infty \varepsilon_n < \infty$ ($\varepsilon_n = 2^{-n}$ would do) and set $\delta_n = \sum_{k=n}^\infty \varepsilon_k$. Choose $G_j = F_{n_j}$ where $\{n_j\}$ is a subsequence of $\mathbb{N}$ such that

$$\mu(\{\|G_{j+1} - G_j\| > \varepsilon_j\}) \le \varepsilon_j.$$

Let

$$A_N := \cup_{j \ge N} \{\|G_{j+1} - G_j\| > \varepsilon_j\} \text{ and}$$
$$E := \cap_{N=1}^\infty A_N = \{\|G_{j+1} - G_j\| > \varepsilon_j \text{ i.o.}\}.$$

Since $\mu(A_N) \le \delta_N < \infty$ and $A_N \downarrow E$ it follows[1] that $0 = \mu(E) = \lim_{N\to\infty} \mu(A_N)$. For $\omega \notin E$, $\|G_{j+1}(\omega) - G_j(\omega)\| \le \varepsilon_j$ for a.a. j and hence by Lemma 1.21, $F(\omega) := \lim_{j\to\infty} G_j(\omega)$ exists for $\omega \notin E$. Let us define $F(\omega) = 0$ for all $\omega \in E$.

Next we will show $G_N \xrightarrow{\mu} F$ as $N \to \infty$ where F and G_N are as above. If

$$\omega \in A_N^c = \cap_{j \ge N} \{\|G_{j+1} - G\| \le \varepsilon_j\},$$

then

$$\|G_{j+1}(\omega) - G_j(\omega)\| \le \varepsilon_j \text{ for all } j \ge N.$$

Another application of Lemma 1.21 shows $\|F(\omega) - G_j(\omega)\| \le \delta_j$ for all $j \ge N$, i.e.

[1] Alternatively, $\mu(E) = 0$ by the first Borel Cantelli lemma and the fact that $\sum_{j=1}^\infty \mu(\{\|G_{j+1} - G_j\| > \varepsilon_j\}) \le \sum_{j=1}^\infty \varepsilon_j < \infty$.

$$A_N^c \subset \cap_{j\ge N}\{\|F-G_j\|\le \delta_j\} \subset \{|F-G_N|\le \delta_N\}.$$

Therefore, by taking complements of this equation, $\{\|F-G_N\|>\delta_N\}\subset A_N$ and hence

$$\mu(\|F-G_N\|>\delta_N)\le \mu(A_N)\le \delta_N \to 0 \text{ as } N\to\infty$$

and in particular, $G_N \overset{\mu}{\to} F$ as $N\to\infty$.

With this in hand, it is straightforward to show $F_n \overset{\mu}{\to} F$. Indeed, by the usual trick, for all $j\in\mathbb{N}$,

$$\mu(\{\|F_n-F\|>\varepsilon\})\le \mu(\{\|F-G_j\|>\varepsilon/2\})+\mu(\|G_j-F_n\|>\varepsilon/2).$$

Therefore, letting $j\to\infty$ in this inequality gives,

$$\mu(\{\|F_n-F\|>\varepsilon\})\le \limsup_{j\to\infty}\mu(\|G_j-F_n\|>\varepsilon/2)\to 0 \text{ as } n\to\infty,$$

wherein we have used $\{F_n\}_{n=1}^\infty$ is Cauchy in measure and $G_j \overset{\mu}{\to} F$. ■

Theorem 1.23. *For each $p\in[0,\infty)$, the space $(L^p(\mu;X),\|\cdot\|_{L^p})$ is a Banach space.*

Proof. It is straightforward to check that $\|\cdot\|_{L^p}$ is a norm. For example,

$$\begin{aligned}\|F+G\|_{L^p} &= \left(\int_\Omega \|F+G\|_X^p d\mu\right)^{\frac1p} \le \left(\int_\Omega (\|F\|_X+\|G\|_X)^p d\mu\right)^{\frac1p}\\ &\le \|F\|_{L^p}+\|G\|_{L^p}.\end{aligned}$$

So the main point is to prove completeness of the norm.

Let $\{F_n\}_{n=1}^\infty\subset L^p(\mu)$ be a Cauchy sequence. By Chebyshev's inequality $\{F_n\}$ is Cauchy in measure and by Lemma 1.22 there exists a subsequence $\{G_j\}$ of $\{F_n\}$ such that $G_j\to F$ a.e. By Fatou's Lemma,

$$\begin{aligned}\|G_j-F\|_p^p &= \int_\Omega \lim_{k\to\infty}\inf\|G_j-G_k\|^p d\mu \le \lim_{k\to\infty}\inf\int_\Omega \|G_j-G_k\|^p d\mu\\ &= \lim_{k\to\infty}\inf\|G_j-G_k\|_p^p\to 0 \text{ as } j\to\infty.\end{aligned}$$

In particular, $\|F\|_p\le \|G_j-F\|_p+\|G_j\|_p<\infty$ so the $F\in L^p$ and $G_j \xrightarrow{L^p} F$. The proof is finished because,

$$\|F_n-F\|_p\le \|F_n-G_j\|_p+\|G_j-F\|_p\to 0 \text{ as } j,n\to\infty.$$

■

Definition 1.24 (Simple functions). *We say a function $F:\Omega\to X$ is a simple function if F is measurable and has finite range. If F also satisfies, $\mu(F\ne 0)<\infty$ we say that F is a μ – simple function and let $\mathcal{S}(\mu;X)$ denote the vector space of μ – simple functions.*

Proposition 1.25. *For each $1\le p<\infty$ the μ – simple functions, $\mathcal{S}(\mu;X)$, are dense inside of $L^p(\mu;X)$.*

Proof. Let $\mathbb{D}:=\{x_n\}_{n=1}^\infty$ be a countable dense subset of $X\setminus\{0\}$. For each $\varepsilon>0$ and $n\in\mathbb{N}$ let

$$B_n^\varepsilon := \left\{x\in X:\|x-x_n\|\le \min\left(\varepsilon,\frac12\|x_n\|\right)\right\}$$

and then define $A_n^\varepsilon := B_n^\varepsilon\setminus(\cup_{k=1}^n B_k^\varepsilon)$. Thus $\{A_n^\varepsilon\}_{n=1}^\infty$ is a partition of $X\setminus\{0\}$ with the added property that $\|y-x_n\|\le\varepsilon$ and $\frac12\|x_n\|\le\|y\|\le\frac32\|x_n\|$for all $y\in A_n^\varepsilon$.

Given $F\in L^p(\mu;X)$ let

$$F_\varepsilon := \sum_{n=1}^\infty x_n\cdot 1_{F\in A_n^\varepsilon} = \sum_{n=1}^\infty x_n\cdot 1_{F^{-1}(A_n^\varepsilon)}.$$

For $\omega\in F^{-1}(A_n^\varepsilon)$, i.e. $F(\omega)\in A_n^\varepsilon$, we have

$$\|F_\varepsilon(\omega)\| = \|x_n\|\le 2\|F(\omega)\| \text{ and}$$
$$\|F_\varepsilon(\omega)-F(\omega)\| = \|x_n-F(\omega)\|\le\varepsilon.$$

Putting these two estimates together shows,

$$\|F_\varepsilon-F\|\le\varepsilon \text{ and } \|F_\varepsilon-F\|\le\|F_\varepsilon\|+\|F\|\le 3\|F\|.$$

Hence we may now apply the dominated convergence theorem in order to show

$$\lim_{\varepsilon\downarrow 0}\|F-F_\varepsilon\|_{L^p(\mu;X)}=0.$$

As the F_ε – have countable range we have not yet completed the proof. To remedy this defect, to each $N\in\mathbb{N}$ let

$$F_\varepsilon^N := \sum_{n=1}^N x_n\cdot 1_{F^{-1}(A_n^\varepsilon)}.$$

Then it is clear that $\lim_{N\to\infty}F_\varepsilon^N=F_\varepsilon$ and that $\left\|F_\varepsilon^N\right\|\le\|F_\varepsilon\|\le 2\|F\|$ for all N. Therefore another application of the dominated convergence theorem implies, $\lim_{N\to\infty}\left\|F_\varepsilon^N-F_\varepsilon\right\|_{L^p(\mu;X)}=0$. Thus any $F\in L^p(\mu;X)$ may be arbitrarily well approximated by one of the $F_\varepsilon^N\in\mathcal{S}(\mu;X)$ with ε sufficiently small and N sufficiently large. ■

For later purposes it will be useful to record a result based on the partitions $\{A_n^\varepsilon\}_{n=1}^\infty$ of $X\setminus\{0\}$ introduced in the above proof.

Lemma 1.26. *Suppose that $F : \Omega \to X$ is a measurable function such that $\mu(F \neq 0) > 0$. Then there exists $B \in \mathcal{F}$ and $\varphi \in X^*$ such that $\mu(B) > 0$ and $\inf_{\omega \in B} \varphi \circ F(\omega) > 0$.*

Proof. Let $\varepsilon > 0$ be chosen arbitrarily, for example you might take $\varepsilon = 1$ and let $\{A_n := A_n^\varepsilon\}_{n=1}^\infty$ be the partition of $X \setminus \{0\}$ introduced in the proof of Proposition 1.25 above. Since $\{F \neq 0\} = \sum_{n=1}^\infty \{F \in A_n\}$ and $\mu(F \neq 0) > 0$, it follows that that $\mu(F \in A_n) > 0$ for some $n \in \mathbb{N}$. We now let $B := \{F \in A_n\} = F^{-1}(A_n)$ and choose $\varphi \in X^*$ such that $\varphi(x_n) = \|x_n\|$ and $\|\varphi\|_{X^*} = 1$. For $\omega \in B$ we have $F(\omega) \in A_n$ and therefore $\|F(\omega) - x_n\| \leq \frac{1}{2}\|x_n\|$ and hence,

$$|\varphi(F(\omega)) - \|x_n\|| = |\varphi(F(\omega)) - \varphi(x_n)| \leq \|\varphi\|_{X^*} \|F(\omega) - x_n\| \leq \frac{1}{2}\|x_n\|.$$

From this inequality we see that $\varphi(F(\omega)) \geq \frac{1}{2}\|x_n\| > 0$ for all $\omega \in B$. ∎

Definition 1.27. *To each $F \in \mathcal{S}(\mu; X)$, let*

$$\begin{aligned} I(F) &= \sum_{x \in X} x\mu(F^{-1}(\{x\})) = \sum_{x \in X} x\mu(\{F = x\}) \\ &= \sum_{x \in F(\Omega)} x\mu(F = x) \in X. \end{aligned}$$

The following proposition is straightforward to prove.

Proposition 1.28. *The map $I : \mathcal{S}(\mu; X) \to X$ is linear and satisfies for all $F \in \mathcal{S}(\mu; X)$,*

$$\|I(F)\|_X \leq \int_\Omega \|F\| d\mu \text{ and} \tag{1.16}$$

$$\varphi(I(F)) = \int_\Omega \varphi \circ F \, d\mu \; \forall \varphi \in X^*. \tag{1.17}$$

More generally, if $T \in B(X, Y)$ where Y is another Banach space then

$$TI(F) = I(TF).$$

Proof. If $0 \neq c \in \mathbb{R}$ and $F \in \mathcal{S}(\mu; X)$, then

$$\begin{aligned} I(cF) &= \sum_{x \in X} x\mu(cF = x) = \sum_{x \in X} x\mu\left(F = \frac{x}{c}\right) \\ &= \sum_{y \in X} cy \; \mu(F = y) = cI(F) \end{aligned}$$

and if $c = 0$, $I(0F) = 0 = 0I(F)$. If $F, G \in \mathcal{S}(\mu; X)$,

$$\begin{aligned} I(F + G) &= \sum_x x\mu(F + G = x) \\ &= \sum_x x \sum_{y+z=x} \mu(F = y, G = z) \\ &= \sum_{y,z} (y + z)\mu(F = y, G = z) \\ &= \sum_y y\mu(F = y) + \sum_z z\mu(G = z) = I(F) + I(G). \end{aligned}$$

Equation (1.16) is a consequence of the following computation:

$$\|I(F)\|_X = \|\sum_{x \in X} x\mu(F = x)\| \leq \sum_{x \in X} \|x\| \mu(F = x) = \int_\Omega \|F\| d\mu$$

and Eq. (1.17) follows from:

$$\begin{aligned} \varphi(I(F)) &= \varphi(\sum_{x \in X} x\mu(\{F = x\})) \\ &= \sum_{x \in X} \varphi(x)\,\mu(\{F = x\}) = \int_\Omega \varphi \circ F \, d\mu. \end{aligned}$$

∎

The next elementary theorem (referred to as the bounded linear transformation theorem, or B.L.T. theorem for short) is often useful when constructing bounded linear transformations.

Theorem 1.29 (B. L. T. Theorem). *Suppose that Z is a normed space, X is a Banach space, and $\mathcal{S} \subset Z$ is a dense linear subspace of Z. If $T : \mathcal{S} \to X$ is a bounded linear transformation (i.e. there exists $C < \infty$ such that $\|Tz\| \leq C\|z\|$ for all $z \in \mathcal{S}$), then T has a unique extension to an element $\bar{T} \in L(Z, X)$ and this extension still satisfies*

$$\|\bar{T}z\| \leq C\|z\| \text{ for all } z \in \bar{\mathcal{S}}.$$

Proof. The proof is left to the reader. ∎

Theorem 1.30 (Bochner Integral). *There is a unique continuous linear map $\bar{I} : L^1(\Omega, \mathcal{F}, \mu; X) \to X$ such that $\bar{I}|_{\mathcal{S}(\mu;X)} = I$ where I is defined in Definition 1.27. Moreover, for all $F \in L^1(\Omega, \mathcal{F}, \mu; X)$,*

$$\|\bar{I}(F)\|_X \leq \int_\Omega \|F\| d\mu \tag{1.18}$$

and $\bar{I}(F)$ is the unique element in X such that

$$\varphi(\bar{I}(F)) = \int_\Omega \varphi \circ F \; d\mu \; \forall \varphi \in X^*. \tag{1.19}$$

The map $\bar{I}(F)$ will be denoted suggestively by $\int_\Omega F d\mu$ or $\mu(F)$ so that Eq. (1.19) may be written as

$$\varphi\left(\int_\Omega F d\mu\right) = \int_\Omega \varphi \circ F \; d\mu \; \forall \varphi \in X^* \text{ or}$$
$$\varphi(\mu(F)) = \mu(\varphi \circ F) \;\; \forall \varphi \in X^*$$

It is also true that if $T \in B(X,Y)$ where Y is another Banach space, then

$$\int_\Omega TF d\mu = T \int_\Omega F d\mu$$

where one should interpret $TF : \Omega \to \overline{TX}$ which is a separable subspace of Y even is Y is not separable.

Proof. The existence of a continuous linear map $\bar{I} : L^1(\Omega, \mathcal{F}, \mu; X) \to X$ such that $\bar{I}|_{\mathcal{S}(\mu;X)} = I$ and Eq. (1.18) holds follows from Propositions 1.28 and 1.25 and the bounded linear transformation Theorem 1.29. If $\varphi \in X^*$ and $F \in L^1(\Omega, \mathcal{F}, \mu; X)$, choose $F_n \in \mathcal{S}(\mu; X)$ such that $F_n \to F$ in $L^1(\Omega, \mathcal{F}, \mu; X)$ as $n \to \infty$. Then $\bar{I}(F) = \lim_{n\to\infty} I(F_n)$ and hence by Eq. (1.17),

$$\varphi(\bar{I}(F)) = \varphi(\lim_{n\to\infty} I(F_n)) = \lim_{n\to\infty} \varphi(I(F_n)) = \lim_{n\to\infty} \int_\Omega \varphi \circ F_n \; d\mu.$$

This proves Eq. (1.19) since

$$\begin{aligned}\left|\int_\Omega (\varphi \circ F - \varphi \circ F_n) d\mu\right| &\le \int_\Omega |\varphi \circ F - \varphi \circ F_n| \, d\mu \\ &\le \int_\Omega \|\varphi\|_{X^*} \|\varphi \circ F - \varphi \circ F_n\|_X \, d\mu \\ &= \|\varphi\|_{X^*} \|F - F_n\|_{L^1} \to 0 \text{ as } n \to \infty.\end{aligned}$$

The fact that $\bar{I}(F)$ is determined by Eq. (1.19) is a consequence of the Hahn – Banach theorem. ■

Example 1.31. Suppose that $x \in X$ and $f \in L^1(\mu; \mathbb{R})$, then $F(\omega) := f(\omega) x$ defines an element of $L^1(\mu; X)$ and

$$\int_\Omega F d\mu = \left(\int_\Omega f d\mu\right) x. \tag{1.20}$$

To prove this just observe that $\|F\| = |f| \, \|x\| \in L^1(\mu)$ and for $\varphi \in X^*$ we have

$$\begin{aligned}\varphi\left(\left(\int_\Omega f d\mu\right) x\right) &= \left(\int_\Omega f d\mu\right) \cdot \varphi(x) \\ &= \left(\int_\Omega f \varphi(x) \, d\mu\right) = \int_\Omega \varphi \circ F \; d\mu.\end{aligned}$$

Since $\varphi\left(\int_\Omega F d\mu\right) = \int_\Omega \varphi \circ F \; d$ for all $\varphi \in X^*$ it follows that Eq. (1.20) is correct.

Definition 1.32 (Essential Range). *Suppose that $(\Omega, \mathcal{F}, \mu)$ is a measure space, (Y, ρ) is a metric space, and $q : \Omega \to Y$ is a measurable function. We then define the* ***essential range*** *of q to be the set,*

$$\mathrm{essran}_\mu(q) = \{y \in Y : \mu(\{\rho(q,y) < \varepsilon\}) > 0 \;\; \forall \varepsilon > 0\}.$$

In other words, $y \in Y$ is in $\mathrm{essran}_\mu(q)$ iff q lies in $B_\rho(y, \varepsilon)$ with positive μ – measure.

Remark 1.33. The separability assumption on X may be relaxed by assuming that $F : \Omega \to X$ has separable essential range. In this case we may still define $\int_\Omega F d\mu$ by applying the above formalism with X replaced by the separable Banach space, $X_0 := \overline{\mathrm{span}(\mathrm{essran}_\mu(F))}$. For example if Ω is a compact topological space and $F : \Omega \to X$ is a continuous map, then $\int_\Omega F d\mu$ is always defined.

Theorem 1.34 (DCT). *If $\{u_n\} \subset L^1(\mu; X)$ is such that $u(\omega) = \lim_{n\to\infty} u_n(\omega)$ exists for μ-a.e. x and there exists $g \in L^1(\mu)$ such that $\|u_n\|_X \le g$ a.e. for all n, then $u \in L^1(\mu; X)$ and $\lim_{n\to\infty} \|u - u_n\|_1 = 0$ and in particular,*

$$\left\|\int_\Omega u d\mu - \int_\Omega u_n d\mu\right\|_X \le \|u - u_n\|_1 \to 0 \text{ as } n \to \infty.$$

Proof. Since $\|u(\omega)\|_X = \lim_{n\to\infty} \|u_n(\omega)\| \le g(\omega)$ for a.e. ω, it follows that $u \in L^1(\mu, X)$. Moreover, $\|u - u_n\|_X \le 2g$ a.e. and $\lim_{n\to\infty} \|u - u_n\|_X = 0$ a.e. and therefore by the real variable dominated convergence theorem it follows that

$$\|u - u_n\|_1 = \int_\Omega \|u - u_n\|_X \, d\mu \to 0 \text{ as } n \to \infty.$$

■

1.4 Strong Bochner Integrals

Let us again assume that X is a separable Banach space but now suppose that $C : \Omega \to B(X)$ is the type of function we wish to integrate. As $B(X)$ is typically not separable, we can not directly apply the theory of the last section. However, there is an easy solution which will briefly describe here.

Definition 1.35. *We say $C : \Omega \to B(X)$ is strongly measurable if $\Omega \ni \omega \to C(\omega)x$ is measurable for all $x \in X$.*

Lemma 1.36. *If $C : \Omega \to B(X)$ is strongly measurable, then $\Omega \ni \omega \to \|C(\omega)\|_{op}$ is measurable.*

Proof. Let $\mathbb{D}$ be a dense subset of the unit vectors in X. Then

$$\|C(\omega)\|_{op} = \sup_{x\in\mathbb{D}} \|C(\omega)x\|_X$$

is measurable. ∎

Lemma 1.37. *Suppose that $u : \Omega \to X$ is measurable and $C : \Omega \to B(X)$ is strongly measurable, then $\Omega \ni \omega \to C(\omega)u(\omega) \in X$ is measurable.*

Proof. Using the ideas in Proposition 1.25 we may find simple functions $u_n : \Omega \to X$ so that $u = \lim_{n\to\infty} u_n$. It is easy to verify that $C(\cdot)u_n(\cdot)$ is measurable for all n and that $C(\cdot)u(\cdot) = \lim_{n\to\infty} C(\cdot)u_n(\cdot)$. The result now follows Corollary 1.17. ∎

Corollary 1.38. *Suppose $C, D : \Omega \to B(X)$ are strongly measurable, then $\Omega \ni \omega \to C(\omega)D(\omega) \in X$ is strongly measurable.*

Proof. For $x \in X$, let $u(\omega) := D(\omega)x$ which is measurable by assumption. Therefore, $C(\cdot)D(\cdot)x = C(\cdot)u(\cdot)$ is measurable by Lemma 1.37. ∎

Definition 1.39. *We say $C : \Omega \to B(X)$ is* ***integrable*** *and write $C \in L^1(\mu : B(X))$ if C is strongly measurable and*

$$\|C\|_1 := \int_\Omega \|C(\omega)\| d\mu(\omega) < \infty.$$

In this case we further define $\mu(C) = \int_\Omega C(\omega)d\mu(\omega)$ to be the unique element $B(X)$ such that

$$\mu(C)x = \int_\Omega C(\omega)xd\mu(\omega) \text{ for all } x \in X.$$

It is easy to verify that this integral again has all of the usual properties of integral. In particular,

$$\|\mu(C)x\| \le \int_\Omega \|C(\omega)x\| d\mu(\omega) \le \int_\Omega \|C(\omega)\| \|x\| d\mu(\omega) = \|C\|_1 \|x\|$$

from which it follows that $\|\mu(C)\|_{op} \le \|C\|_1$.

Theorem 1.40. *Suppose that $\left(\tilde\Omega, \nu\right)$ is another measure space and $D \in L^1(\tilde\mu : B(X))$. Then*

$$\mu(C)\nu(D) = \mu\otimes\nu(C\otimes D)$$

where $\mu\otimes\nu$ is product measure and

$$C\otimes D(\omega,\tilde\omega) := C(\omega)D(\tilde\omega).$$

Proof. Let $\pi_1 : \Omega\times\tilde\Omega \to \Omega$ and $\pi_2 : \Omega\times\tilde\Omega \to \tilde\Omega$ be the natural projection maps. Since $C\otimes D = [C\circ\pi_1][D\circ\pi_2]$, we conclude from Corollary 1.38 that $C\otimes D$ is measurable on the product space. We further have

$$\begin{aligned}
&\int_{\Omega\times\tilde\Omega} \|C\otimes D(\omega,\tilde\omega)\|_{op} d\mu(\omega)d\nu(\tilde\omega) \\
&\quad= \int_{\Omega\times\tilde\Omega} \|C(\omega)D(\tilde\omega)\|_{op} d\mu(\omega)d\nu(\tilde\omega) \\
&\quad\le \int_{\Omega\times\tilde\Omega} \|C(\omega)\|_{op}\|D(\tilde\omega)\|_{op} d\mu(\omega)d\nu(\tilde\omega) \\
&\quad= \int_\Omega \|C(\omega)\|_{op} d\mu(\omega)\cdot\int_{\tilde\Omega}\|D(\tilde\omega)\|_{op} d\nu(\tilde\omega) < \infty
\end{aligned}$$

and therefore $\mu\otimes\nu(C\otimes D)$ is well defined.

Now suppose that $x \in X$ and let u_n be simple function in $L^1\left(\tilde\Omega,\nu\right)$ such that $\lim_{n\to\infty}\|u_n - D(\cdot)x\|_{L^1(\nu)} = 0$. If $u_n = \sum_{k=0}^{M_n} a_k 1_{A_k}$ with $\{A_k\}_{k=1}^{M_n}$ being disjoint subsets of $\tilde\Omega$ and $a_k \in X$, then

$$C(\omega)u_n(\tilde\omega) = \sum_{k=0}^{M_n} 1_{A_k}(\tilde\omega)C(\omega)a_k.$$

After another approximation argument for $\omega \to C(\omega)a_k$, we find,

$$\begin{aligned}
\int_{\Omega\times\tilde\Omega} C(\omega)u_n(\tilde\omega)d[\mu\otimes\nu](\omega,\tilde\omega) &= \sum_{k=0}^{M_n}\nu(A_k)\int_\Omega C(\omega)a_k d\mu(w) \\
&= \sum_{k=0}^{M_n}\nu(A_k)\mu(C)a_k \\
&= \mu(C)\sum_{k=0}^{M_n}\nu(A_k)a_k = \mu(C)\nu(\mu_n).
\end{aligned} \tag{1.21}$$

Since,

$$\begin{aligned}\int_{\Omega\times\tilde{\Omega}} &\|C(\omega)\,u_n(\tilde{\omega}) - C(\omega)\,D(\tilde{\omega})\,x\|\,d[\mu\otimes\nu](\omega,\tilde{\omega})\\ &\le \int_{\Omega\times\tilde{\Omega}} \|C(\omega)\|_{op}\,\|u_n(\tilde{\omega}) - D(\tilde{\omega})\,x\|\,d\mu(\omega)\,d\nu(\tilde{\omega})\\ &= \|C\|_1\cdot\|u_n - D(\cdot)\,x\|_{L^1(\nu)} \to 0 \text{ as } n\to\infty,\end{aligned}$$

we may pass to the limit in Eq. (1.21) in order to find

$$\begin{aligned}\mu\otimes\nu\,(C\otimes D)\,x &= \int_{\Omega\times\tilde{\Omega}} C(\omega)\,D(\tilde{\omega})\,x\;d[\mu\otimes\nu](\omega,\tilde{\omega})\\ &= \mu(C)\int_{\tilde{\Omega}} D(\tilde{\omega})\,x d\nu(\tilde{\omega}) = \mu(C)\,\nu(D)\,x.\end{aligned}$$

As $x\in X$ was arbitrary the proof is complete. ■

Exercise 1.5. Suppose that U is an open subset of $\mathbb{R}$ or $\mathbb{C}$ and $F: U\times\Omega\to X$ is a measurable function such that;

1. $U\ni z\to F(z,\omega)$ is (complex) differentiable for all $\omega\in\Omega$.
2. $F(z,\cdot)\in L^1(\mu:X)$ for all $z\in U$.
3. There exists $G\in L^1(\mu:\mathbb{R})$ such that

$$\left\|\frac{\partial F(z,\omega)}{\partial z}\right\| \le G(\omega) \text{ for all } (z,\omega)\in U\times\Omega.$$

Show

$$U\ni z\to\int_\Omega F(z,\omega)\,d\mu(\omega)\in X$$

is differentiable and

$$\frac{d}{dz}\int_\Omega F(z,\omega)\,d\mu(\omega) = \int_\Omega \frac{\partial F(z,\omega)}{\partial z}d\mu(\omega)\,.$$

1.5 Weak integrals for Hilbert Spaces

This section may be read independently of the previous material of this chapter. Although you should still learn about the fundamental theorem of calculus in Section **??** above at least for Hilbert space valued functions.

In this section, let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$, H be a separable Hilbert space over $\mathbb{F}$, and $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ be two σ – finite measures spaces.

Definition 1.41. *A function $\psi: X\to H$ is said to be* ***weakly measurable*** *if $X\ni x\to\langle h,\psi(x)\rangle\in\mathbb{F}$ is $\mathcal{M}$ – measurable for all $h\in H$.*

Notice that if ψ is weakly measurable, then $\|\psi(\cdot)\|$ is measurable as well. Indeed, if D is a countable dense subset of $H\setminus\{0\}$, then

$$\|\psi(x)\| = \sup_{h\in D}\frac{|\langle h,\psi(x)\rangle|}{\|h\|}.$$

Definition 1.42. *A function $\psi: X\to H$ is* ***weakly-integrable*** *if ψ is weakly measurable and*

$$\|\psi\|_1 := \int_X \|\psi(x)\|\,d\mu(x) < \infty.$$

We let $L^1(X,\mu:H)$ denote the space of weakly integrable functions.

For $\psi\in L^1(X,\mu:H)$, let

$$f_\psi(h) := \int_X \langle h,\psi(x)\rangle\,d\mu(x)$$

and notice that $f_\psi\in H^*$ with

$$|f_\psi(h)| \le \int_X |\langle h,\psi(x)\rangle|\,d\mu(x) \le \|h\|_H\int_X \|\psi(x)\|_H\,d\mu(x) = \|\psi\|_1\cdot\|h\|_H\,.$$

Thus by the Riesz theorem, there exists a unique element $\bar{\psi}\in H$ such that

$$\left\langle h,\bar{\psi}\right\rangle = f_\psi(h) = \int_X \langle h,\psi(x)\rangle\,d\mu(x) \text{ for all } h\in H.$$

We will denote this element, $\bar{\psi}$, as

$$\bar{\psi} = \int_X \psi(x)\,d\mu(x)\,.$$

Theorem 1.43. *There is a unique linear map,*

$$L^1(X,\mu:H)\ni\psi\to\int_X\psi(x)\,d\mu(x)\in H,$$

such that

$$\left\langle h,\int_X\psi(x)\,d\mu(x)\right\rangle = \int_X\langle h,\psi(x)\rangle\,d\mu(x) \text{ for all } h\in H.$$

Moreover this map satisfies;

1.
$$\left\|\int_X \psi(x)\, d\mu(x)\right\|_H \le \|\psi\|_{L^1(\mu:H)}.$$

2. If $B \in L(H,K)$ is a bounded linear operator from H to K, then
$$B\int_X \psi(x)\, d\mu(x) = \int_X B\psi(x)\, d\mu(x).$$

3. If $\{e_n\}_{n=1}^\infty$ is any orthonormal basis for H, then
$$\int_X \psi(x)\, d\mu(x) = \sum_{n=1}^\infty \left[\int_X \langle \psi(x), e_n\rangle\, d\mu(x)\right] e_n.$$

Proof. We take each item in turn.

1. We have
$$\begin{aligned}\left\|\int_X \psi(x)\, d\mu(x)\right\|_H &= \sup_{\|h\|=1}\left|\left\langle h, \int_X \psi(x)\, d\mu(x)\right\rangle\right| \\ &= \sup_{\|h\|=1}\left|\int_X \langle h, \psi(x)\rangle\, d\mu(x)\right| \le \|\psi\|_1.\end{aligned}$$

2. If $k \in K$, then
$$\begin{aligned}\left\langle B\int_X \psi(x)\, d\mu(x), k\right\rangle &= \left\langle \int_X \psi(x)\, d\mu(x), B^*k\right\rangle = \int_X \langle \psi(x), B^*k\rangle\, d\mu(x) \\ &= \int_X \langle B\psi(x), k\rangle\, d\mu(x) = \left\langle \int_X B\psi(x)\, d\mu(x), k\right\rangle\end{aligned}$$
and this suffices to verify item 2.

3. Lastly,
$$\begin{aligned}\int_X \psi(x)\, d\mu(x) &= \sum_{n=1}^\infty \left\langle \int_X \psi(x)\, d\mu(x), e_n\right\rangle e_n \\ &= \sum_{n=1}^\infty \left[\int_X \langle \psi(x), e_n\rangle\, d\mu(x)\right] e_n.\end{aligned}$$

■

Definition 1.44. *A function $C : (X, \mathcal{M}, \mu) \to B(H)$ is said to be a **weakly measurable** operator if $x \to \langle C(x) v, w\rangle \in \mathbb{C}$ is measurable for all $v, w \in H$.*

Again if C is weakly measurable, then
$$X \ni x \to \|C(x)\|_{op} := \sup_{h,k\in D} \frac{|\langle C(x) h, k\rangle|}{\|h\|\cdot\|k\|}$$
is measurable as well.

Definition 1.45. *A function $C : X \to B(H)$ is **weakly-integrable** if C is weakly measurable and*
$$\|C\|_1 := \int_X \|C(x)\|\, d\mu(x) < \infty.$$
We let $L^1(X, \mu : B(H))$ denote the space of weakly integrable $B(H)$-valued functions.

Theorem 1.46. *If $C \in L^1(\mu : B(H))$, then there exists a unique $\bar{C} \in B(H)$ such that*
$$\bar{C}v = \int_X [C(x) v]\, d\mu(x) \text{ for all } v \in H \tag{1.22}$$
and $\|\bar{C}\| \le \|C\|_1$.

Proof. By very definition, $X \ni x \to C(x) v \in H$ is weakly measurable for each $v \in H$ and moreover
$$\int_X \|C(x) v\|\, d\mu(x) \le \int_X \|C(x)\|\, \|v\|\, d\mu(x) = \|C\|_1 \|v\| < \infty. \tag{1.23}$$
Therefore the integral in Eq. (1.22) is well defined. By the linearity of the weak integral on H – valued functions one easily checks that $\bar{C} : H \to H$ defined by Eq. (1.22) is linear and moreover by Eq. (1.23) we have
$$\|\bar{C}v\| \le \int_X \|C(x) v\|\, d\mu(x) \le \|C\|_1 \|v\|$$
which implies $\|\bar{C}\| \le \|C\|_1$. ■

Notation 1.47 (Weak Integrals) *We denote the $\bar{C}$ in Theorem 1.46 by either $\mu(C)$ or $\int_X C(x)\, d\mu(x)$.*

Theorem 1.48. *Let $C \in L^1(\mu : B(H))$. The weak integral, $\mu(C)$, has the following properties;*

1. *$\|\mu(C)\|_{op} \le \|C\|_1$.*
2. *For all $v, w \in H$,*
$$\langle \mu(C) v, w\rangle = \left\langle \int_X C(x)\, d\mu(x) v, w\right\rangle = \int_X \langle C(x) v, w\rangle\, d\mu(x).$$

3. $\mu(C^*) = \mu(C)^*$ *, i.e.*

$$\int_X C(x)^* \, d\mu(x) = \left[\int_X C(x) \, d\mu(x)\right]^*.$$

4. If $\{e_i\}_{i=1}^\infty$ *is an orthonormal basis for* H*, then*

$$\mu(C) v = \sum_{i=1}^\infty \left(\int_X \langle C(x) v, e_i\rangle \, d\mu(x)\right) e_i \ \forall \ v \in H. \tag{1.24}$$

5. If $D \in L^1(\nu : B(H))$ *, then*

$$\mu(C) \nu(D) = \mu \otimes \nu (C \otimes D) \tag{1.25}$$

where $\mu \otimes \nu$ *is the product measure on* $X \times Y$ *and* $C \otimes D \in L^1(\mu \otimes \nu : B(H))$ *is the operator defined by*

$$C \otimes D(x, y) := C(x) D(y) \ \ \forall \ x \in X \text{ and } y \in Y.$$

6. For $v, w \in H$*,*

$$\langle \mu(C) v, \nu(D) w\rangle = \int_{X \times Y} d\mu(x) \, d\nu(y) \, \langle C(x) v, D(y) w\rangle.$$

Proof. We leave the verifications of items 1., 2., and 4. to the reader.

Item 3. For $v, w \in H$ we have,

$$\begin{aligned}\langle \mu(C)^* v, w\rangle &= \overline{\langle \mu(C) w, v\rangle} = \overline{\int_X \langle C(x) w, v\rangle \, d\mu(x)} \\ &= \int_X \overline{\langle C(x) w, v\rangle} d\mu(x) = \int_X \langle v, C(x) w\rangle \, d\mu(x) \\ &= \int_X \langle C^*(x) v, w\rangle \, d\mu(x) = \langle \mu(C^*) v, w\rangle.\end{aligned}$$

Item 5. First observe that for $v, w \in H$,

$$\langle C \otimes D(x, y) v, w\rangle = \langle C(x) D(y) v, w\rangle = \sum_{i=1}^\infty \langle D(y) v, e_i\rangle \langle C(x) e_i, w\rangle \tag{1.26}$$

where $\{e_i\}_{i=1}^\infty$ is an orthonormal basis for H. From this relation it follows that $C \otimes D$ is still weakly measurable. Since

$$\begin{aligned}&\int_{X \times Y} \|C \otimes D(x, y)\|_{op} \, d\mu(x) \, d\nu(y) \\ &\quad = \int_{X \times Y} \|C(x) D(y)\|_{op} \, d\mu(x) \, d\nu(y) \\ &\quad \le \int_{X \times Y} \|C(x)\|_{op} \|D(y)\|_{op} \, d\mu(x) \, d\nu(y) = \|C\|_{L^1(\mu)} \|D\|_{L^1(\nu)} < \infty,\end{aligned}$$

we see $C \otimes D \in L^1(\mu \otimes \nu : B(H))$ and hence $\mu \otimes \nu (C \otimes D)$ is well defined. So it only remains to verify the identity in Eq. (1.25). However, making use of Eq. (1.26) and the estimates,

$$\begin{aligned}g(x, y) &:= \sum_{i=1}^\infty |\langle D(y) v, e_i\rangle| \, |\langle C(x) e_i, w\rangle| \\ &\le \sqrt{\sum_{i=1}^\infty |\langle D(y) v, e_i\rangle|^2 \sum_{i=1}^\infty |\langle C(x) e_i, w\rangle|^2} \\ &= \sqrt{\|D(y) v\|^2 \cdot \left\|C(x)^* w\right\|^2} \\ &\le \|D(y)\|_{op} \|C^*(x)\|_{op} \|v\| \, \|w\| \\ &= \|D(y)\|_{op} \|C(x)\|_{op} \|v\| \, \|w\|,\end{aligned}$$

it follows that $g \in L^1(\mu \otimes \nu)$. Using this observations we may easily justify the following computation,

$$\begin{aligned}\langle \mu \otimes \nu (C \otimes D) v, w\rangle &= \int_{X \times Y} d\mu(x) \, d\nu(y) \, \langle C(x) D(y) v, w\rangle \\ &= \int_{X \times Y} d\mu(x) \, d\nu(y) \sum_{i=1}^\infty \langle D(y) v, e_i\rangle \langle C(x) e_i, w\rangle \\ &= \sum_{i=1}^\infty \int_{X \times Y} d\mu(x) \, d\nu(y) \, \langle D(y) v, e_i\rangle \langle C(x) e_i, w\rangle \\ &= \sum_{i=1}^\infty \langle \nu(D) v, e_i\rangle \langle \mu(C) e_i, w\rangle = \langle \mu(C) \nu(D) v, w\rangle.\end{aligned}$$

Item 6. By the definition of $\mu(C)$ and $\nu(D)$,

$$\begin{aligned}\langle \mu(C) v, \nu(D) w\rangle &= \int_X d\mu(x) \, \langle C(x) v, \nu(D) w\rangle \\ &= \int_X d\mu(x) \int_Y d\nu(y) \, \langle C(x) v, D(y) w\rangle.\end{aligned}$$

■

Exercise 1.6. Let us continue to use the notation in Theorem 1.48. If $B \in B(H)$ is a linear operator such that $[C(x), B] = 0$ for μ − a.e. x, show $[\mu(C), B] = 0$.

Part II

Basics of Banach and C^*-Algebras

In this part, we will only begin to scratch the surface on the topic of Banach algebras. For an encyclopedic view of the subject, the reader is referred to Palmer [27, 28].For general Banach and C^*-algebra stuff have a look at [26, 53]. Also see the lecture notes in [33, 52]. Putnam's file looked quite good. For a very detailed statements see [9, See bottom of p. 45]

Banach Algebras and Linear ODE

2.1 Basic Definitions, Examples, and Properties

Definition 2.1. *An associative algebra over a field is a vector space over with a bilinear, associative multiplication: i.e.,*

$$\begin{aligned}(ab)c &= a(bc)\\ a(b+c) &= ab+ac\\ (a+b)c &= ac+bc\\ a(\lambda c) &= (\lambda a)c = \lambda(ac).\end{aligned}$$

As usual, from now on we assume that $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. Later in this chapter we will restrict to the complex case.

Definition 2.2. *A* ***Banach Algebra****, $\mathcal{A}$, is an $\mathbb{F}$ – Banach space which is an associative algebra over $\mathbb{F}$ satisfying,*

$$\|ab\| \le \|a\| \ \|b\| \ \forall\ a,b \in \mathcal{A}.$$

[It is typically the case that if $\mathcal{A}$ has a unit element, $\mathbf{1}$, then $\|\mathbf{1}\| = 1$. I will bake this into the definition!]

Exercise 2.1 (The unital correction). Let $\mathcal{A}$ be a Banach algebra with a unit, $\mathbf{1}$, with $\mathbf{1} \neq 0$. Suppose that we do not assume $\|\mathbf{1}\| = 1$. Show;

1. $\|\mathbf{1}\| \ge 1$.
2. For $a \in \mathcal{A}$, let $L_a \in B(\mathcal{A})$ be left multiplication by a, i.e. $L_a x = ax$ for all $x \in \mathcal{A}$. Now define
$$|a| = \|L_a\|_{B(\mathcal{A})} = \sup\{\|ax\| : x \in \mathcal{A} \text{ with } \|x\| = 1\}.$$
Show
$$\frac{1}{c}\|a\| \le |a| \le \|a\| \text{ for all } a \in \mathcal{A},$$
$|\mathbf{1}| = 1$ and $(\mathcal{A}, |\cdot|)$ is again a Banach algebra.

Examples 2.3 *Here are some examples of Banach algebras. The first example is the prototype for the definition.*

1. *Suppose that X is a Banach space, $\mathcal{B}(X)$ denote the collection of bounded operators on X. Then $\mathcal{B}(X)$ is a Banach algebra in operator norm with identity. $\mathcal{B}(X)$ is not commutative if $\dim X > 1$.*
2. *Let X be a topological space, $BC(X,\mathbb{F})$ be the bounded $\mathbb{F}$-valued, continuous functions on X, with $\|f\| = \sup_{x\in X}|f(x)|$. $BC(X,\mathbb{F})$ is a commutative Banach algebra under pointwise multiplication. The constant function $\mathbf{1}$ is an identity element.*
3. *If we assume that X is a locally compact Hausdorff space, then $C_0(X,\mathbb{F})$ – the space of continuous $\mathbb{F}$ – valued functions on X vanishing at infinity is a Banach sub-algebra of $BC(X,\mathbb{F})$. If X is non-compact, then $BC(X,\mathbb{F})$ is a Banach algebra without unit.*
4. *If $(\Omega,\mathcal{F},\mu)$ is a measure space then $L^\infty(\mu) := L^\infty(\Omega,\mathcal{F},\mu : \mathbb{C})$ is a commutative complex Banach algebra with identity. In this case $\|f\| = \|f\|_{L^\infty(\mu)}$ is the essential supremum of $|f|$ defined by*
$$\|f\|_{L^\infty(\mu)} = \inf\{M > 0 : |f| \le M \ \mu\text{-a.e.}\}.$$
5. *$\mathcal{A} = L^1(\mathbb{R}^1)$ with multiplication being convolution is a commutative Banach algebra* ***without*** *identity.*
6. *If $\mathcal{A} = \ell^1(\mathbb{Z})$ with multiplication given by convolution is a commutative Banach algebra* ***with*** *identity which is this case is the function*
$$\delta_0(n) := \begin{cases} 1 \text{ if } n = 0 \\ 0 \text{ if } n \neq 0 \end{cases}.$$
This example is generalized and expanded on in the next proposition.

Proposition 2.4 (Group Algebra). *Let G be a discrete group (i.e. finite or countable), $\mathcal{A} := \ell^1(G)$, and for $g \in G$ let $\delta_g \in \mathcal{A}$ be defined by*

$$\delta_g(x) := \begin{cases} 1 \text{ if } x = g \\ 0 \text{ if } x \neq g \end{cases}.$$

Then there exists a unique multiplication $(\cdot)$ on $\mathcal{A}$ which makes $\mathcal{A}$ into a Banach algebra with unit such that $\delta_g \circledast \delta_k = \delta_{gk}$ for all $g,k \in G$ which is given by

$$(u \circledast v)(x) = \sum_{g\in G} u(g)\, v\left(g^{-1}x\right) = \sum_{k\in G} u\left(xk^{-1}\right) v(k). \tag{2.1}$$

[The unit in $\mathcal{A}$ is δ_e where e is the identity element of G.]

Proof. If $u,v\in\ell^1(G)$ then

$$u=\sum_{g\in G}u(g)\,\delta_g \text{ and } v=\sum_{k\in G}v(k)\,\delta_k$$

where the above sums are convergent in $\mathcal{A}$. As we are requiring ($\circledast$) to be continuous we must have

$$u\circledast v=\sum_{g,k\in G}u(g)\,v(k)\,\delta_g\delta_k=\sum_{g,k\in G}u(g)\,v(k)\,\delta_{gk}.$$

Making the change of variables $x=gk$, i.e. $g=xk^{-1}$ or $k=g^{-1}x$ then shows,

$$u\circledast v=\sum_{g,x\in G}u(g)\,v\left(g^{-1}x\right)\delta_x=\sum_{k,x\in G}u\left(xk^{-1}\right)v(k)\,\delta_x.$$

This leads us to define $u\circledast v$ as in Eq. (2.1). Notice that

$$\sum_{x\in G}\sum_{k\in G}\left|u\left(xk^{-1}\right)\right|\left|v(k)\right|=\|u\|_1\|v\|_1$$

which shows that $u\circledast v$ is well defined and satisfies, $\|u\circledast v\|_1\le\|u\|_1\|v\|_1$. The reader may now verify that $(\mathcal{A},\circledast)$ is a Banach algebra. ∎

Remark 2.5. By construction, we have $\delta_g\circledast\delta_k=\delta_{gk}$ and so $(\mathcal{A},\circledast)$ is commutative iff G is commutative. Moreover for $k\in G$ and $u\in\ell^1(G)$ we have,

$$\delta_k\circledast u=\sum_{g\in G}u(g)\,\delta_{kg}=\sum_{g\in G}u\left(k^{-1}g\right)\delta_g=u\left(k^{-1}(\cdot)\right)$$

and

$$u\circledast\delta_k=\sum_{g\in G}u(g)\,\delta_{gk}=\sum_{g\in G}u\left(gk^{-1}\right)\delta_g=u\left((\cdot)\,k^{-1}\right).$$

In particular it follows that $\delta_e\circledast u=u=u\circledast\delta_e$ where $e\in G$ is the identity element.

Proposition 2.6. *Let $\mathcal{A}$ be a (complex) Banach algebra without identity. Let*

$$\mathcal{B}=\{(a,\alpha):a\in\mathcal{A},\alpha\in\mathbb{C}\}=\mathbb{A}\oplus\mathbb{C}.$$

Define

$$(a,\alpha)(b,\beta)=(ab+\alpha b+\beta a,\alpha\beta)$$

and

$$\|(a,\alpha)\|=\|a\|+|\alpha|. \tag{2.2}$$

Then $\mathcal{B}$ is a Banach algebra with identity $e=(0,1)$, and the map $a\to(a,0)$ is an isometric isomorphism onto a closed two sided ideal in $\mathcal{B}$.

Proof. Straightforward. ∎

Remark 2.7. If $\mathcal{A}$ is a C^*-algebra as in Definition 2.50 below it is better to defined the norm on $\mathcal{B}$ by

$$\|(a,\alpha)\|=\sup\{\|ab+\alpha b\|:b\in\mathcal{A}\text{ with }\|b\|\le1\} \tag{2.3}$$

rather than Eq. (2.2). The above definition is motivated by the fact that $a\in\mathcal{A}\hookrightarrow L_a\in B(\mathcal{A})$ is an isometry, where $L_ab=ab$ for all $a,b\in\mathcal{A}$. Indeed, $\|L_ab\|=\|ab\|\le\|a\|\,\|b\|$ with equality when $b=a^*$ so that $\|L_a\|_{B(\mathcal{A})}=\|a\|$. The definition in Eq. (2.3) has been crafted so that

$$\|(a,\alpha)\|=\|L_a+\alpha I\|_{B(\mathcal{A})}$$

which shows $\|(a,\alpha)\|$ is a norm and $a\in\mathcal{A}\hookrightarrow(a,0)\in\mathcal{B}\hookrightarrow B(\mathcal{A})$ are all isometric embeddings.

The advantage of this choice of norm is that $\mathcal{B}$ is still a C^*-algebra. Indeed

$$\begin{aligned}\|ab+\alpha b\|^2&=\|(ab+\alpha b)^*(ab+\alpha b)\|=\|(b^*a^*+\bar{\alpha}b^*)(ab+\alpha b)\|\\&=\|b^*a^*ab+\bar{\alpha}b^*ab+\alpha b^*a^*b+|\alpha|^2b^*b\|\\&\le\|b^*\|\,\|(a^*a+\bar{\alpha}a+\alpha a^*)\,b+|\alpha|^2b\|\end{aligned}$$

and so taking the sup of this expression over $\|b\|\le1$ implies

$$\|(a,\alpha)\|^2\le\left\|\left(a^*a+\bar{\alpha}a+\alpha a^*,|\alpha|^2\right)\right\|=\|(a,\alpha)^*(a,\alpha)\|\le\|(a,\alpha)^*\|\,\|(a,\alpha)\|. \tag{2.4}$$

Eq. (2.4) implies $\|(a,\alpha)\|\le\|(a,\alpha)^*\|$ and by symmetry $\|(a,\alpha)^*\|\le\|(a,\alpha)\|$. Thus the inequalities in Eq. (2.4) are equalities and this shows $\|(a,\alpha)\|^2=\|(a,\alpha)^*(a,\alpha)\|$. Moreover $\mathcal{A}$ is still embedded in $\mathcal{B}$ isometrically. because for $a\in\mathcal{A}$,

$$\|a\|=\left\|a\,\frac{a^*}{\|a\|}\right\|\le\sup\{\|ab\|:b\in\mathcal{A}\text{ with }\|b\|\le1\}\le\|a\|$$

which combined with Eq. (2.3) implies $\|(a,0)\|=\|a\|$.

Definition 2.8. *Let $\mathcal{A}$ be a Banach algebra with identity, 1. If $a\in\mathcal{A}$, then a is* ***right (left)*** *invertible if there exists $b\in\mathcal{A}$ such that $ab=1$ ($ba=1$) in which case we call b a* ***right (left) inverse*** *of a. The element a is called* ***invertible*** *if it has both a left and a right inverse.*

Note if $ab=1$ and $ca=1$, then $c=cab=b$. Therefore if a has left and right inverses then they are equal and such inverses are unique. When a is invertible, we will write a^{-1} for the unique left and right inverse of a. The next lemma shows that notion of inverse given here is consistent with the notion of algebraic inverses when $\mathcal{A}=B(X)$ for some Banach space X.

Lemma 2.9 (Inverse Mapping Theorem). *If X, Y are Banach spaces and $T \in L(X, Y)$ is invertible (i.e. a bijective linear transformation) then the inverse map, T^{-1}, is **bounded**, i.e. $T^{-1} \in B(Y, X)$. (Note that T^{-1} is automatically linear.) In other words algebraic invertibility implies topological invertibility.*

Proof. If T is surjective, we know by the open mapping theorem that T is an open mapping and form this it follows that the algebraic inverse of T is continuous. ∎

Corollary 2.10 (Closed ranges). *Let X and Y be Banach spaces and $T \in L(X, Y)$. Then $\mathrm{Nul}(T) = \{0\}$ and $\mathrm{Ran}(T)$ is closed in Y iff*

$$\varepsilon := \inf_{\|x\|_X = 1} \|Tx\|_Y > 0. \tag{2.5}$$

Proof. If $\mathrm{Nul}(T) = \{0\}$ and $\mathrm{Ran}(T)$ is closed then T thought of an operator in $B(X, \mathrm{Ran}(T))$ is an invertible map with inverse denoted by $S : \mathrm{Ran}(T) \to X$. Since $\mathrm{Ran}(T)$ is a closed subspace of a Banach space it is itself a Banach space and so by Corollary 2.9 we know that S is a bounded operator, i.e.

$$\|Sy\|_X \le \|S\|_{op} \cdot \|y\|_Y \ \ \forall\, y \in \mathrm{Ran}(T).$$

Taking $y = Tx$ in the above inequality shows,

$$\|x\|_X \le \|S\|_{op} \cdot \|Tx\|_Y \ \ \forall\, x \in X$$

from which we learn $\varepsilon = \|S\|_{op}^{-1} > 0$.

Conversely if $\varepsilon > 0$ (ε as in Eq. (2.5)), then by scaling, it follows that

$$\|Tx\|_Y \ge \varepsilon \|x\|_X \quad \forall \ \ x \in X.$$

This last inequality clearly implies $\mathrm{Nul}(T) = \{0\}$. Moreover if $\{x_n\} \subset X$ is a sequence such that $y := \lim_{n\to\infty} Tx_n$ exists in Y, then

$$\begin{aligned}\|x_n - x_m\| &\le \frac{1}{\varepsilon}\|T(x_n - x_m)\|_Y = \frac{1}{\varepsilon}\|Tx_n - Tx_m\|_Y \\ &\to \frac{1}{\varepsilon}\|y - y\|_Y = 0 \text{ as } m, n \to \infty.\end{aligned}$$

Therefore $x := \lim_{n\to\infty} x_n$ exists in X and $y = \lim_{n\to\infty} Tx_n = Tx$ which shows $\mathrm{Ran}(T)$ is closed. ∎

Example 2.11. Let $X = \ell^1(\mathbb{N}_0)$ and $T : X \to C([0,1])$ be defined by $Ta = \sum_{n=0}^{\infty} a_n x^n$. Now let $Y := \mathrm{Ran}(T)$ so that $T : X \to Y$ is bijective. The inverse map is again not bounded. For example consider $a = (1, -1, 1, -1, \dots, \pm 1, 0, 0, 0, \dots)$ so that

$$Ta = \sum_{k=0}^{n} (-x)^k = \frac{(-x)^{n+1} - 1}{-x - 1} = \frac{1 + (-1)^n x^{n+1}}{1 + x}.$$

We then have $\|Ta\|_\infty \le 2$ while $\|a\|_X = n + 1$. Thus $\left\|T^{-1}\right\|_{op} = \infty$. This shows that range space in the open mapping theorem must be complete as well.

The next elementary proposition shows how to use geometric series in order to construct inverses.

Proposition 2.12. *Let $\mathcal{A}$ be a Banach algebra with identity and $a \in \mathcal{A}$. If $\sum_{n=0}^{\infty} \|a^n\| < \infty$ then $1 - a$ is invertible and*

$$\left\|(1-a)^{-1}\right\| \le \sum_{n=0}^{\infty} \|a^n\|.$$

In particular, if $\|a\| < 1$, then $1 - a$ is invertible and

$$\left\|(1-a)^{-1}\right\| \le \frac{1}{1 - \|a\|}.$$

Proof. Let $b = \sum_{n=0}^{\infty} a^n$ which, by assumption, is absolutely convergent and so satisfies, $\|b\| \le \sum_{n=0}^{\infty} \|a^n\|$. It is easy to verify that $(1-a)b = b(1-a) = 1$ which implies $(1-a)^{-1} = b$ which proves the first assertion. Then second assertion now follows from the first and the simple estimates, $\|a^n\| \le \|a\|^n$, and geometric series identity, $\sum_{n=0}^{\infty} \|a\|^n = 1/(1 - \|a\|)$. ∎

Notation 2.13 *Let $\mathcal{A}_{inv}$ denote the **invertible elements** for $\mathcal{A}$ and by **convention** we write λ instead of $\lambda 1$.*

Remark 2.14. The invertible elements, $\mathcal{A}_{inv}$, form a multiplicative system, i.e. if $a, b \in \mathcal{A}_{inv}$, then $ab \in \mathcal{A}_{inv}$. As usual we have $(ab)^{-1} = b^{-1}a^{-1}$ as is easily verified.

Corollary 2.15. *If $x \in \mathcal{A}_{inv}$ and $h \in \mathcal{A}$ satisfy $\left\|x^{-1}h\right\| < 1$, show $x + h \in \mathcal{A}_{inv}$ and*

$$\left\|(x+h)^{-1}\right\| \le \left\|x^{-1}\right\| \cdot \frac{1}{1 - \|x^{-1}h\|}. \tag{2.6}$$

In particular this shows $\mathcal{A}_{inv}$ of invertible is an open subset of $\mathcal{A}$. We further have

$$\begin{aligned}(x+h)^{-1} &= \sum_{n=0}^{\infty} (-1)^n \left(x^{-1}h\right)^n x^{-1} \\ &= x^{-1} - x^{-1}hx^{-1} + x^{-1}hx^{-1}hx^{-1} - x^{-1}hx^{-1}hx^{-1}hx^{-1} + \dots. \\ &= \sum_{n=0}^{N} (-1)^n \left(x^{-1}h\right)^n x^{-1} + R_N\end{aligned}$$

where
$$\|R_N\| \le \left\|\left(x^{-1}h\right)^{N+1}\right\| \left\|x^{-1}\right\| \frac{1}{1-\|x^{-1}h\|}.$$

Proof. By the assumptions and Proposition 2.12, both x and $1+x^{-1}h$ are invertible with
$$\left\|1+x^{-1}h\right\| \le \frac{1}{1-\|x^{-1}h\|}.$$
As $(x+h) = x\left(1+x^{-1}h\right)$, it follows that $x+h$ is invertible and
$$(x+h)^{-1} = \left(1+x^{-1}h\right)^{-1} x^{-1}.$$
Taking norms of this equation then gives the estimate in Eq. (2.6). The series expansion now follows from the previous equation and the geometric series representation in Proposition 2.12. Lastly the remainder estimate is easily obtained as follows;
$$\begin{aligned} R_N &= \sum_{n>N} \left(-x^{-1}h\right)^n x^{-1} = \left(-x^{-1}h\right)^{N+1} \left[\sum_{n=0}^{\infty} \left(-x^{-1}h\right)^n\right] x^{-1} \\ &= \left(-x^{-1}h\right)^{N+1} \left(1+x^{-1}h\right)^{-1} x^{-1} \end{aligned}$$
so that
$$\begin{aligned} \|R_N\| &\le \left\|x^{-1}\right\| \left\|\left(1+x^{-1}h\right)^{-1}\right\| \left\|\left(x^{-1}h\right)^{N+1}\right\| \\ &\le \left\|\left(x^{-1}h\right)^{N+1}\right\| \left\|x^{-1}\right\| \frac{1}{1-\|x^{-1}h\|}. \end{aligned}$$
■

In the sequel the following simple identity is often useful; if $b, c \in \mathcal{A}_{inv}$, then
$$b^{-1} - c^{-1} = b^{-1}(c-b)c^{-1}. \tag{2.7}$$
This identity is the non-commutative form of adding fractions by using a common denominator. Here is a simple (redundant in light of Corollary 2.15) application.

Corollary 2.16. *The map,* $\mathcal{A}_{inv} \ni x \to x^{-1} \in \mathcal{A}_{inv}$ *is continuous. [This map is in fact* C^∞*, see Exercise 2.2 below.]*

Proof. Suppose that $x \in \mathcal{A}_{inv}$ and $h \in \mathcal{A}$ is sufficiently small so that $\left\|x^{-1}h\right\| \le \left\|x^{-1}\right\| \|h\| < 1$. Then $x+h$ is invertible by Corollary 2.15 and we find the identity,
$$(x+h)^{-1} - x^{-1} = (x+h)^{-1}(x-(x+h))x^{-1} = -(x+h)^{-1}hx^{-1}. \tag{2.8}$$
From Eq. (2.8) and Corollary 2.15 it follows that
$$\left\|(x+h)^{-1} - x^{-1}\right\| \le \left\|x^{-1}\right\| \left\|(x+h)^{-1}\right\| \|h\| \le \left\|x^{-1}\right\|^2 \cdot \frac{\|h\|}{1-\|x^{-1}h\|} \to 0 \text{ as } h \to 0.$$
■

2.2 Calculus in Banach Algebras

Exercise 2.2. Show that the inversion map $f : \mathcal{A}_{inv} \to \mathcal{A}_{inv} \subset \mathcal{A}$ defined by $f(x) = x^{-1}$ is differentiable with
$$f'(x)h = (\partial_h f)(x) = -x^{-1}hx^{-1}$$
for all $x \in \mathcal{A}_{inv}$ and $h \in \mathcal{A}$. **Hint:** iterate the identity
$$(x+h)^{-1} = x^{-1} - (x+h)^{-1} hx^{-1} \tag{2.9}$$
that was derived in the lecture notes. [Again this exercise is somewhat redundant in light of light of Corollary 2.15.]

Exercise 2.3. Suppose that $a \in \mathcal{A}$ and $t \in \mathbb{R}$ (or $\mathbb{C}$ if $\mathcal{A}$ is a complex Banach algebra). Show directly that:

1. $e^{ta} := \sum_{n=0}^{\infty} \frac{t^n}{n!} a^n$ is an absolutely convergent series and $\|e^{ta}\| \le e^{|t|\|a\|}$.
2. e^{ta} is differentiable in t and that $\frac{d}{dt}e^{ta} = ae^{ta} = e^{ta}a$. [Suggestion; you could prove this by scratch or make use of Exercise 1.2.]

Corollary 2.17. *For* $a, b \in \mathcal{A}$ *commute, i.e.* $ab = ba$*, then* $e^a e^b = e^{a+b} = e^b e^a$.

Proof. In the proof to follows we will use $e^{ta}b = be^{ta}$ for all $t \in \mathbb{R}$. [Proof is left to the reader.] Let $f(t) := e^{-ta}e^{t(a+b)}$, then by the product rule,
$$\dot{f}(t) = -e^{-ta}ae^{t(a+b)} + e^{-ta}(a+b)e^{t(a+b)} = e^{-ta}be^{t(a+b)} = be^{-ta}e^{t(a+b)} = bf(t).$$
Therefore, $\frac{d}{dt}\left[e^{-tb}f(t)\right] = 0$ and hence $e^{-tb}f(t) = e^{-0b}f(0) = 1$. Altogether we have shown,
$$e^{-tb}e^{-ta}e^{t(a+b)} = e^{-tb}f(t) = 1.$$
Taking $t = \pm 1$ and $b = 0$ in this identity shows $e^{-a}e^a = 1 = e^a e^{-a}$, i.e. $(e^a)^{-1} = e^{-a}$. Knowing this fact it then follows from the previously displayed equation that $e^{t(a+b)} = e^{ta}e^{tb}$ which at $t = 1$ gives, $e^a e^b = e^{a+b}$. Interchanging the roles of a and b then completes the proof. ■

Corollary 2.18. *Suppose that $A \in \mathcal{A}$, then the solution to*

$$\dot{y}(t) = Ay(t) \text{ with } y(0) = 1$$

is given by $y(t) = e^{tA}$ where

$$e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n. \tag{2.10}$$

Moreover,

$$e^{(t+s)A} = e^{tA} e^{sA} \text{ for all } s, t \in \mathbb{R}. \tag{2.11}$$

We also have the following converse to this corollary whose proof is outlined in Exercise 2.16 below.

Theorem 2.19. *Suppose that $T_t \in \mathcal{A}$ for $t \geq 0$ satisfies*

1. *(Semi-group property.) $T_0 = 1 \in \mathcal{A}$ and $T_t T_s = T_{t+s}$ for all $s, t \geq 0$.*
2. *(Norm Continuity) $t \to T_t$ is continuous at 0, i.e. $\|T_t - I\|_{\mathcal{A}} \to 0$ as $t \downarrow 0$.*

Then there exists $A \in \mathcal{A}$ such that $T_t = e^{tA}$ where e^{tA} is defined in Eq. (2.10).

Exercise 2.4. Let $a, b \in \mathcal{A}$ and $f(t) := e^{t(a+b)} - e^{ta} e^{tb}$ and then show

$$\ddot{f}(0) = ab - ba.$$

[Therefore if $e^{t(a+b)} = e^{ta} e^{tb}$ for t near 0, then $ab = ba$.]

Exercise 2.5. If $\mathcal{A}_0$ is a unital commutative Banach algebra, show $\exp(a) = e^a$ is a differentiable function with differential,

$$\exp'(a) b = e^a b = b e^a.$$

Exercise 2.6. If $t \to c(t) \in \mathcal{A}$ is a C^1-function such that $[c(s), c(t)] = 0$ for all $s, t \in \mathbb{R}$, then show

$$\frac{d}{dt} e^{c(t)} = \dot{c}(t) e^{c(t)}.$$

Notation 2.20 *For $a \in \mathcal{A}$, let $\operatorname{ad}_a \in B(\mathcal{A})$ be defined by $\operatorname{ad}_a b = ab - ba$.*

Notice that

$$\|\operatorname{ad}_a b\| \leq 2 \|a\| \|b\| \ \forall b \in \mathcal{A}$$

and hence $\|\operatorname{ad}_a\|_{op} \leq 2\|a\|$.

Proposition 2.21. *If $a, b \in \mathcal{A}$, then*

$$e^a b e^{-a} = e^{\operatorname{ad}_a}(b) = \sum_{n=0}^{\infty} \frac{1}{n!} \operatorname{ad}_a^n b.$$

where e^{ad_a} is computed by working in the Banach algebra, $B(\mathcal{A})$.

Proof. Let $f(t) := e^{ta} b e^{-ta}$, then

$$\dot{f}(t) = a e^{ta} b e^{-ta} - e^{ta} b e^{-ta} a = \operatorname{ad}_a f(t) \text{ with } f(0) = b.$$

Thus it follows that

$$\frac{d}{dt}\left[e^{-t \operatorname{ad}_a} f(t)\right] = 0 \implies e^{-t \operatorname{ad}_a} f(t) = e^{-0 \operatorname{ad}_a} f(0) = b.$$

From this we conclude,

$$e^{ta} b e^{-ta} = f(t) = e^{t \operatorname{ad}_a}(b).$$

■

Corollary 2.22. *Let $a, b \in \mathcal{A}$ and suppose that $[a, b] := ab - ba$ commutes with both a and b. Then*

$$e^a e^b = e^{a+b+\frac{1}{2}[a,b]}.$$

Proof. Let $u(t) := e^{ta} e^{tb}$ and then compute,

$$\begin{aligned} \dot{u}(t) &= a e^{ta} e^{tb} + e^{ta} b e^{tb} = a e^{ta} e^{tb} + e^{ta} b e^{-ta} e^{ta} e^{tb} \\ &= \left[a + e^{t \operatorname{ad}_a}(b)\right] u(t) = c(t) u(t) \text{ with } u(0) = 1, \end{aligned} \tag{2.12}$$

where

$$c(t) = a + e^{t \operatorname{ad}_a}(b) = a + b + t[a, b]$$

because

$$\operatorname{ad}_a^2 b = [a, [a, b]] = 0 \text{ by assumption.}$$

Furthermore, our assumptions imply for all $s, t \in \mathbb{R}$ that

$$\begin{aligned} [c(t), c(s)] &= [a + b + t[a, b], a + b + s[a, b]] \\ &= [t[a, b], a + b + s[a, b]] = st[[a, b], [a, b]] = 0. \end{aligned}$$

Therefore the solution to Eq. (2.12) is given by

$$u(t) = e^{\int_0^t c(\tau) d\tau} = e^{t(a+b) + \frac{1}{2} t^2 [a, b]}.$$

Taking $t = 1$ complete the proof. ■

Remark 2.23 (Baker-Campbell-Dynkin-Hausdorff formula). In general the Baker-Campbell-Dynkin-Hausdorff formula states there is a function $\Gamma(a,b) \in \mathcal{A}$ defined for $\|a\|_{\mathcal{A}} + \|b\|_{\mathcal{A}}$ sufficiently small such that

$$e^a e^b = e^{\Gamma(a,b)} \text{ where}$$
$$\Gamma(a,b) = a + b + \frac{1}{2}[a,b] + \frac{1}{12}\left(ad_a^2 b + ad_b^2 a\right) + \dots$$

where all of the higher order terms are linear combinations of terms of the form $ad_{x_1} \dots ad_{x_n} x_0$ with $x_i \in \{a,b\}$ for $0 \le i \le n$ and $n \ge 3$.

Exercise 2.7. Suppose that $a(s,t) \in \mathcal{A}$ is a C^2-function (s,t) near $(s_0,t_0) \in \mathbb{R}^2$, show $(s,t) \to e^{a(s,t)} \in \mathcal{A}$ is still C^2. Hints:

1. Let $f_n(s,t) := \frac{a(s,t)^n}{n!}$ and then verify
$$\left\|\dot{f}_n\right\| \le \frac{1}{(n-1)!}\|a\|^{n-1}\|\dot{a}\|,$$
$$\|f_n'\| \le \frac{1}{(n-1)!}\|a\|^{n-1}\|a'\|,$$
$$\left\|\ddot{f}_n\right\| \le \frac{1}{(n-2)!}\|a\|^{n-2}\|\dot{a}\|^2 + \frac{1}{(n-1)!}\|a\|^{n-1}\|\ddot{a}\|$$
$$\left\|\dot{f}_n'\right\| \le \frac{1}{(n-2)!}\|a\|^{n-2}\|\dot{a}\|\|a'\| + \frac{1}{(n-1)!}\|a\|^{n-1}\|\dot{a}'\|$$
$$\|f_n''\| \le \frac{1}{(n-2)!}\|a\|^{n-2}\|a'\|^2 + \frac{1}{(n-1)!}\|a\|^{n-1}\|a''\|$$
where $\dot{f} := \partial f/\partial t$ and $f' = \frac{\partial f}{\partial s}$.
2. Use the above estimates along with repeated applications of Exercise 1.2 in order to conclude that $f(s,t) = e^{a(s,t)}$ is C^2 near (s_0,t_0).

Theorem 2.24 (Differential of e^a). *For any $a, b \in \mathcal{A}$,*

$$\partial_b e^a := \frac{d}{ds}|_0 e^{a+sb} = e^a \int_0^1 e^{-ta} b e^{ta} dt.$$

Proof. The function, $u(s,t) := e^{t(a+sb)}$ is C^2 by Exercise 2.7 and therefore we find,

$$\frac{d}{dt}u_s(0,t) = \frac{\partial}{\partial s}|_0 \dot{u}(s,t) = \frac{\partial}{\partial s}|_0 [(a+sb)u(s,t)]$$
$$= bu(s,t) + au_s(0,t) \text{ with } u_s(0,0) = 0.$$

To solve this equation we consider,

$$\frac{d}{dt}\left[e^{-ta}u_s(0,t)\right] = e^{-ta}bu(0,t) = e^{-ta}be^{ta}$$

which upon integration,

$$e^{-a}\left[\partial_b e^a\right] = e^{-a}u_s(0,1) = \int_0^1 e^{-ta}be^{ta}dt$$

and hence

$$\partial_b e^a = e^a \int_0^1 e^{-ta}be^{ta}dt.$$

■

Corollary 2.25. *The map $a \to e^a$ is differentiable. More precisely,*

$$\left\|e^{a+b} - e^a - \partial_b e^a\right\| = O\left(\|b\|^2\right).$$

Proof. From Theorem 2.24,

$$\frac{d}{ds}e^{a+sb} = \frac{d}{d\varepsilon}|_0 e^{a+sb+\varepsilon b} = e^{a+sb}\int_0^1 e^{-t(a+sb)}be^{t(a+sb)}dt$$

and therefore,

$$e^{a+b} - e^a - \partial_b e^a = \int_0^1 ds e^{a+sb}\int_0^1 dt e^{-t(a+sb)}be^{t(a+sb)} - e^a\int_0^1 e^{-ta}be^{ta}dt$$
$$= \int_0^1 ds \int_0^1 dt \left[e^{(1-t)(a+sb)}be^{t(a+sb)} - e^{(1-t)a}be^{ta}\right]$$

and so

$$\left\|e^{a+b} - e^a - \partial_b e^a\right\| \le \int_0^1 ds \int_0^1 dt \left\|e^{(1-t)(a+sb)}be^{t(a+sb)} - e^{(1-t)a}be^{ta}\right\|.$$

To estimate right side, let

$$g(s,t) := e^{(1-t)(a+sb)}be^{t(a+sb)} - e^{(1-t)a}be^{ta}.$$

Then by Theorem 2.24,

$$\|g'(s,t)\| = \left\|\frac{d}{ds}\left[e^{(1-t)(a+sb)}be^{t(a+sb)}\right]\right\| \le C\|b\|^2$$

and since $g(0,t) = 0$, we conclude that $\|g(s,t)\| \le C\|b\|^2$. Hence it follows that

$$\left\|e^{a+b} - e^a - \partial_b e^a\right\| = O\left(\|b\|^2\right).$$

■

2.3 General Linear ODE in $\mathcal{A}$

There is a bit of change of notation in this section as we use both capital and lower case letters for possible elements of $\mathcal{A}$. Let us now work with more general linear differential equations on $\mathcal{A}$ where again $\mathcal{A}$ is a Banach algebra with identity. Further let $J=(a,b)\subset\mathbb{R}$ be an open interval. Further suppose that $h,A\in C(J,\mathcal{A})$, $s\in J$, and $x\in\mathcal{A}$ are give then we wish to solve the ordinary differential equation,

$$\dot{y}(t)=A(t)\,y(t)+h(t)\ \text{ with } y(s)=x\in\mathcal{A}, \tag{2.13}$$

for a function, $y\in C^1(J,\mathcal{A})$. This equation may be written in its equivalent (as the reader should verify) integral form, namely we are looking for $y\in C(J,\mathcal{A})$ such that

$$y(t)=\int_s^t A(\tau)\,y(\tau)\,d\tau+x+\int_s^t h(\tau)\,d\tau. \tag{2.14}$$

Notation 2.26 *For $\varphi\in C(J,\mathcal{A})$, let $\|\varphi\|_\infty:=\max_{t\in J}\|\varphi(t)\|\in[0,\infty]$. We further let*

$$BC(J,\mathcal{A}):=\{\varphi\in C(J,\mathcal{A}):\|\varphi\|_\infty<\infty\}$$

denote the bounded functions in $C(J,\mathcal{A})$.

The reader should verify that $BC(J,\mathcal{A})$ with $\|\cdot\|_\infty$ is again a Banach algebra. If we let

$$(\Lambda_s y)(t)=\left(\Lambda_s^A y\right)(t):=\int_s^t A(\tau)\,y(\tau)\,d\tau \text{ and} \tag{2.15}$$

$$\varphi(t):=x+\int_s^t h(\tau)\,d\tau$$

then these equations may be written as

$$y=\Lambda_s y+\varphi \iff (\mathcal{I}-\Lambda_s)\,y=\varphi.$$

Thus we see these equations will have a unique solution provided $(\mathcal{I}-\Lambda_s)^{-1}$ is invertible. To simplify the exposition without real loss of generality we are going to now assume

$$\|A\|_1:=\int_J\|A(\tau)\|\,d\tau<\infty. \tag{2.16}$$

The point of this assumption if Λ_s is defined as in Eq. (2.15), then for $y\in BC(J,\mathcal{A})$ and $t\in J$,

$$\|(\Lambda y)(t)\|\le\left|\int_0^t\|A(\tau)\,y(\tau)\|\,d\tau\right|\le\left|\int_0^t\|A(\tau)\|\,d\tau\right|\|y\|_\infty\le\int_J\|A(\tau)\|\,d\tau\cdot\|y\|_\infty. \tag{2.17}$$

This inequality then immediately implies $\Lambda_s:BC(J,\mathcal{A})\to BC(J,\mathcal{A})$ is a bounded operator with $\|\Lambda_s\|_{op}\le\|A\|_1$. In fact we will see below in Corollary 2.29 that more generally we have

$$\|\Lambda_s^n\|_{op}\le\frac{1}{n!}\left(\|A\|_1\right)^n$$

which is the key to showing $(\mathcal{I}-\Lambda_s)^{-1}$ is invertible.

Lemma 2.27. *For all $n\in\mathbb{N}$,*

$$(\Lambda_s^n\varphi)(t)=\int_s^t d\tau_n\int_s^{\tau_n}d\tau_{n-1}\cdots\int_s^{\tau_2}d\tau_1 A(\tau_n)\dots A(\tau_1)\,\varphi(\tau_1).$$

Proof. The proof is by induction with the induction step being,

$$\begin{aligned}(\Lambda_s^{n+1}\varphi)(t)&=(\Lambda_s^n\Lambda_s\varphi)(t)\\&=\int_s^t d\tau_n\int_s^{\tau_n}d\tau_{n-1}\cdots\int_s^{\tau_2}d\tau_1 A(\tau_n)\dots A(\tau_1)\,(\Lambda_s\varphi)(\tau_1)\\&=\int_s^t d\tau_n\int_s^{\tau_n}d\tau_{n-1}\cdots\int_s^{\tau_2}d\tau_1 A(\tau_n)\dots A(\tau_1)\int_s^{\tau_1}A(\tau_0)\,\varphi(\tau_0)\,d\tau_0\\&=\int_s^t d\tau_n\int_s^{\tau_n}d\tau_{n-1}\cdots\int_s^{\tau_1}d\tau_0 A(\tau_n)\dots A(\tau_1)\,A(\tau_0)\,\varphi(\tau_0).\end{aligned}$$

■

Lemma 2.28. *Suppose that $\psi\in C(J,\mathbb{R})$, then*

$$\int_s^t d\tau_n\int_s^{\tau_n}d\tau_{n-1}\cdots\int_s^{\tau_2}d\tau_1\psi(\tau_n)\dots\psi(\tau_1)=\frac{1}{n!}\left(\int_s^t\psi(\tau)\,d\tau\right)^n. \tag{2.18}$$

Proof. The proof will go by induction on n with $n=1$ assertion obviously being true. Now let $\Psi(t):=\int_s^t\psi(\tau)\,d\tau$ so that the right side of Eq. (2.18) is $\Psi(t)^n/n!$ and $\dot{\Psi}(t)=\psi(t)$. We now complete the induction step;

$$\begin{aligned}&\int_s^t d\tau_n\int_s^{\tau_n}d\tau_{n-1}\cdots\int_s^{\tau_1}d\tau_0\psi(\tau_n)\dots\psi(\tau_0)\\&\quad=\frac{1}{n!}\int_s^t d\tau_n\psi(\tau_n)\,[\Psi(\tau_n)]^n=\frac{1}{n!}\int_s^t d\tau\,[\Psi(\tau)]^n\,\dot{\Psi}(\tau)\\&\quad=\frac{1}{(n+1)!}\,[\Psi(\tau)]^{n+1}\,|_{\tau=s}^{\tau=t}=\frac{1}{(n+1)!}\,[\Psi(t)]^{n+1}.\end{aligned}$$

■

Corollary 2.29. *For all* $n \in \mathbb{N}$,

$$\|\Lambda_s^n\|_{op} \le \frac{1}{n!}\|A\|_1^n = \frac{1}{n!}\left[\int_J \|A(\tau)\| d\tau\right]^n$$

and therefore $(\mathcal{I} - \Lambda_s)$ *is invertible with*

$$\left\|(\mathcal{I} - \Lambda_s)^{-1}\right\|_{op} \le \exp(\|A\|_1) = \exp\left(\int_J \|A(\tau)\| d\tau\right).$$

Proof. This follows by the simple estimate along with Lemma 2.27 that for any $t \in J$,

$$\begin{aligned}\|(\Lambda_s^n \varphi)(t)\| &\le \left|\int_s^t d\tau_n \int_s^{\tau_n} d\tau_{n-1} \cdots \int_s^{\tau_2} d\tau_1 \|A(\tau_n)\dots A(\tau_1)\varphi(\tau_1)\|\right| \\ &\le \left|\int_s^t d\tau_n \int_s^{\tau_n} d\tau_{n-1} \cdots \int_s^{\tau_2} d\tau_1 \|A(\tau_n)\|\dots\|A(\tau_1)\|\right| \|\varphi\|_\infty \\ &= \frac{1}{n!}\left|\int_s^t \|A(\tau)\| d\tau\right|^n \|\varphi\|_\infty \le \frac{1}{n!}\left(\int_J \|A(\tau)\| d\tau\right)^n \|\varphi\|_\infty.\end{aligned}$$

Taking the supremum over $t \in J$ then shows

$$\|\Lambda_s^n \varphi\|_\infty \le \frac{1}{n!}\left(\int_J \|A(\tau)\| d\tau\right)^n \|\varphi\|_\infty$$

which completes the proof. ■

Theorem 2.30. *For all* $\varphi \in BC(J, \mathcal{A})$, *there exists a unique solution,* $y \in BC(J, \mathcal{A})$, *to* $y = \Lambda_s y + \varphi$ *which is given by*

$$\begin{aligned}y(t) &= \left((\mathcal{I} - \Lambda_s)^{-1}\varphi\right)(t) \\ &= \varphi(t) + \sum_{n=1}^\infty \int_s^t d\tau_n \int_s^{\tau_n} d\tau_{n-1} \cdots \int_s^{\tau_2} d\tau_1 A(\tau_n)\dots A(\tau_1)\varphi(\tau_1).\end{aligned}$$

Notation 2.31 *For* $s, t \in J$, *let* $u_0^A(t,s) = \mathbf{1}$ *and for* $n \in \mathbb{N}$ *let*

$$u_n^A(t,s) := \int_s^t d\tau_n \int_s^{\tau_n} d\tau_{n-1} \cdots \int_s^{\tau_2} d\tau_1 A(\tau_n)\dots A(\tau_1). \tag{2.19}$$

Definition 2.32 (Fundamental Solutions). *For* $s, t \in J$, *let*

$$\begin{aligned}u^A(t,s) &:= \left((\mathcal{I} - \Lambda_s)^{-1}\mathbf{1}\right)(t) = \sum_{n=0}^\infty u_n^A(t,s) && (2.20) \\ &= \mathbf{1} + \sum_{n=1}^\infty \int_s^t d\tau_n \int_s^{\tau_n} d\tau_{n-1} \cdots \int_s^{\tau_2} d\tau_1 A(\tau_n)\dots A(\tau_1). && (2.21)\end{aligned}$$

Equivalently $u^A(t,s)$ *is the unique solution to the ODE,*

$$\frac{d}{dt}u^A(t,s) = A(t)\,u^A(t,s) \ \text{ with } u^A(s,s) = \mathbf{1}.$$

Proposition 2.33 (Group Property). *For all* $s, \sigma, t \in J$ *we have*

$$u^A(t,s)\,u^A(s,\sigma) = u^A(t,\sigma). \tag{2.22}$$

Proof. Both sides of Eq. (2.22) satisfy the same ODE, namely the ODE

$$\dot{y}(t) = A(t)\,y(t) \ \text{ with } y(s) = u^A(s,\sigma).$$

The uniqueness of such solutions completes the proof. ■

Lemma 2.34 (A Fubini Result). *Let* $s, t \in J$, $n \in \mathbb{N}$ *and* $f(\tau_n, \dots, \tau_1, \tau_0)$ *be a continuous function with values in* $\mathcal{A}$, *then*

$$\begin{aligned}&\int_s^t d\tau_n \int_s^{\tau_n} d\tau_{n-1} \cdots \int_s^{\tau_2} d\tau_1 \int_s^{\tau_1} d\tau_0 f(\tau_n, \dots, \tau_1, \tau_0) \\ &\quad = \int_s^t d\tau_0 \int_{\tau_0}^t d\tau_n \int_{\tau_0}^{\tau_n} d\tau_{n-1} \cdots \int_{\tau_0}^{\tau_2} d\tau_1 f(\tau_n, \dots, \tau_1, \tau_0).\end{aligned}$$

Proof. We simply use Fubini's theorem to change the order of integration while referring to Figure (2.1) in order to work out the correct limits of integration. ■

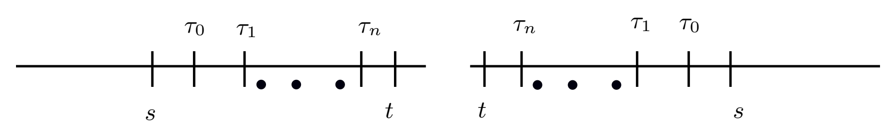

Fig. 2.1. This figures shows how to find the new limits of integration when $t > s$ and $t < s$ respectively.

Lemma 2.35. *If* $n \in \mathbb{N}_0$ *and* $s, t \in J$, *then in general,*

$$\left(\Lambda_s^{n+1}\varphi\right)(t) = \int_s^t u_n^A(t,\sigma)\,A(\sigma)\,\varphi(\sigma)\,d\sigma. \tag{2.23}$$

and if $H(t) := \int_s^\cdot h(\tau)\,d\tau$, *then*

$$(\Lambda_s^n H)(t) = \int_s^t u_n^A(t,\sigma)\,h(\sigma)\,d\sigma. \tag{2.24}$$

Proof. Using Lemma 2.34 shows,

$$\begin{aligned}\left(\Lambda_s^{n+1}\varphi\right)(t) &= \int_s^t d\tau_n \cdots \int_s^{\tau_2} d\tau_1 \int_s^{\tau_1} d\tau_0 A(\tau_n)\dots A(\tau_1) A(\tau_0)\varphi(\tau_0) \\ &= \int_s^t d\tau_0 \left[\int_{\tau_0}^t d\tau_n \int_{\tau_0}^{\tau_n} d\tau_{n-1} \cdots \int_{\tau_0}^{\tau_2} d\tau_1 A(\tau_n)\dots A(\tau_1)\right] A(\tau_0)\varphi(\tau_0) \\ &= \int_s^t u_n^A(t,\sigma)\left[A(\sigma)\varphi(\sigma)\right] d\sigma.\end{aligned}$$

Similarly,

$$\begin{aligned}\left(\Lambda_s^{n}H\right)(t) &= \int_s^t d\tau_n \cdots \int_s^{\tau_2} d\tau_1 \int_s^{\tau_1} d\tau_0 A(\tau_n)\dots A(\tau_1) \int_s^{\tau_1} h(\tau_0)\, d\tau_0 \\ &= \int_s^t d\tau_0 \left[\int_{\tau_0}^t d\tau_n \int_{\tau_0}^{\tau_n} d\tau_{n-1} \cdots \int_{\tau_0}^{\tau_2} d\tau_1 A(\tau_n)\dots A(\tau_1)\right] h(\tau_0) \\ &= \int_s^t u_n^A(t,\sigma)\, h(\sigma)\, d\sigma.\end{aligned}$$

■

Proposition 2.36 (Dual Equation). *The fundamental solution, u^A also satisfies*

$$u^A(t,s) = \mathbf{1} + \int_s^t u^A(t,\sigma)\, A(\sigma)\, d\sigma \tag{2.25}$$

which is equivalent to solving the ODE,

$$\frac{d}{ds}u^A(t,s) = -u^A(t,s)\, A(s) \text{ with } u^A(t,t) = \mathbf{1}. \tag{2.26}$$

Proof. Summing Eq. (2.23) on n shows,

$$\begin{aligned}\sum_{n=0}^\infty \left(\Lambda_s^{n+1}\varphi\right)(t) &= \sum_{n=0}^\infty \int_s^t u_n^A(t,\sigma)\, A(\sigma)\,\varphi(\sigma)\, d\sigma \\ &= \int_s^t \sum_{n=0}^\infty u_n^A(t,\sigma)\, A(\sigma)\,\varphi(\sigma)\, d\sigma \\ &= \int_s^t u^A(t,\sigma)\, A(\sigma)\,\varphi(\sigma)\, d\sigma\end{aligned}$$

and hence

$$\begin{aligned}\left((\mathcal{I}-\Lambda_s)^{-1}\varphi\right)(t) &= \varphi(t) + \sum_{n=0}^\infty \left(\Lambda_s^{n+1}\varphi\right)(t) \\ &= \varphi(t) + \int_s^t u^A(t,\sigma)\, A(\sigma)\,\varphi(\sigma)\, d\sigma \end{aligned}\tag{2.27}$$

which specializes to Eq. (2.25) when $\varphi(t) = \mathbf{1}$.Differentiating Eq. (2.25) on s then gives Eq. (2.26). Another proof of Eq. (2.26) may be given using Proposition 2.33 to conclude that $u(t,s) = u(s,t)^{-1}$ and then differentiating this equation shows

$$\begin{aligned}\frac{d}{ds}u(t,s) = \frac{d}{ds}u(s,t)^{-1} &= -u(s,t)^{-1}\left(\frac{d}{ds}u(s,t)\right)u(s,t)^{-1} \\ &= -u(s,t)^{-1}A(s)\,u(s,t)\,u(s,t)^{-1} = -u(s,t)^{-1}A(s).\end{aligned}$$

■

Theorem 2.37 (Duhamel's principle). *The unique solution to Eq. (2.13) is*

$$y(t) = u^A(t,s)\, x + \int_s^t u^A(t,\sigma)\, h(\sigma)\, d\sigma. \tag{2.28}$$

Proof. First Proof. Let

$$\varphi(t) = x + H(t) \text{ with } H(t) = \int_s^t h(\tau)\, d\tau.$$

Then we know that the unique solution to Eq. (2.13) is given by

$$\begin{aligned}y &= (\mathcal{I}-\Lambda_s)^{-1}\varphi = (\mathcal{I}-\Lambda_s)^{-1}x + (\mathcal{I}-\Lambda_s)^{-1}H \\ &= u^A(\cdot,s)\, x + \sum_{n=0}^\infty \Lambda_s^n H,\end{aligned}$$

where by summing Eq. (2.24),

$$\begin{aligned}\left(\left(\mathcal{I}-\Lambda_s^A\right)^{-1}H\right)(t) &= \sum_{n=0}^\infty \left(\Lambda_s^n H\right)(t) = \sum_{n=0}^\infty \int_s^t u_n^A(t,\sigma)\, h(\sigma)\, d\sigma \\ &= \int_s^t \sum_{n=0}^\infty u_n^A(t,\sigma)\, h(\sigma)\, d\sigma = \int_s^t u^A(t,\sigma)\, h(\sigma)\, d\sigma \end{aligned}\tag{2.29}$$

and the proof is complete.

Second Proof. We need only verify that y defined by Eq. (2.28) satisfies Eq. (2.13). The main point is that the chain rule, FTC, and differentiation past the integral implies

$$\begin{aligned}
&\frac{d}{dt}\int_s^t u^A(t,\sigma)\,h(\sigma)\,d\sigma\\
&\quad=\frac{d}{d\varepsilon}|_0\int_s^{t+\varepsilon} u^A(t,\sigma)\,h(\sigma)\,d\sigma+\frac{d}{d\varepsilon}|_0\int_s^t u^A(t+\varepsilon,\sigma)\,h(\sigma)\,d\sigma\\
&\quad=u^A(t,t)\,h(t)+\int_s^t\frac{d}{dt}u^A(t,\sigma)\,h(\sigma)\,d\sigma\\
&\quad=h(t)+\int_s^t A(t)\,u^A(t,\sigma)\,h(\sigma)\,d\sigma\\
&\quad=h(t)+A(t)\int_s^t u^A(t,\sigma)\,h(\sigma)\,d\sigma.
\end{aligned}$$

Thus it follows that

$$\begin{aligned}
\dot{y}(t)&=A(t)\,u^A(t,s)\,x+A(t)\int_s^t u^A(t,\sigma)\,h(\sigma)\,d\sigma+h(t)\\
&=A(t)\,y(t)+h(t)\ \text{ with } y(s)=x.
\end{aligned}$$

■

The last main result of this section is to show that $u^A(t,s)$ is a differentiable function of A.

Theorem 2.38. *The map, $A\to u^A(t,s)$ is differentiable and moreover,*

$$\partial_B u^A(t,s)=\int_s^t u^A(t,\sigma)\,B(\sigma)\,u^A(\sigma,s)\,d\sigma. \tag{2.30}$$

Proof. Since $\partial_B\Lambda_s^A=\Lambda_s^B$ and

$$u^A(\cdot,s)=\left(\mathcal{I}-\Lambda_s^A\right)^{-1}\mathbf{1}$$

we conclude form Exercise 2.2 that

$$\partial_B u^A(\cdot,s)=\left(\mathcal{I}-\Lambda_s^A\right)^{-1}\Lambda_s^B\left(\mathcal{I}-\Lambda_s^A\right)^{-1}\mathbf{1}.$$

Equation (2.30) now follows from Eq. (2.29) with $h(\sigma)=B(\sigma)\,u^A(\sigma,s)$ so that and

$$H(t)=\int_s^t B(\sigma)\,u^A(\sigma,s)\,d\sigma=\left(\Lambda_s^B\left(\mathcal{I}-\Lambda_s^A\right)^{-1}\mathbf{1}\right)(t).$$

■

Remark 2.39 (Constant coefficient case). When $A(t)=A$ is constant, then

$$u_n^A(t,s)=\int_s^t d\tau_n\int_s^{\tau_n}d\tau_{n-1}\cdots\int_s^{\tau_2}d\tau_1 A^n=\frac{(t-s)^n}{n!}A^n$$

and hence $u^A(t,s)=e^{(t-s)A}$. In this case Eqs. (2.28) (2.30) reduce to

$$y(t)=e^{(t-s)A}x+\int_s^t e^{(t-\sigma)A}h(\sigma)\,d\sigma,$$

and for $B\in\mathcal{A}$,

$$\partial_B e^{(t-s)A}=\int_s^t e^{(t-\sigma)A}B(\sigma)\,e^{(\sigma-s)A}d\sigma.$$

Taking $s=0$ in this last equation gives the familiar formula,

$$\partial_B e^{tA}=\int_0^t e^{(t-\sigma)A}B(\sigma)\,e^{\sigma A}d\sigma.$$

2.4 Logarithms

Our goal in this section is to find an explicit local inverse to the exponential function, $A\to e^A$ for A near zero. The existence of such an inverse can be deduced from the inverse function theorem although we will not need this fact here. We begin with the real variable fact that

$$\ln(1+x)=\int_0^1\frac{d}{ds}\ln(1+sx)\,ds=\int_0^1 x\,(1+sx)^{-1}\,ds.$$

Definition 2.40. *When $A\in\mathcal{A}$ satisfies $1+sA$ is invertible for $0\le s\le 1$ we define*

$$\ln(1+A)=\int_0^1 A\,(1+sA)^{-1}\,ds. \tag{2.31}$$

The invertibility of $1+sA$ for $0\le s\le 1$ is satisfied if;

1. A is nilpotent, i.e. $A^N=0$ for some $N\in\mathbb{N}$ or more generally if
2. $\sum_{n=0}^\infty\|A^n\|<\infty$ (for example assume that $\|A\|<1$), of
3. if X is a Hilbert space and $A^*=A$ with $A\ge 0$.

In the first two cases

$$(1+sA)^{-1}=\sum_{n=0}^\infty(-s)^n A^n.$$

Proposition 2.41. *If $1+sA$ is invertible for $0\le s\le 1$, then*

$$\partial_B\ln(1+A)=\int_0^1(1+sA)^{-1}B\,(1+sA)^{-1}\,ds. \tag{2.32}$$

If $0=[A,B]:=AB-BA$, Eq. (2.32) reduces to

$$\partial_B\ln(1+A)=B\,(1+A)^{-1}. \tag{2.33}$$

Proof. Differentiating Eq. (2.31) shows

$$\begin{aligned}\partial_B \ln(1+A) &= \int_0^1 \left[B(1+sA)^{-1} - A(1+sA)^{-1} sB(1+sA)^{-1}\right] ds \\ &= \int_0^1 \left[B - sA(1+sA)^{-1} B\right](1+sA)^{-1} ds.\end{aligned}$$

Combining this last equality with

$$sA(1+sA)^{-1} = (1+sA-1)(1+sA)^{-1} = 1-(1+sA)^{-1}$$

gives Eq. (2.32). In case $[A,B]=0$,

$$\begin{aligned}(1+sA)^{-1} B (1+sA)^{-1} &= B(1+sA)^{-2} \\ &= B\frac{d}{ds}\left[-A^{-1}(1+sA)^{-1}\right]\end{aligned}$$

and so by the fundamental theorem of calculus

$$\begin{aligned}\partial_B \ln(1+A) &= B\int_0^1 (1+sA)^{-2} ds = B\left[-A^{-1}(1+sA)^{-1}\right]_{s=0}^{s=1} \\ &= B\left[A^{-1} - A^{-1}(1+A)^{-1}\right] = BA^{-1}\left[1-(1+A)^{-1}\right] \\ &= B\left[A^{-1}(1+A) - A^{-1}\right](1+A)^{-1} = B(1+A)^{-1}.\end{aligned}$$

■

Corollary 2.42. *Suppose that $t \to A(t) \in \mathcal{A}$ is a C^1 – function $1+sA(t)$ is invertible for $0 \le s \le 1$ for all $t \in J = (a,b) \subset \mathbb{R}$. If $g(t) := 1+A(t)$ and $t \in J$, then*

$$\frac{d}{dt}\ln(g(t)) = \int_0^1 (1-s+sg(t))^{-1} \dot{g}(t)(1-s+sg(t))^{-1} ds. \tag{2.34}$$

Moreover if $[A(t), A(\tau)] = 0$ for all $t,\tau \in J$ then,

$$\frac{d}{dt}\ln(g(t)) = \dot{A}(t)(1+A(t))^{-1}. \tag{2.35}$$

Proof. Differentiating past the integral and then using Eq. (2.32) gives

$$\begin{aligned}\frac{d}{dt}\ln(g(t)) &= \int_0^1 (1+sA(t))^{-1} \dot{A}(t)(1+sA(t))^{-1} ds \\ &= \int_0^1 (1+s(g(t)-1))^{-1} \dot{g}(t)(1+s(g(t)-1))^{-1} ds \\ &= \int_0^1 (1-s+sg(t))^{-1} \dot{g}(t)(1-s+sg(t))^{-1} ds.\end{aligned}$$

For the second assertion we may use Eq. (2.33) instead Eq. (2.32) in order to immediately arrive at Eq. (2.35). ■

Theorem 2.43. *If $A \in \mathcal{A}$ satisfies, $1+sA$ is invertible for $0 \le s \le 1$, then*

$$e^{\ln(I+A)} = I + A. \tag{2.36}$$

If $C \in \mathcal{A}$ satisfies $\sum_{n=1}^\infty \frac{1}{n!}\|C^n\|^n < 1$ (for example assume $\|C\| < \ln 2$, i.e. $e^{\|C\|} < 2$), then

$$\ln e^C = C. \tag{2.37}$$

This equation also holds of C is nilpotent or if X is a Hilbert space and $C = C^*$- with $C \ge 0$.

Proof. For $0 \le t \le 1$ let

$$C(t) = \ln(I+tA) = t\int_0^1 A(1+stA)^{-1} ds.$$

Since $[C(t), C(\tau)] = 0$ for all $\tau, t \in [0,1]$, if we let $g(t) := e^{C(t)}$, then

$$\dot{g}(t) = \frac{d}{dt}e^{C(t)} = \dot{C}(t)e^{C(t)} = A(1+tA)^{-1} g(t) \text{ with } g(0) = I.$$

Noting that $g(t) = 1+tA$ solves this ordinary differential equation, it follows by uniqueness of solutions to ODE's that $e^{C(t)} = g(t) = 1+tA$. Evaluating this equation at $t=1$ implies Eq. (2.36).

Now let $C \in \mathcal{A}$ as in the statement of the theorem and for $t \in \mathbb{R}$ set

$$A(t) := e^{tC} - 1 = \sum_{n=1}^\infty \frac{t^n}{n!}C^n.$$

Therefore,

$$1 + sA(t) = 1 + s\sum_{n=1}^\infty \frac{t^n}{n!}C^n$$

with

$$\left\| s\sum_{n=1}^\infty \frac{t^n}{n!}C^n \right\| \le s\sum_{n=1}^\infty \frac{t^n}{n!}\|C^n\|^n < 1 \text{ for } 0 \le s,t \le 1.$$

Because of this observation, $\ln\left(e^{tC}\right) := \ln(1+A(t))$ is well defined and because $[A(t), A(\tau)] = 0$ for all τ and t we may use Eq. (2.35) to learn,

$$\frac{d}{dt}\ln\left(e^{tC}\right) := \dot{A}(t)(1+A(t))^{-1} = Ce^{tC}e^{-tC} = C \text{ with } \ln\left(e^{0C}\right) = 0.$$

The unique solution to this simple ODE is $\ln\left(e^{tC}\right) = tC$ and evaluating this at $t=1$ gives Eq. (2.37). ■

2.5 C^*-algebras

We now are going to introduce the notion of "star" structure on a complex Banach algebra. We will be primarily motivated by the example of closed $*$-sub-algebras of the bounded linear operators on (in) a Hilbert space. For the rest of this section and essentially the rest of these notes we will assume that $\mathcal{B}$ is a **complex** Banach algebra.

Definition 2.44. *An **involution** on a complex Banach algebra, $\mathcal{B}$, is a map $a \in \mathcal{B} \to a^* \in \mathcal{B}$ satisfying:*

1. *involutory* $a^{**} = a$
2. *additive* $(a+b)^* = a^* + b^*$
3. *conjugate homogeneous* $(\lambda a)^* = \overline{\lambda} a^*$
4. *anti–automorphic* $(ab)^* = b^* a^*$.

If $*$ is an involution on $\mathcal{B}$ and $\mathbf{1} \in \mathcal{B}$, then automatically we have $\mathbf{1}^* = \mathbf{1}$. Indeed, applying the involution to the identity, $\mathbf{1}^* = \mathbf{1} \cdot \mathbf{1}^*$ gives

$$\mathbf{1} = \mathbf{1}^{**} = (\mathbf{1} \cdot \mathbf{1}^*)^* = \mathbf{1}^{**} \cdot \mathbf{1}^* = \mathbf{1} \cdot \mathbf{1}^* = \mathbf{1}^*.$$

For the rest of this section we let $\mathcal{B}$ be a Banach algebra with involution, $*$.

Definition 2.45. *If $a \in \mathcal{B}$ we say;*

1. *a is **hermitian** if $a = a^*$.*
2. *a is **normal** if $a^* a = a a^*$, i.e. $[a, a^*] = 0$ where $[a,b] := ab - ba$.*
3. *a is **unitary** if $a^* = a^{-1}$.*

Example 2.46. Let G be a discrete group and $\mathcal{B} = \ell^1(G, \mathbb{C})$ as in Proposition 2.4. We define $*$ on $\mathcal{B}$ so that $\delta_g^* = \delta_{g^{-1}}$. In more detail if $f = \sum_{g \in G} f(g)\, \delta_g$, then

$$f^* = \sum_{g \in G} \overline{f(g)} \delta_g^* = \sum_{g \in G} \overline{f(g)} \delta_{g^{-1}} \implies f^*(g) := \overline{f(g^{-1})}.$$

Notice that

$$(\delta_g \delta_h)^* = \delta_{gh}^* = \delta_{(gh)^{-1}} = \delta_{h^{-1} g^{-1}} = \delta_{h^{-1}} \delta_{g^{-1}} = \delta_h^* \delta_g^*.$$

Using this or by direct verification one shows $(f \cdot h)^* = h^* \cdot f^*$. The other properties of $*$ – are now easily verified.

Definition 2.47 (C^*-condition). *A Banach $*$ algebra $\mathcal{B}$ is*

1. *$*$ multiplicative if $\|a^* a\| = \|a^*\| \ \|a\|$*
2. *$*$ isometric if $\|a^*\| = \|a\|$*
3. *$*$ quadratic if $\|a^* a\| = \|a\|^2$.*

   We refer to item 3. as the C^*-condition.

Lemma 2.48. *Conditions 1) and 2) in Definition 2.47 are equivalent to condition 3), i.e. $*$ is multiplicative & isometric iff $*$ is quadratic.*

Proof. Clearly $*$ is multiplicative & isometric implies that $*$ is quadratic. For the reverse implication; if $\|a^* a\| = \|a\|^2$ for all $a \in \mathcal{B}$, then

$$\|a\|^2 \le \|a^*\| \, \|a\| \implies \|a\| \le \|a^*\|.$$

Replacing a by a^* in this inequality shows $\|a\| = \|a^*\|$ and hence Thus $\|a^* a\| = \|a\|^2 = \|a\| \, \|a^*\|$ ∎

Remark 2.49. It is fact the case that seemingly weaker condition 1. in Definition 2.47 by itself implies condition 3 but the implication 1. $\implies$ 3. is quite non-trivial. See Theorem 16.1 on page 45 of [9]. [That this result holds under the additional assumption that $\mathcal{B}$ is commutative and "symmetric" is contained in Theorem 8.14 below.] Historically condition 1. is called the C^*-condition on a norm and condition 3. is called the B^*-condition on a norm, see the Wikipedia[1] article for information about B^*-algebras being the same as C^*-algebras.

Definition 2.50. *A C^*-**algebra** is a $*$ quadratic algebra, i.e. $\mathcal{B}$ is a C^*-algebra if $\mathcal{B}$ is a Banach algebra with involution $*$ such that $\|a^* a\| = \|a\|^2$ for all $a \in \mathcal{B}$.*

The next proposition gives the primary motivating examples of C^*-algebras.

Proposition 2.51. *Let H be a Hilbert space and $\mathcal{B}$ be a $*-$closed and operator norm-closed sub-algebra of $B(H)$, where A^* is the adjoint of $A \in B(H)$. Then $(\mathcal{B}, *)$ is a C^*–algebra.*

Proof. From the basic properties of the adjoint, $B(H)$, is a $*$-algebra so the main point is to verify the C^*-condition, which we now do in two steps.

1. If $k \in H$, then

$$\begin{aligned} \|A^* k\|_H &= \sup_{\|h\|_H = 1} |\langle A^* k, h\rangle| = \sup_{\|h\|_H=1} |\langle k, Ah\rangle| \\ &\le \sup_{\|h\|_H=1} \|k\|_H \, \|Ah\|_H = \|A\|_{op} \, \|k\|_H. \end{aligned}$$

   From this inequality it follows that $\|A^*\|_{op} \le \|A\|_{op}$. Applying this inequality with A replaced by A^* shows $\|A\|_{op} \le \|A^*\|_{op}$ and hence $\|A^*\| = \|A\|$ which prove that $*$ is an isometry.

[1] https://en.wikipedia.org/wiki/C*-algebra

2. Given item 1., we find the inequality,

$$\|A^*A\| \le \|A^*\|\|A\| = \|A\|^2.$$

However we also have for any $x \in H$ that

$$\|Ax\|^2 = \langle A^*Ax, x\rangle \le \|A^*A\| \|x\|^2 \implies \|A\|^2 \le \|A^*A\|.$$

Combining the last two displayed inequalities verifies the C^*-condition, $\|A^*A\| = \|A\|^2$.

Alternate proof. Using the Rayleigh quotient in Theorem A.26, we have for any $A \in B(H)$,

$$\|A\|_{op}^2 = \sup_{\|f\|=1} \|Af\|^2 = \sup_{\|f\|=1} \langle Af, Af\rangle = \sup_{\|f\|=1} \langle A^*Af, f\rangle = \|A^*A\|_{op}.$$

■

Remark 2.52. Irvine Segal's original definition of C^*-algebra was in fact a $*$-**C**losed sub-algebra of $B(H)$ for some Hilbert space H. The letter "C" used here indicated that the sub-algebra was closed under the operator norm topology. Later, the definition was abstracted to the C^*-algebra definition we have given above. It is however a (standard) fact that by the "GNS construction," every abstract C^*-algebra may be "represented" by a "concrete" (i.e. sub-algebra of $B(H)$) C^*-algebra. The "GNS construction" along with appropriate choices of states shows that in fact every abstract C^*-algebra has a faithful representation as a C^*-subalgebra in the sense of Segal, see Conway [7, Theorem 5.17, p. 253]. The B^*-terminology has fallen out of favour. [Incidentally, a von Neumann algebra is a w.o.t. (or s.o.t.) closed $*$-subalgebra of $B(H)$ and is often called a W^*- algebra.] See the Appendix 2.5.4 to this section for some examples of embedding commutative C^*-algebras into $B(H)$.

2.5.1 Examples

Here are a few more examples of C^*-algebras.

Example 2.53. If X is a compact Hausdorff space then $\mathcal{B} := C(X, \mathbb{C})$ with

$$\|f\| = \sup_{x\in X} |f(x)| \text{ and } f^*(x) := \overline{f(x)}$$

is a C^*-algebra with identity. If X is only locally compact, then $\mathcal{B} := C_0(X, \mathbb{C})$ is a C^*-algebra without identity. We will see that these are, up to isomorphism, all of the commutative C^*-algebras.

Example 2.54. Let $\mathcal{B}$ be a C^*-subalgebra of $B(H)$ and then set

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} : A \in \mathcal{B} \right\} \subset B(H \oplus H).$$

Clearly,

$$\mathcal{B} \ni A \to \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \in \mathcal{B}_1$$

is a C^*-isomorphism. This example shows that $\mathcal{B}$ and $\mathcal{B}_1$ are the same as abstract C^*-algebras. This example shows that the C^*-algebra structure of $\mathcal{B}$ is not necessarily the whole story when one cares about how $\mathcal{B}$ is embedded inside of the bounded operators on a Hilbert space.

Example 2.55. If $(\Omega, \mathcal{F}, \mu)$ is a measure space then $L^\infty(\mu) := L^\infty(\Omega, \mathcal{F}, \underline{\mu : \mathbb{C}})$ is a commutative complex C^*-algebra with identity. Again we let $f^*(\omega) = \overline{f(\omega)}$. The C^*-condition is

$$\begin{aligned} \|f^*f\| &= \sup\left\{M > 0 : |f|^2 \le M \text{ a.e. }\right\} \\ &= \sup\left\{M^2 > 0 : |f| \le M \text{ a.e. }\right\} = \|f\|^2. \end{aligned}$$

Notation 2.56 (Bounded Multiplication Operators) *Given a measure space $(\Omega, \mathcal{F}, \mu)$ and a bounded measurable function $q : \Omega \to \mathbb{C}$, let $M_q : L^2(\mu) \to L^2(\mu)$ denote the operation of multiplication by q, i.e. $M_q : L^2(\mu) \to L^2(\mu)$ is defined by $M_q f = qf$ for all $f \in L^2(\mu)$.*

Definition 2.57 (Atoms). *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. A set $A \in \mathcal{F}$ is said to be an* ***atom of*** *μ if $\mu(A) > 0$ and $\mu(A \cap B)$ is either $\mu(A)$ or 0 for every $B \in \mathcal{F}$. We say A is an* ***infinite atom*** *if it is an atom such that $\mu(A) = \infty$.*

Theorem 2.58. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space with no infinite atoms and*

$$\mathcal{B} = \{M_f : f \in L^\infty(\mu)\} =: M_{L^\infty(\mu)} \tag{2.38}$$

which we view as a $*$-subalgebra of $B\left(L^2(\mu)\right)$. Then $\mathcal{B}$ is a C^*-subalgebra of $B\left(L^2(\mu)\right)$ and the map,

$$L^\infty(\mu) \ni f \xrightarrow{M_{(\cdot)}} M_f \in \mathcal{B} \tag{2.39}$$

is a C^*-isometric isomorphism. Explicitly that isometry condition means,

$$\|M_f\|_{op} = \|f\|_\infty \text{ for all } f \in L^\infty(\mu). \tag{2.40}$$

Proof. Given $f, g \in L^\infty(\mu)$ and $\lambda \in \mathbb{C}$, one readily shows,

$$M_f + M_g = M_{f+g},\ M_{\lambda f} = \lambda M_f,\ M_f M_g = M_{fg}, \text{ and } M_f^* = M_{\bar{f}},$$

i.e. $M_{(\cdot)} : L^\infty(\mu) \to B(L^2(\mu))$ is a $*$-algebra homomorphism. Since $\|M_f g\|_2 = \|fg\|_2 \le \|f\|_\infty \|g\|_2$, it follows that $\|M_f\|_{op} \le \|f\|_\infty$ with equality when $\|f\|_\infty = 0$. For the reverse inequality we may assume that $\|f\|_\infty > 0$. If $0 < k < \|f\|_\infty$, then $\mu(|f| \ge k) > 0$ and since μ has not infinite atoms we may find $A \subset \{|f| \ge k\}$ such that $0 < \mu(A) < \infty$. It then follows that $\|1_A\|_2 = \sqrt{\mu(A)} \in (0, \infty)$ and

$$\|M_f\|_{op} \ge \frac{\|f 1_A\|_2}{\|1_A\|_2} \ge k.$$

As this holds for all $k < \|f\|_\infty$ we conclude that $\|M_f\|_{op} \ge \|f\|_\infty$ and so Eq. (2.40) has been proved.

Since $\mathcal{B}$ is the image of $M_{(\cdot)}$, $M_{(\cdot)}$ is a linear isometry, and $L^\infty(\mu)$ is complete, it follows that $\mathcal{B}$ is complete and hence closed in $B(L^2(\mu))$. Thus $\mathcal{B}$ is a C^*-subalgebra of $B(L^2(\mu))$ and the proof is done. ∎

Example 2.59. If $T_1, \dots, T_n \in B(H)$, let $\mathcal{A}(T_1, \dots, T_n)$ be the smallest subalgebra of $B(H)$ containing $\{T_1, \dots, T_n\}$, i.e. $\mathcal{A}$ consists of linear combination of words in $\{T_1, \dots, T_n\}$. With this notation, $\mathcal{A}(T_1, \dots, T_n, T_1^*, \dots, T_n^*)$ is the smallest $*$ -sub-algebra of $B(H)$ which contains $\{T_1, \dots, T_n\}$. We let

$$C^*(T_1, \dots, T_n) := \overline{\mathcal{A}(T_1, \dots, T_n, T_1^*, \dots, T_n^*)}^{\|\cdot\|_{op}}$$

be the C^*-algebra generated by $\{T_1, \dots, T_n\}$.

Example 2.60. If $T_1, \dots, T_n \in B(H)$ are commuting self-adjoint operators, then

$$\mathcal{A}(T_1, \dots, T_n) := \{p(T_1, \dots, T_n) : p \in \mathbb{C}[z_1, \dots, z_n] \ni p(\mathbf{0}) = 0\}$$

is a commutative $*-$ sub-algebra of $B(H)$. We also have

$$\mathcal{A}(I, T_1, \dots, T_n) := \{p(T_1, \dots, T_n) : p \in \mathbb{C}[z_1, \dots, z_n]\}$$

where if $p(z_1, \dots z_n) = p_0 + q(z_1, \dots z_n)$ with $q(\mathbf{0}) = 0$ we let

$$p(T_1, \dots, T_n) = p_0 I + q(T_1, \dots, T_n).$$

For most of this chapter we will mostly interested in the commutative $*$-sub-algebra, $\mathcal{A}(I, T)$ where $T \in B(H)$ with $T^* = T$.

Proposition 2.61. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $\mathcal{B} = L^\infty(\mu)$ be the C^*-algebra of essentially bounded functions, $\{f_j\}_{j=1}^n \subset \mathcal{B}$, $\mathbf{f} = (f_1, \dots, f_n) : \Omega \to \mathbb{C}^n$, and $\mathrm{essran}_\mu(\mathbf{f})$ be the essential range of $\mathbf{f}$ (see Definition 1.32). Then $\hat{\mathbf{f}} : C(\mathrm{essran}_\mu(\mathbf{f})) \to L^\infty(\mu)$ defined by $\hat{\mathbf{f}}(\psi) = \psi(\mathbf{f})$ for all $\psi \in C(\mathrm{essran}_\mu(\mathbf{f}))$ is an isometric C^*-isomorphism onto $C^*(\mathbf{f}, 1)$.*

Proof. Let us first show that

$$\|\psi(\mathbf{f})\|_\infty = \|\psi\|_{C(\mathrm{essran}_\mu(\mathbf{f}))} \text{ for all } \psi \in C(\mathrm{essran}_\mu(\mathbf{f})). \tag{2.41}$$

It is clear that $\|\psi(\mathbf{f})\|_\infty \le \|\psi\|_{C(\mathrm{essran}_\mu(\mathbf{f}))}$. If $M < \|\psi\|_{C(\mathrm{essran}_\mu(\mathbf{f}))}$, then there exists $\mathbf{z} \in \mathrm{essran}_\mu(\mathbf{f})$ so that $M < |\psi(\mathbf{z})|$ and for this $\mathbf{z}$, $\mu(\|\mathbf{f} - \mathbf{z}\| < \varepsilon) > 0$ for all $\varepsilon > 0$. By the continuity of ψ there exists $\varepsilon > 0$ so that $|\psi(\mathbf{w})| > M$ for $\|\mathbf{w} - \mathbf{z}\| < \varepsilon$ and hence

$$\mu(|\psi(\mathbf{f})| > M) \ge \mu(\|\mathbf{f} - \mathbf{z}\| < \varepsilon) > 0$$

from which it follows that $\|\psi(\mathbf{f})\|_\infty \ge M$. As $M < \|\psi\|_{C(\mathrm{essran}_\mu(\mathbf{f}))}$ was arbitrary, it follows that $\|\psi(\mathbf{f})\|_\infty \ge \|\psi\|_{C(\mathrm{essran}_\mu(\mathbf{f}))}$ and Eq. (2.41) is proved.

Let $\mathcal{B}_0 := \hat{\mathbf{f}}(C(\mathrm{essran}_\mu(\mathbf{f})))$ be the image of $\hat{\mathbf{f}}$ which, as $\hat{\mathbf{f}}$ is a isometric C^*-homomorphism, is a closed $*$-subalgebra of $\mathcal{B}$. To finish the proof we must show $\mathcal{B}_0 = C^*(\mathbf{f}, 1)$.

Given $\psi \in C(\mathrm{essran}_\mu(\mathbf{f}))$, there exists $p_k \in \mathbb{C}[z_1, \dots, z_n, \bar{z}_1, \dots, \bar{z}_n]$ such that

$$\lim_{n\to\infty} \max_{\mathbf{z} \in \mathrm{essran}_\mu(\mathbf{f})} |\psi(\mathbf{z}) - p_n(\mathbf{z}, \bar{\mathbf{z}})| = 0.$$

Using

$$p(\mathbf{f}, \bar{\mathbf{f}}) := p(f_1, \dots, f_n, \bar{f}_1, \dots, \bar{f}_n) \in C^*(\mathbf{f}, 1),$$

along with the isometry property in Eq. (2.41), it follows that

$$\left\|\psi(\mathbf{f}) - p_k(\mathbf{f}, \bar{\mathbf{f}})\right\|_\infty = \max_{\mathbf{z} \in \mathrm{essran}_\mu(\mathbf{f})} |\psi(\mathbf{z}) - p_k(\mathbf{z}, \bar{\mathbf{z}})| \to 0 \text{ as } k \to \infty,$$

which implies $\psi(\mathbf{f}) \in C^*(\mathbf{f}, 1)$, i.e. $\mathcal{B}_0 \subset C^*(\mathbf{f}, 1)$. For the opposite inclusion simply observe that if we let $\psi_i(\mathbf{z}) = z_i$ for $i \in [n]$, then $f_i = \hat{\mathbf{f}}(\psi_i) \in \mathcal{B}_0$ for each $i \in [n]$. As $\mathcal{B}_0$ is a C^*-algebra we must also have that $C^*(\mathbf{f}, 1) \subset \mathcal{B}_0$ and the proof is complete. ∎

Remark 2.62. It is also easy to verify that

$$C^*(\mathbf{f}) = \{\psi(\mathbf{f}) : \psi \in C(\mathrm{essran}_\mu(\mathbf{f})) \ \ni\ \psi(0, \dots, 0) = 0\}$$

and that

$$\{\psi \in C(\mathrm{essran}_\mu(\mathbf{f})) \ \ni\ \psi(0, \dots, 0) = 0\} \to \psi(f_1, \dots, f_n) \in C^*(\mathbf{f})$$

is a isomorphism of C^*-algebras." We leave the details to the reader.

The next result is a direct corollary of Theorem 2.58 and Proposition 2.61.

Corollary 2.63. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space with no infinite atoms, $\mathcal{B} = M_{L^\infty(\mu)}$ as in Theorem 2.58, $\{f_j\}_{j=1}^n \subset L^\infty(\mu)$, and $\mathbf{f} = (f_1, \dots, f_n) : \Omega \to \mathbb{C}^n$. Then the map*

$$C(\mathrm{essran}_\mu(\mathbf{f})) \ni \psi \to M_{\psi(\mathbf{f})} \in C^*(M_{f_1}, \dots, M_{f_n}, 1) \subset \mathcal{B}$$

is an isometric isomorphism of C^*-algebras.

2.5.2 Some Consequences of the C^*-condition

Let us now explore some of the consequence of the C^*-condition. The following simple lemma turns out to be a very important consequence of the C^*-condition which will be used in Proposition 4.3 in order to show;

$$\|a\| = \sup\{|\lambda| : \lambda \in \sigma(a)\} \text{ when } a \text{ is normal.}$$

Lemma 2.64. *If $\mathcal{B}$ is a C^*-algebra and b is a normal element of $\mathcal{B}$, then $\|b^2\| = \|b\|^2$.*

Proof. This is easily proved as follows;

$$\left\|b^2\right\|^2 \overset{C^*\text{-cond.}}{=} \left\|\left(b^2\right)^* b^2\right\| \overset{\text{Normal}}{=} \left\|(b^*b)^2\right\| \overset{C^*\text{-cond.}}{=} \|b^*b\|^2 \overset{C^*\text{-cond.}}{=} \|b\|^4.$$

■

Lemma 2.65. *If $\mathcal{B}$ is a unital C^*-algebra and $u \in \mathcal{B}$ is unitary, then $\|u\| = 1$. Moreover, if $u, v \in \mathcal{B}$ are unitary, then $\|uav\| = \|a\|$ for all $a \in \mathcal{B}$.*

Proof. Since $1 = u^*u$, it follows by the C^*-condition that $1 = \|1\| = \|u^*u\| = \|u\|^2$ from which it follows that $\|u\| = 1$. If $a \in \mathcal{B}$, then

$$\|uav\| \le \|u\| \|a\| \|v\| = \|a\|.$$

By replacing a by u^*av^* in the above inequality we also find that $\|a\| \le \|u^*av^*\|$. We may replace u by u^* and v by v^* in the last inequality in order to show $\|a\| \le \|uav\|$ which along with the previously displayed equation completes the proof. ■

Example 2.66. If $A \in \mathcal{B}$ is a C^*-algebra, then using the fact that $*$ is an isometry, it follows that

$$\left(e^A\right)^* = \sum_{n=0}^\infty \left(\frac{1}{n!} A^n\right)^* = \sum_{n=0}^\infty \frac{1}{n!} (A^*)^n = e^{A^*}.$$

Thus if $A^* = A$, we find

$$\left(e^{iA}\right)^* = e^{-iA^*} = e^{-iA} = \left(e^{iA}\right)^{-1},$$

which shows e^{iA} is unitary. This result is generalized in the following proposition.

Proposition 2.67. *Suppose that $\mathcal{B}$ is a C^*-algebra with identity and $t \to A(t) \in \mathcal{B}$ is continuous and $A(t)^* = -A(t)$ for all $t \in \mathbb{R}$. If $u(t)$ is the unique solution to*

$$\dot{u}(t) = A(t)u(t) \text{ with } u(0) = 1 \tag{2.42}$$

then $u(t)$ is unitary.

Proof. Let $u(t,s)$ denote the solution to

$$\dot{u}(t,s) = A(t)u(t,s) \text{ with } u(s,s) = 1$$

so that $u(t) = u(t,0)$. From Proposition 2.33 it follows that $u(t)^{-1} = u(0,t)$ and from Proposition 2.36 we conclude that

$$\frac{d}{dt}u(t)^{-1} = \frac{d}{dt}u(0,t) = -u(0,t)A(t) = -u(t)^{-1}A(t) = u(t)^{-1}A(t)^*.$$

On the other hand taking the adjoint of Eq. (2.42) shows

$$\dot{u}^*(t) = u(t)^* A(t)^* \text{ with } u^*(0) = 1.$$

So by uniqueness of solutions we conclude that $u^*(t) = u(t)^{-1}$. ■

Theorem 2.68 (Fuglede-Putnam Theorem, see Conway, p. 278). *Let $\mathcal{B}$ be a C^*-algebra with identity and M and N be normal elements in $\mathcal{B}$ and $B \in \mathcal{B}$ satisfy $NB = BM$, then $N^*B = BM^*$. In particular, taking $M = N$ implies $[N,B] = 0$ implies $[N^*, B] = 0$. [Note well that B is not assumed to be normal here.]*

Proof. Given $w \in \mathbb{C}$ let

$$u(t) := e^{twN} B e^{-twM}.$$

Then $u(0) = B$ and

$$\dot{u}(t) = we^{twN}[NB - BM]e^{-twM} = 0$$

and hence $u(t) = B$ for all t, i.e. $e^{wN}Be^{-wM} = B$ for all $w \in \mathbb{C}$.

Now for $z \in \mathbb{C}$ let $f : \mathbb{C} \to \mathcal{B}$ be the analytic function,

$$f(z) = e^{izN^*} B e^{-izM^*}.$$

Using what we have just proved and the normality assumptions[2] on N and M we have for any $w \in \mathbb{C}$ that

$$f(z) = e^{izN^*} e^{wN} B e^{-wM} e^{-izM^*} = e^{[izN^* + wN]} B e^{-[wM + izM^*]}.$$

We now take $w = i\bar{z}$ to find,

$$f(z) = e^{i[zN^* + \bar{z}N]} B e^{-i[\bar{z}M + zM^*]}$$

and hence by Example 2.66 and Lemma 2.65,

$$\|f(z)\| = \left\| e^{i[zN^* + \bar{z}N]} B e^{-i[\bar{z}M + zM^*]} \right\| = \|B\|$$

wherein we have used both, $zN^* + \bar{z}N$ and $\bar{z}M + zM^*$ are Hermitian elements. By an application of Liouville's Theorem (see Corollary 1.12) we conclude $f(z) = f(0) = B$ for all $z \in \mathbb{C}$, i.e.

$$e^{izN^*} B e^{-izM^*} = B.$$

Differentiating this identity at $z = 0$ then shows $N^* B = BM^*$. ■

Corollary 2.69. *Again suppose $\mathcal{B}$ is a unital C^*-algebra, $M \in \mathcal{B}$ is normal and $B \in \mathcal{B}$ is arbitrary. If $[M, B] = 0$, then $[\{M, M^*\}, B] = \{0\} = [\{M, M^*\}, B^*]$.*

Proof. By Theorem 2.68 we know that $0 = [M^*, B]$ and taking adjoints of this equation then shows $0 = -[M, B^*]$. Finally by one more application of Theorem 2.68 it follows that $[M^*, B^*] = 0$ as well. ■

Note well that under the assumption that M is normal and $[M, B] = 0$, $C^*(M, B, I)$ will be commutative iff B is normal.

Definition 2.70. *If $\mathcal{B}$ is a C^*-algebra and $\mathcal{S} \subset \mathcal{B}$ is a non-empty set, we define $C^*(\mathcal{S})$ to be the smallest C^*-subalgebra of $\mathcal{B}$. [Please note that we require $C^*(\mathcal{S})$ to be closed under $A \to A^*$.]*

Corollary 2.71. *Suppose that $\mathcal{B}$ is a unital C^*-algebra with identity and $\mathbf{T} := \{T_j\}_{j=1}^n \subset \mathcal{B}$ are commuting normal operators, then $\mathbf{T} \cup \mathbf{T}^* := \left\{T_j, T_j^*\right\}_{j=1}^n$ is a list of pairwise commuting operators and $C^*(\mathbf{T}, 1)$ is the norm closure of all elements of $\mathcal{B}$ of the form $p(\mathbf{T}, \mathbf{T}^*)$ where $p(z_1, \dots, z_n, w_1, \dots, w_n)$ is a polynomial in $2n$-variables. Moreover, $C^*(\mathbf{T}, 1)$ is a commutative C^*-subalgebra of $\mathcal{B}$.*

[2] The normality assumptions allows us to conclude $e^{[izN^* + wN]} = e^{izN^*} e^{wN}$

Remark 2.72. For the fun of it, here are two elementary proofs of Theorem 2.68 for $\mathcal{B} = B(H)$ when $\dim H < \infty$.

First proof. The key point here is that $H = \oplus_{\lambda \in \mathbb{C}}^{\perp} E_\lambda^M$ where $E_\lambda^M := \mathrm{Nul}(M - \lambda I)$ and for $u \in E_\lambda^M$ we have for $v \in E_\alpha^M$ that

$$\langle M^* u, v \rangle = \langle u, Mv \rangle = \bar{\alpha} \langle u, v \rangle$$

from which it follows that $\langle M^* u, v \rangle = 0$ if $\alpha \neq \lambda$ or if $\alpha = \lambda$ and $u \perp v$. Thus we may conclude that $M^* u = \bar{\lambda} u$ for all $u \in E_\lambda^M$. With this preparation, $NBu = BMu = B\lambda u = \lambda Bu$ and therefore $Bu \in E_\lambda^N$. Therefore it follows that

$$N^* Bu = \bar{\lambda} Bu = B\bar{\lambda} u = BM^* u.$$

As $u \in E_\lambda^M$ was arbitrary and $\lambda \in \mathbb{C}$ was arbitrary it follows that $N^* B = BM^*$.

Second proof. A key point of M being normal is that for all $\lambda \in \mathbb{C}$ and $u \in H$,

$$\begin{aligned}
\|(M - \lambda) u\|^2 &= \langle (M - \lambda) u, (M - \lambda) u \rangle = \left\langle u, (M - \lambda)^* (M - \lambda) u \right\rangle \\
&= \left\langle u, (M - \lambda)(M - \lambda)^* u \right\rangle = \left\langle (M - \lambda)^* u, (M - \lambda)^* u \right\rangle \\
&= \left\| (M - \lambda)^* u \right\|^2.
\end{aligned}$$

Thus if $\{u_j\}_{j=1}^{\dim H}$ is an orthonormal basis of eigenvectors of M with $Mu_j = \lambda_j u_j$ then $M^* u_j = \bar{\lambda}_j u_j$. Thus if we apply $NB = BM$ to u_j we find,

$$NBu_j = BMu_j = \lambda_j Bu_j$$

and therefore as N is normal, $N^* Bu_j = \bar{\lambda}_j Bu_j$. Since M is normal we also have

$$N^* Bu_j = B\bar{\lambda}_j u_j = BM^* u_j.$$

As this holds for all j, we conclude that $N^* B = BM^*$.

2.5.3 Symmetric Condition

Definition 2.73. *An involution $*$ in a Banach algebra $\mathcal{B}$ with unit is **symmetric** if $1 + a^* a$ is invertible for all $a \in \mathcal{B}$.*

Lemma 2.74. *If H is a complex Hilbert space, then $B(H)$, then $B(H)$ is symmetric. [It is in fact true that any C^*-subalgebra, $\mathcal{B}$, of $B(H)$ is symmetric but this requires more proof than we can give at this time. See Theorem 9.4 below for the missing ingredient.]*

Proof. It clearly suffices to show $B(H)$ is symmetric, i.e. that $I + A^*A$ is invertible for any $A \in B(H)$. The key point is that for any $h \in H$,

$$\|h\|^2 \le \|h\|^2 + \|Ah\|^2 = \langle (I + A^*A) h, h\rangle \le \|(I + A^*A) h\| \, \|h\|$$

and hence

$$\|(I + A^*A) h\| \ge \|h\|. \tag{2.43}$$

This inequality clearly shows $\mathrm{Nul}(I + A^*A) = \{0\}$ and that $I + A^*A$ has closed range, see Corollary 2.10. Therefore we conclude that

$$\mathrm{Ran}(I + A^*A) = \overline{\mathrm{Ran}(I + A^*A)} = \mathrm{Nul}(I + A^*A)^{\perp} = H$$

and so $I + A^*A$ is algebraically invertible and hence invertible in $B(H)$ by Lemma 2.9. In fact, because of Eq. (2.43) we have the estimate, $\left\|(I + A^*A)^{-1}\right\|_{op} \le 1$.

If we have Theorem 9.4 at our disposal, then we may conclude that $(I + A^*A)^{-1} \in C^*(A^*A, I) \subset C^*(A, I)$ and with this result we may assert that theorem holds for any C^*-subalgebra, $\mathcal{B}$, of $B(H)$. ∎

Example 2.75. Referring to Example 2.46 with $G = \mathbb{Z}$, we claim that $\ell^1(\mathbb{Z})$ with convolution for multiplication is an abelian $*$-Banach algebra which is not a C^*-algebra. For example, let $f := \delta_0 - \delta_1 - \delta_2$, then

$$\begin{aligned} f^* f &= (\delta_0 - \delta_{-1} - \delta_{-2})(\delta_0 - \delta_1 - \delta_2) \\ &= \delta_0 - \delta_1 - \delta_2 + (-\delta_{-1} + \delta_0 + \delta_1) + (-\delta_{-2} + \delta_{-1} + \delta_0) \\ &= 3\delta_0 - \delta_2 - \delta_{-2} \end{aligned}$$

and hence

$$\|f^* f\| = 3 + 1 + 1 = 5 < 9 = 3^2 = \|f\|^2.$$

As a consequence of Lemma 2.74 and assuming Remark 2.52, every C^*-algebra is symmetric[3] and so this example implies $\ell^1(\mathbb{Z})$ is not a C^*-algebra. See Remark 8.13 below for some more information about the symmetry condition on a Banach algebra. See Exercise 8.2 for more on this example.

2.5.4 Appendix: Embeddings of function C^*-algebras into $B(H)$

The next example is a special case of the GNS construction in disguise. See Remark 2.52 for more comments and references in this direction.

[3] We will explicitly prove this fact for commutative C^*-algebras below in Lemma 8.12.

Example 2.76. Suppose that X is a compact Hausdorff space, μ is counting measure on X, and $H = L^2(X, \mu)$. Then

$$\mathcal{C} := \{M_f \in B(H) : f \in C(X) := C(X, \mathbb{C})\} \subset B(H)$$

is a C^*-algebra. Indeed $\mathcal{C}$ is a $*-$ algebra since, $M_f + kM_g = M_{f+kg}$, $M_f M_g = M_{fg}$, and $M_f^* = M_{\bar{f}}$ for all $f, g \in C(X)$. Moreover, we have

$$\|M_f\|_{op} = \sup_{x \in X} |f(x)| = \|f\|_u \tag{2.44}$$

from which it follows that $\mathcal{C}$ is closed in $B(H)$ in the operator norm. In this case H may be a highly non-separable Hilbert space. However the above construction also works for any measure no infinite atom measure, μ on $\mathcal{B}_X$, such that $\mathrm{supp}(\mu) = X$. In particular μ is a σ-finite measure on open sets and X is separable, then $L^2(X, \mu)$ will be separable as well.

For an explicit choice of measure, $D = \{x_n\}_{n=1}^{\infty}$ is a countable dense subset of X, let

$$\mu := \sum_{n=1}^{\infty} \delta_{x_n}$$

in which case $\mathrm{supp}(\mu) = X$ and take $H = \hat{H} = L^2(X, \mathcal{B}_X, \mu)$ in the above construction. In this special case one directly checks Eq. (2.44) using,

$$\|M_f\|_{op} = \sup_{x \in D} |f(x)| = \sup_{x \in X} |f(x)| = \|f\|_u \ \ \forall f \in C(X).$$

2.6 Exercises

Exercise 2.8. To each $A \in \mathcal{A}$, we may define $L_A, R_A : \mathcal{A} \to \mathcal{A}$ by

$$L_A B = AB \text{ and } R_A B = BA \text{ for all } B \in \mathcal{A}.$$

Show $L_A, R_A \in L(\mathcal{A})$ and that

$$\|L_A\|_{L(\mathcal{A})} = \|A\|_{\mathcal{A}} = \|R_A\|_{L(\mathcal{A})}.$$

Exercise 2.9. Suppose that $A : \mathbb{R} \to \mathcal{A}$ is a continuous function and $U, V : \mathbb{R} \to \mathcal{A}$ are the unique solution to the linear differential equations

$$\dot{V}(t) = A(t) V(t) \text{ with } V(0) = I \tag{2.45}$$

and

$$\dot{U}(t) = -U(t) A(t) \text{ with } U(0) = I. \tag{2.46}$$

Prove that $V(t)$ is invertible and that $V^{-1}(t) = U(t)$, where by abuse of notation I am writing $V^{-1}(t)$ for $[V(t)]^{-1}$. **Hints:** 1) show $\frac{d}{dt}[U(t)V(t)] = 0$ (which is sufficient if $\dim(X) < \infty$) and 2) show $y(t) := V(t)U(t)$ solves a linear differential ordinary differential equation that has $y \equiv I$ as an obvious solution. (The results of Exercise 2.8 may be useful here.) Then use the uniqueness of solutions to linear O.D.E.s

Exercise 2.10. Suppose that $A \in \mathcal{A}$ and $v \in X$ is an eigenvector of A with eigenvalue λ, i.e. that $Av = \lambda v$. Show $e^{tA}v = e^{t\lambda}v$. Also show that if $X = \mathbb{R}^n$ and A is a diagonalizable $n \times n$ matrix with

$$A = SDS^{-1} \text{ with } D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

then $e^{tA} = Se^{tD}S^{-1}$ where $e^{tD} = \operatorname{diag}(e^{t\lambda_1}, \dots, e^{t\lambda_n})$. Here $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$ denotes the diagonal matrix Λ such that $\Lambda_{ii} = \lambda_i$ for $i = 1, 2, \dots, n$.

Exercise 2.11. Suppose that $A, B \in \mathcal{A}$ let $ad_A B = [A, B] := AB - BA$. Show $e^{tA}Be^{-tA} = e^{t ad_A}(B)$. In particular, if $[A, B] = 0$ then $e^{tA}Be^{-tA} = B$ for all $t \in \mathbb{R}$.

Exercise 2.12. Suppose that $A, B \in \mathcal{A}$ and $[A, B] := AB - BA = 0$. Show that $e^{(A+B)} = e^A e^B$.

Exercise 2.13. Suppose $A \in C(\mathbb{R}, \mathcal{A})$ satisfies $[A(t), A(s)] = 0$ for all $s, t \in \mathbb{R}$. Show

$$y(t) := e^{\left(\int_0^t A(\tau)d\tau\right)}x$$

is the unique solution to $\dot{y}(t) = A(t)y(t)$ with $y(0) = x$.

Exercise 2.14. Compute e^{tA} when

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and use the result to prove the formula

$$\cos(s + t) = \cos s \cos t - \sin s \sin t.$$

Hint: Sum the series and use $e^{tA}e^{sA} = e^{(t+s)A}$. **Alternatively, compute** $\frac{d^2}{dt^2}e^{tA} = -e^{tA}$ and then solve this equation.

Exercise 2.15. Compute e^{tA} when

$$A = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$$

with $a, b, c \in \mathbb{R}$. Use your result to compute $e^{t(\lambda I + A)}$ where $\lambda \in \mathbb{R}$ and I is the 3×3 identity matrix. **Hint:** Sum the series.

Exercise 2.16 (L. Gårding's trick I.). Prove Theorem 2.19, i.e. suppose that $T_t \in \mathcal{A}$ for $t \geq 0$ satisfies;

1. (Semi-group property.) $T_0 = Id_X$ and $T_t T_s = T_{t+s}$ for all $s, t \geq 0$.
2. (Norm Continuity at 0+) $t \to T_t$ is continuous at 0, i.e. $\|T_t - I\|_{\mathcal{A}} \to 0$ as $t \downarrow 0$.

Then show there exists $A \in \mathcal{A}$ such that $T_t = e^{tA}$ where e^{tA} is defined in Eq. (2.10). Here is an outline of a possible proof based on L. Gårding's "trick."

1. Using the right continuity at 0 and the semi-group property for T_t, show there are constants M and C such that $\|T_t\|_{\mathcal{A}} \leq MC^t$ for all $t > 0$.
2. Show $t \in [0, \infty) \to T_t \in \mathcal{A}$ is continuous.
3. For $\varepsilon > 0$, let
$$S_\varepsilon := \frac{1}{\varepsilon}\int_0^\varepsilon T_\tau d\tau \in \mathcal{A}.$$
Show $S_\varepsilon \to I$ as $\varepsilon \downarrow 0$ and conclude from this that S_ε is invertible when $\varepsilon > 0$ is sufficiently small. For the remainder of the proof fix such a small $\varepsilon > 0$.
4. Show
$$T_t S_\varepsilon = \frac{1}{\varepsilon}\int_t^{t+\varepsilon} T_\tau d\tau = S_\varepsilon T_t$$
and conclude using the fundamental theorem of calculus that
$$\frac{d}{dt}T_t S_\varepsilon = \frac{1}{\varepsilon}\left[T_{t+\varepsilon} - T_t\right] \text{ for } t > 0 \text{ and}$$
$$\frac{d}{dt}|_{0+}T_t S_\varepsilon := \lim_{t \downarrow 0}\left(\frac{T_t - I}{t}\right)S_\varepsilon = \frac{1}{\varepsilon}\left[T_\varepsilon - I\right].$$
5. Using the fact that S_ε is invertible, conclude $A = \lim_{t\downarrow 0} t^{-1}(T_t - I)$ exists in $\mathcal{A}$ and that
$$A = \frac{1}{\varepsilon}(T_\varepsilon - I)S_\varepsilon^{-1}$$
and moreover,
$$\frac{d}{dt}T_t = AT_t \text{ for } t > 0.$$
6. Using step 5., show $\frac{d}{dt}e^{-tA}T_t = 0$ for all $t > 0$ and therefore $e^{-tA}T_t = e^{-0A}T_0 = I$.

Exercise 2.17 (Duhamel' s Principle). Suppose that $A : \mathbb{R} \to \mathcal{A}$ is a continuous function and $V : \mathbb{R} \to \mathcal{A}$ is the unique solution to the linear differential equation (2.45) which we repeat here;

$$\dot{V}(t) = A(t)V(t) \text{ with } V(0) = I.$$

Let $W_0 \in \mathcal{A}$ and $H \in C(\mathbb{R}, \mathcal{A})$ be given. Show that the unique solution to the differential equation:

$$\dot{W}(t) = A(t) W(t) + H(t) \text{ with } W(0) = W_0 \tag{2.47}$$

is given by

$$W(t) = V(t) W_0 + V(t) \int_0^t V(\tau)^{-1} H(\tau)\, d\tau. \tag{2.48}$$

Hint: compute $\frac{d}{dt}[V^{-1}(t) W(t)]$.

3

Spectrum of a Single Element

Convention. Henceforth all Banach algebras, $\mathcal{B}$, are complex and have an identity.

Definition 3.1. *For $a \in \mathcal{B}$;*

1. *The* ***spectrum*** *of a is*
$$\sigma(a) := \{\lambda \in \mathbb{C} : a - \lambda \text{ is not invertible}\},$$
2. *the* ***resolvent set*** *of a is*
$$\rho(a) := \{\lambda \in \mathbb{C} : a - \lambda \text{ is invertible}\} = \sigma(a)^c,$$
and
3. *the* ***spectral radius*** *of a is*
$$r(a) := \sup\{|\lambda| : \lambda \in \sigma(a)\}.$$

We will see later in Corollary 3.41 that $\sigma(a) \neq \emptyset$.

Proposition 3.2. *For all $a \in \mathcal{B}$, $\sigma(a)$ is compact and $r(a) \leq \|a\|$.*

Proof. Since $\lambda \in \mathbb{C} \to a - \lambda \in \mathcal{B}$ is continuous and $\rho(a) = \{\lambda : a - \lambda \in \mathcal{B}_{inv}\}$, $\rho(a)$ is open by Corollary 2.15 and hence $\sigma(a) = \rho(a)^c$ is closed. If $|\lambda| > \|a\|$, then $\|\lambda^{-1}a\| < 1$ and hence
$$a - \lambda = \lambda\left(\lambda^{-1}a - 1\right) \in \mathcal{B}_{inv}.$$
Therefore if $|\lambda| > \|a\|$ then $\lambda \in \rho(a)$ from which we conclude that $r(a) \leq \|a\|$ and so $\sigma(a)$ is compact. ■

Lemma 3.3. *If $\mathcal{B}$ is a $*$-algebra with unit then*
$$\sigma(a^*) = \overline{\sigma(a)} = \left\{\bar{\lambda} : \lambda \in \sigma(a)\right\}.$$

Proof. The point is that $a \in \mathcal{B}$ is invertible iff a^* is invertible since $[a^*]^{-1} = \left(a^{-1}\right)^*$. Thus $\lambda \in \rho(a)$ iff $a - \lambda 1$ is invertible iff $a^* - \bar{\lambda}1 = (a - \lambda 1)^*$ is invertible iff $\bar{\lambda} \in \rho(a^*)$. ■

Notation 3.4 *If $\mathcal{B}$ is a Banach subalgebra of $\mathcal{A}$ with $\mathbf{1} \in \mathcal{B}$ and a is an element of $\mathcal{B}$, then we let $\sigma_{\mathcal{A}}(a)$ and $\sigma_{\mathcal{B}}(a)$ be the spectrum of a computed in $\mathcal{A}$ and $\mathcal{B}$ respectively.*

Remark 3.5. Continuing the notation above, we always have $\sigma_{\mathcal{B}}(a) \subset \sigma_{\mathcal{A}}(a)$ for all $a \in \mathcal{A}$. Indeed, if $\lambda \notin \sigma_{\mathcal{A}}(a)$, then $a - \lambda$ is invertible in $\mathcal{A}$ and hence also in $\mathcal{B}$, i.e. $\lambda \notin \sigma_{\mathcal{B}}(a)$. See Proposition 3.14 and Theorem 3.12 to see that $\sigma_{\mathcal{B}}(a) \subsetneq \sigma_{\mathcal{A}}(a)$ is possible.

Proposition 3.6. *Let $\mathbf{1} \in \mathcal{A} \subset \mathcal{B}$ be as in Notation 3.4. Then $\sigma_{\mathcal{A}}(a) = \sigma_{\mathcal{B}}(a)$ for all $a \in \mathcal{A}$ iff $\mathcal{A} \cap \mathcal{B}_{inv} = \mathcal{A}_{inv}$ iff $\mathcal{A} \cap \mathcal{B}_{inv} \subset \mathcal{A}_{inv}$. Put another way, $\sigma_{\mathcal{A}}(a) = \sigma_{\mathcal{B}}(a)$ if whenever $a \in \mathcal{A}$ is invertible in $\mathcal{B}$, then a is also invertible in $\mathcal{A}$.*

Proof. Suppose that $\sigma_{\mathcal{A}}(a) = \sigma_{\mathcal{B}}(a)$ for all $a \in \mathcal{A}$. Then if $a \in \mathcal{A} \cap \mathcal{B}_{inv}$, we have $a \notin \sigma_{\mathcal{B}}(a) = \sigma_{\mathcal{A}}(a)$, i.e. $a \in \mathcal{A}_{inv}$ which shows $\mathcal{A} \cap \mathcal{B}_{inv} \subset \mathcal{A}_{inv}$. The opposite inclusion is trivial.

Conversely, suppose that $\mathcal{A} \cap \mathcal{B}_{inv} = \mathcal{A}_{inv}$. Because of Remark 3.5 we must show for any $a \in \mathcal{A}$ that $\sigma_{\mathcal{A}}(a) \subset \sigma_{\mathcal{B}}(a)$. If $\lambda \notin \sigma_{\mathcal{B}}(a)$, then $a - \lambda \in \mathcal{A} \cap \mathcal{B}_{inv} = \mathcal{A}_{inv}$ and hence $\lambda \notin \sigma_{\mathcal{A}}(a)$ and the proof is complete. ■

3.1 Spectrum Examples

Before continuing the formal development it may be useful to consider a few examples and some more properties of the spectrum of elements of a Banach algebra, $\mathcal{B}$.

3.1.1 Finite Dimensional Examples

Exercise 3.1. Let X be a finite set and $\mathcal{B} = \mathbb{C}^X$ denote the functions, $f : X \to \mathbb{C}$. Clearly f is invertible in $\mathcal{B}$ iff $0 \notin f(X)$ in which case $(f)^{-1} = \frac{1}{f}$. Show that $1/f = p(f)$ for some $p \in \mathbb{C}[z]$ and hence $1/f$ is in the subalgebra of $\mathcal{B}$ generated by f and 1. Use this to conclude that $\sigma_{\mathcal{B}}(f) = \sigma_{\mathcal{A}(f,1)}(f) = f(X)$ where $\mathcal{A}(f,1)$ is the algebra generated by f and 1.

Remark 3.7 (Be careful in infinite dimensions). An easy consequence of Exercise 3.1 is that

$$\sigma_{\mathcal{B}}(f) = \sigma_{\mathcal{B}_0}(f) = f(X)$$

where is $\mathcal{B}_0$ is any unital sub-algebra of $\mathcal{B}$ which contains f. This result does not necessarily extrapolate to infinite dimensional settings as demonstrated in Proposition 3.14 below, see also Theorem 3.12 and Remark 3.13.

A similar result holds for finite dimensional matrix algebras as well. In this case we will need to use the following Cayley Hamilton theorem.

Theorem 3.8 (Cayley Hamilton Theorem). *Let B be an $n \times n$ matrix and*

$$p(\lambda) := \det(\lambda I - B) = \sum_{j=0}^{n} p_j \lambda^j$$

be it characteristic polynomial. Then $p(B) = \mathbf{0}$ where $\mathbf{0}$ is the zero $n \times n$ matrix.

Proof. This result is easy to understand if B has a basis $\{v_j\}_{j=1}^n$ of eigenvectors with respective eigenvalues $\{\lambda_j\}_{j=1}^n$. Since $p(\lambda_j) = 0$ for all j it follows that

$$p(B) v_j = p(\lambda_j) v_j = 0 \text{ for all } j$$

which implies $p(B)$ is the zero matrix. For completeness we give a proof of the general case below.

For the general case, let $\operatorname{adj}(M)$ be the classical adjoint of M which is the transpose of the cofactor matrix. This matrix satisfies,

$$\operatorname{adj}(M) M = M \operatorname{adj}(M) = \det(M) I.$$

Taking $M = \lambda I - B$ in this equation shows,

$$(\lambda I - B) \operatorname{adj}(\lambda I - B) = p(\lambda) I = \sum_{j=0}^{n} p_j I \lambda^j.$$

Writing out

$$\operatorname{adj}(\lambda I - B) = \sum_{k=0}^{n-1} \lambda^k C_k \text{ where } C_k \in \mathbb{F}^{n \times n},$$

we have

$$\begin{aligned}
\sum_{j=0}^{n} p_j I \lambda^j &= (\lambda I - B) \sum_{k=0}^{n-1} \lambda^k C_k \\
&= \sum_{k=0}^{n-1} \lambda^{k+1} C_k - \sum_{k=0}^{n-1} \lambda^k B C_k \\
&= \sum_{k=1}^{n} \lambda^k C_{k-1} - \sum_{k=0}^{n-1} \lambda^k B C_k \\
&= \lambda^n C_{n-1} + \sum_{k=1}^{n-1} \lambda^k [C_{k-1} - B C_k] - B C_0.
\end{aligned}$$

Comparing coefficients of λ^j then implies,

$$\begin{aligned}
p_n I &= C_{n-1}, \\
p_k I &= [C_{k-1} - B C_k] \text{ for } 1 \le k \le n-1, \\
p_0 I &= -B C_0
\end{aligned}$$

and hence

$$\begin{aligned}
B^n p_n I &= B^n C_{n-1}, \\
B^k p_k I &= B^k [C_{k-1} - B C_k] \text{ for } 1 \le k \le n-1, \\
p_0 I &= -B C_0.
\end{aligned}$$

Summing these identities then shows,

$$\begin{aligned}
p(B) = p(P) I &= B^n C_{n-1} + \sum_{k=1}^{n-1} B^k [C_{k-1} - B C_k] - B C_0 \\
&= B^n C_{n-1} + \sum_{k=1}^{n-1} B^k C_{k-1} - \sum_{k=1}^{n-1} B^{k+1} C_k - B C_0 \\
&= \sum_{k=1}^{n} B^k C_{k-1} - \sum_{k=0}^{n-1} B^{k+1} C_k = \mathbf{0}.
\end{aligned}$$

■

Lemma 3.9. *Let B be an invertible $n \times n$ matrix, then there exists a degree $n-1$ polynomial, q, such that $B^{-1} = q(B)$. In other words B^{-1} is in the sub-algebra of $\operatorname{End}(\mathbb{C}^n)$ generated by B and I.*

Proof. Let p be the characteristic polynomial of B, i.e.

$$p(\lambda) := \det(\lambda I - B) = \sum_{j=0}^{n} a_j \lambda^j = \lambda r(\lambda) + a_0$$

where $a_n = 1$, $a_0 = (-1)^n \det B$, and

$$r(\lambda) := \sum_{j=1}^{n} a_j \lambda^{j-1}.$$

By the Cayley Hamilton Theorem, which means explicitly that

$$\mathbf{0} = p(B) = Br(B) + a_0 I$$

and so

$$B^{-1} = -\frac{1}{a_0} r(B) = q(B).$$

■

Corollary 3.10. *Let $n \in N$ and suppose that $\mathcal{B}$ is any subalgebra of $B(\mathbb{F}^n)$ which contains I. (As usual $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$.) Then for all $S \in \mathcal{B}$, $\sigma_{\mathcal{B}}(S) = \sigma_{B(\mathbb{F}^n)}(S)$ is the set of eigenvalues of S.*

3.1.2 Function Space and Multiplication Operator Examples

Lemma 3.11. *Let $\mathcal{B} := C(X)$ where X is a compact Hausdorff space. Then $f \in \mathcal{B}_{inv}$ iff $0 \notin \operatorname{Ran}(f) = f(X)$ and in this case $f^{-1} = 1/f \in C^*(f,1)$. Consequently, $\sigma_{\mathcal{B}}(f) = f(X) = \sigma_{C^*(f,1)}(f)$.*

Proof. If $f \in \mathcal{B}_{inv}$ and $g = f^{-1} \in \mathcal{B}$, then $f(x)g(x) = 1$ for all $x \in X$ which implies $f(x) \neq 0$ for all x, i.e. $0 \notin \operatorname{Ran}(f)$. Conversely if $0 \notin \operatorname{Ran}(f)$, then $\varepsilon := \min_{x \in X} |f(x)| > 0$ and hence $1/f \in \mathcal{B}$ from which it follows that $f \in \mathcal{B}_{inv}$. By the Weierstrass approximation theorem, there exists $p_n \in \mathbb{C}[z, \bar{z}]$ such that $p_n(z, \bar{z}) \to \frac{1}{z}$ uniformly on $\varepsilon \le |z| \le \|f\|_u$ and therefore

$$\frac{1}{f} = \|\cdot\|_\infty - \lim_{n\to\infty} p_n\left(f, \bar{f}\right) \implies \frac{1}{f} \in C^*(f,1)$$

■

We now are going to take $X = S = \{z \in \mathbb{C} : |z| = 1\}$ in the next couple of results.

Theorem 3.12. *Let $\mathcal{B} = C\left(S^1; \mathbb{C}\right)$ and $\mathcal{A}$ be the Banach subalgebra (not C^*-subalgebra) generated by $u(z) = z$, i.e.*

$$\mathcal{A} = \overline{\{p(z) : p \in \mathbb{C}[z]\}}^{\mathcal{B}}.$$

Then

$$\mathcal{A} = \left\{ f \in \mathcal{B} : \int_{-\pi}^{\pi} f\left(e^{i\theta}\right) e^{in\theta} d\theta = 0 \text{ for all } n \in \mathbb{N} \right\}. \tag{3.1}$$

Proof. Let $\mathcal{A}_0$ denote the right side of Eq. (3.1). It is clear that if $p(z) = \sum_{k=0}^{n} p_k z^k$ is a polynomial in z, then

$$\int_{-\pi}^{\pi} p\left(e^{i\theta}\right) e^{in\theta} d\theta = \sum_{k=0}^{n} p_k \int_{-\pi}^{\pi} e^{ik\theta} e^{in\theta} d\theta = 0 \text{ for all } n \in \mathbb{N}$$

which shows that $p \in \mathcal{A}_0$. As $\mathcal{A}_0$ is a closed subspace of $\mathcal{B}$ we may conclude that $\mathcal{A} \subset \mathcal{A}_0$.

To prove the reverse inclusion, suppose that $f \in \mathcal{A}_0$ and let

$$p_k := \frac{1}{2\pi} \int_{-\pi}^{\pi} f\left(e^{i\theta}\right) e^{-ik\theta} d\theta \text{ for all } k \in \mathbb{Z}$$

and then, for each $n \in \mathbb{N}_0$, let

$$p_n(z) := \sum_{|k| \le n} p_k z^k = \sum_{k=0}^{n} p_k z^k$$

wherein we have used $p_{-k} = 0$ for all $k \in \mathbb{N}$ because $f \in \mathcal{A}$. By the theory of the Fourier series (using the Féjer kernel[1]) we know that

$$q_N(z) := \frac{1}{N+1} \sum_{n=0}^{N} p_n(z) \to f(z) \text{ uniformly in } z,$$

which shows that $f \in \mathcal{A}$.

Alternatively: we can easily show, for any $0 < r < 1$, that

$$\sum_{k=0}^{\infty} p_k r^k z^k = \lim_{N\to\infty} \sum_{k=0}^{N} p_k r^k z^k$$

is a uniform limit and hence $\sum_{k=0}^{\infty} p_k r^k z^k \in \mathcal{A}$. However it is well know that

$$\sum_{k=0}^{\infty} p_k r^k z^k = \sum_{k=-\infty}^{\infty} p_k r^k z^k = (p_r * f)(z)$$

where p_r is the Poisson kernel.[2] This kernel had the property that $(p_r * f)(z) \to f(z)$ as $r \uparrow 1$, uniformly in z, for any continuous function on S^1. Thus we again find $f \in \mathcal{A}$. Incidentally, this proof shows that every $f \in \mathcal{A}$ is the boundary value of an **analytic** function in $D = D(0,1)$. ■

[1] Google Fejér kernel and find the corresponding Wikipedia site for the required details.

[2] Google Poisson kernel and find the corresponding Wikipedia site for the required details.

Remark 3.13. Notice that $\mathcal{B} = C\left(S^1;\mathbb{C}\right) = C^*(u,1)$ while $\mathcal{A}$ is "holomorphic" subalgebra of $\mathcal{B}$, i.e. is the Banach algebra generated by u.

Proposition 3.14. *Continuing the notation above we have*

$$\sigma_{\mathcal{B}}(u) = S^1 \subsetneq \bar{D} = \sigma_{\mathcal{A}}(u).$$

[See Conway [7], p.p. 205- 207 and in particular Theorem 5.4 for some related general theory. We will come back to this example again in Example 8.18 below.]

Proof. We know that $\sigma_{\mathcal{B}}(u) = u\left(S^1\right) = S^1$ by Lemma 3.11. Let us not work out $\sigma_{\mathcal{A}}(u)$. Since $\|u\| \le 1$, we know that $S^1 = \sigma_{\mathcal{B}}(u) \subset \sigma_{\mathcal{A}}(u) \subset \bar{D}$. So to complete the proof we must show $D \subset \sigma_{\mathcal{A}}(u)$.

Let $\lambda \in D$ and

$$v_\lambda := (u-\lambda)^{-1} = \frac{1}{u-\lambda} \in \mathcal{B}.$$

For sake of contradiction assume that $v_\lambda \in \mathcal{A}$, i.e. there exists polynomials, $\{p_n\}_{n=1}^\infty$ such that

$$p_n(z) \xrightarrow{\text{unif.}} v_\lambda(z) = \frac{1}{z-\lambda} \text{ as } n \to \infty.$$

Under this assumption we find, by basic complex analysis, that

$$2\pi i = \oint_{S^1} \frac{1}{z-\lambda} dz = \lim_{n\to\infty} \oint_{S^1} p_n(z)\, dz = \lim_{n\to\infty} 0 = 0$$

which is a contradiction. Thus we have shown $\nu_\lambda \notin \mathcal{A}$ and hence $\lambda \in \sigma_{\mathcal{A}}(u)$. ■

The following definition is a special case of Definition 1.32 above.

Definition 3.15. *If $q \in L^\infty(\Omega,\mathcal{F},\mu)$, the **essential range** of q is the subset of $\mathbb{C}$ defined by*

$$\mathrm{essran}_\mu(q) = \{w \in \mathbb{C} : \mu(q^{-1}(D(w,\varepsilon))) > 0 \text{ for all } \varepsilon > 0\}.$$

Here, as usual,

$$D(w,\varepsilon) = \{z \in \mathbb{C} : |z-w| < \varepsilon\}$$

for all $w \in \mathbb{C}$ and $\varepsilon > 0$.

Lemma 3.16. *Suppose that $(\Omega,\mathcal{F},\mu)$ is a measure space and $f : \Omega \to \mathbb{C}$ is a measurable map such that $\mu(f=0) = 0$ and $M := \left\|\frac{1}{f}\right\|_\infty < \infty$. Then $\mu(|f| < 1/(2M)) = 0$ and in particular $0 \notin \mathrm{essran}_\mu(f)$.*

Proof. If $M := \left\|\frac{1}{f}\right\|_\infty$ then for every $C > M$, $\mu\left(\left|\frac{1}{f}\right| \ge C\right) = 0$ or equivalently $\mu(|f| \le 1/C) = 0$. ■

Theorem 3.17. *Suppose that $(\Omega,\mathcal{F},\mu)$ is a measure space and $f \in L^\infty(\mu)$. Then*

$$\mathrm{essran}_\mu(f) = \sigma_{L^\infty(\mu)}(f) = \sigma_{C^*(f,1)}(f). \tag{3.2}$$

Proof. We start with the proof of the first equality in Eq. (3.2). If $\lambda \notin \mathrm{essran}_\mu(f)$ iff there exists $\varepsilon > 0$ so that $\mu(\{|f-\lambda| < \varepsilon\}) = 0$. Thus if $\lambda \notin \mathrm{essran}_\mu(f)$, then $\mu\left(\left|\frac{1}{f-\lambda}\right| > \frac{1}{\varepsilon}\right) = 0$ and hence,

$$\left\|\frac{1}{f-\lambda}\right\|_\infty \le \frac{1}{\varepsilon} < \infty$$

which implies $(f-\lambda)^{-1} = \frac{1}{f-\lambda}$ exists in $L^\infty(\mu)$ and so $\lambda \notin \sigma_{L^\infty(\mu)}(f)$.

Conversely, suppose that $\lambda \notin \sigma_{L^\infty(\mu)}(f)$ so that $(f-\lambda)^{-1} = g$ exists in $L^\infty(\mu)$. Then, by definition, we have $g(f-\lambda) = 1$, μ-a.e. and therefore,

$$\frac{1}{f-\lambda} = g \text{ a.e. and } \left\|\frac{1}{f-\lambda}\right\|_\infty = \|g\|_\infty =: M < \infty.$$

By Lemma 3.16, we conclude that $\mu(|f-\lambda| < 1/(2M)) = 0$ and in particular $\lambda \notin \mathrm{essran}_\mu(f)$.

As we automatically know that $\sigma_{L^\infty(\mu)}(f) \subset \sigma_{C^*(f,1)}(f)$ it suffices to show $\sigma_{C^*(f,1)}(f) \subset \sigma_{L^\infty(\mu)}(f)$. So suppose that $\lambda \notin \sigma_{L^\infty(\mu)}(f) = \mathrm{essran}_\mu(f)$ which implies there exists $\varepsilon > 0$ such that $\mu(|f-\lambda| \le \varepsilon) = 0$ and therefore,

$$\varepsilon \le |f-\lambda| \le \|f\|_\infty + |\lambda| =: M \text{ a.e.}$$

Following the proof of Lemma 3.11, there exists $p_n \in \mathbb{C}[z,w]$ such that

$$\lim_{n\to\infty} \left\|p_n\left(f-\lambda, \bar{f}-\bar{\lambda}\right) - \frac{1}{f-\lambda}\right\|_\infty = 0$$

from which it follows that $(f-\lambda)^{-1} \in C^*(f,1)$. This shows $\lambda \notin \sigma_{C^*(f,1)}(f)$ and the proof is complete. ■

Remark 3.18. By Theorem 7.35 or Corollary **??** below or by the spectral theorem, if $\mathcal{B}$ is a unital commutative C^*-subalgebra of $B(H)$, then

$$\sigma_{C^*(T)}(T) = \sigma_{\mathcal{B}}(T) = \sigma_{B(H)}(T)$$

for all $T \in \mathcal{B}$. The real content here is the statement that if $T \in B(H)$ is a normal operator which is invertible, then $T^{-1} \in C^*(I,T)$.

Theorem 3.19. *Let $(\Omega,\mathcal{F},\mu)$ be a measure space with no infinite atoms and $1 \le p < \infty$ and let*

$$\mathcal{B} = \{M_f \in L^p(\mu) : f \in L^\infty(\mu)\} \subset B(L^p(\mu))$$

be the multiplication function subalgebra of $B(L^p(\mu))$. If $M_f \in \mathcal{B}$ is invertible in $\mathcal{B}$ iff it is invertible in $B(L^p(\mu))$. [When $p = 2$, this is a special case of Theorem 7.35 below.]

Proof. Suppose that $T = M_f^{-1}$ exists in $B(L^p(\mu))$. Then for $g \in L^p(\mu)$ we have

$$f \cdot Tg = g = T[fg] \text{ a.e.} \tag{3.3}$$

If $\mu(|f| = 0) > 0$, then (by the no infinite atoms assumption) we may find $A \subset \{|f| = 0\}$ such that $0 < \mu(A) < \infty$. Taking $g = 1_A$ in Eq. (3.3) implies,

$$f \cdot (T1_A) = 1_A \implies 1 = f \cdot (T1_A) = 0 \cdot (T1_A) = 0 \ \mu\text{-a.e. on } A,$$

which is a contradiction. Thus we conclude that in fact $\mu(f = 0) = 0$, and so from Eq. (3.3) it follows that $Tg = \frac{1}{f}g$ a.e. and moreover,

$$\left\|\frac{1}{f}g\right\|_p = \|Tg\|_p \le \|T\|_{op} \|g\|_p \text{ for all } g \in L^p(\mu). \tag{3.4}$$

To finish the proof we need only show $1/f \in L^\infty(\mu)$.

If $0 < M < \infty$ and $\mu(|1/f| \ge M) > 0$, there exists $A \subset \{(|1/f| \ge M)\}$ such that $0 < \mu(A) < \infty$. Then taking $g = 1_A$ in Eq. (3.4) shows,

$$M \|g\|_p \le \left\|\frac{1}{f}g\right\|_p \le \|T\|_{op} \|g\|_p$$

and hence $M \le \|T\|_{op} < \infty$. As this is true for all M such that $\mu(|1/f| \ge M) > 0$, we conclude that $\left\|\frac{1}{f}\right\|_\infty \le \|T\|_{op} < \infty$ and so $T = M_f^{-1} = M_{1/f} \in \mathcal{B}$ and the proof is complete. ∎

Corollary 3.20. *Continuing the notation in Theorem 3.19 with $p = 2$, we have for every $f \in \mathcal{B} = L^\infty(\mu)$ that*

$$\sigma_{B(L^2(\mu))}(M_f) = \sigma_{\mathcal{B}}(M_f) = \sigma_{L^\infty(\mu)}(f) = \sigma_{C^*(f,1)}(f) = \operatorname{essran}_\mu(f).$$

Moreover $C^*(f, 1)$ and $C^*(M_f, 1)$ are isomorphic as C^*-algebras and therefore,

$$\sigma_{C^*(f,1)}(f) = \sigma_{C^*(M_f,1)}(M_f) = \operatorname{essran}_\mu(f).$$

Proof. This is a combination of Theorems 2.58, 3.17 and 3.19. The details are left to the reader. ∎

Example 3.21. Let $\mathbf{q} = (q_1, \ldots, q_n)$ be a vector of bounded measurable functions on some probability space $(\Omega, \mathcal{F}, \mu)$. Let $\mathcal{B}$ be the C^*-algebra generated by $\{1\} \cup \{M_{q_j}\}_{j=1}^n$. Then

$$C(\operatorname{essran}_\mu(\mathbf{q})) \ni f \to M_{f\circ\mathbf{q}} \in \mathcal{B} \subset B(L^2(\mu))$$

is an isometric $*-$isomorphism of Banach algebras. Therefore we conclude and in particular

$$\sigma(M_{f\circ\mathbf{q}}) = f(\operatorname{essran}_\mu(\mathbf{q})).$$

3.1.3 Operators in a Banach Space Examples

For the next couple of definitions and results, let X be a complex Banach space. Recall, by the open mapping theorem, if $T \in B(X)$ is invertible then T^{-1} is bounded, see Lemma 2.9 and Corollary 2.10.

Definition 3.22. *Let X be a complex Banach space and $T \in B(X)$. The set, $\sigma_{ap}(T) \subset \mathbb{C}$, of **approximate eigenvalues** of T is defined by*

$$\sigma_{ap}(T) = \left\{\lambda \in \mathbb{C} : \inf_{\|x\|=1} \|(T - \lambda I)x\| = 0\right\}.$$

*Alternatively stated; $\lambda \in \mathbb{C}$ is $\sigma_{ap}(T)$ iff there exists $\{x_n\}_{n=1}^\infty \subset X$ with $\|x_n\|_X = 1$ such that $\lim_{n\to\infty}(T - \lambda)x_n = 0$. We call such a sequence $\{x_n\}_{n=1}^\infty$ an **approximate eigensequence** for T.*

Proposition 3.23. *If $T \in B(X)$, then $\sigma_{ap}(T)$ is a closed subset of $\sigma(T)$.*

Proof. If $\lambda \notin \sigma(T)$, then $(T - \lambda I)^{-1}$ exists as a bounded operator and therefore with $M := \left\|(T - \lambda I)^{-1}\right\|_{op} < \infty$ we have,

$$\left\|(T - \lambda I)^{-1} x\right\| \le M \|x\| \ \forall x \in X.$$

Replacing x by $(T - \lambda I)x$ in this equation shows,

$$\|(T - \lambda I)x\| \ge \varepsilon \|x\| \ \forall x \in X$$

where $\varepsilon := M^{-1}$. This clearly shows $\lambda \notin \sigma_{ap}(T)$ and hence $\sigma_{ap}(T) \subset \sigma(T)$.

Moreover, if $\lambda \notin \sigma_{ap}(T)$, then there exists $\varepsilon > 0$ so that

$$\|(T - \lambda I)x\| \ge \varepsilon \|x\| \ \forall x \in X.$$

So if $h \in \mathbb{C}$, then

$$\begin{aligned}\|(T - (\lambda + h)I)x\| &= \|(T - \lambda)x - hx\| \ge \|(T - \lambda)x\| - \|hx\| \\ &\ge \varepsilon\|x\| - |h|\|x\| = (\varepsilon - |h|)\|x\|.\end{aligned}$$

Hence we conclude that if $|h| < \varepsilon$, then $(\lambda + h) \notin \sigma_{ap}(T)$ which shows $\mathbb{C}\backslash\sigma_{ap}(T)$ is open and hence $\sigma_{ap}(T)$ is closed. ∎

Example 3.24. Let $D := \{z \in \mathbb{C} : |z| < 1\}$ and $S : \ell^2 \to \ell^2$ be the shift operator, $S(\omega_1, \omega_2, \dots) = (0, \omega_1, \omega_2, \dots)$. Then

$$\sigma_{ap}(S^*) = \sigma(S^*) = \sigma(S) = \bar{D} \text{ and } \sigma_{ap}(S) \subset S^1 \subset \bar{D} = \sigma(S)$$

and hence it can happen that $\sigma_{ap}(S) \subsetneqq \sigma(S)$. [See Exercise 3.2 where you are asked to show $\sigma_{ap}(S) = S^1$.]

Proof. It is easy to see that S is an isometry, the adjoint, S^*, of S is the left shift operator,

$$S^*(\omega_1, \omega_2, \dots) = (\omega_2, \omega_3, \dots),$$

and $\|S\|_{op} = 1 = \|S^*\|_{op}$. Thus we conclude that $\sigma(S) \subset \bar{D}$, and for any $\lambda \in D$,

$$\|(S - \lambda)\psi\| = \|S\psi - \lambda\psi\| \geq |\|S\psi\| - |\lambda|\,\|\psi\|| = (1 - |\lambda|)\,\|\psi\|.$$

The latter inequality shows $\sigma_{ap}(S) \subset \mathbb{C} \setminus D$.

For $\lambda \in D$, $v_\lambda := (1, \lambda, \lambda^2, \dots) \in \ell^2$ and

$$S^* v_\lambda = S^*(1, \lambda, \lambda^2, \dots) = \lambda(1, \lambda, \lambda^2, \dots) = \lambda v_\lambda$$

which shows $D \subset \sigma_{ev}(S^*) \subset \sigma_{ap}(S^*)$. Because $\sigma_{ap}(S^*)$ is closed, $\bar{D} \subset \sigma_{ap}(S^*) \subset \sigma(S^*) \subset \bar{D}$, i.e.

$$\sigma_{ap}(S^*) = \sigma(S^*) = \bar{D} = \sigma(S).$$

Since we have already seen that $\sigma_{ap}(S) \subset \mathbb{C} \setminus D$, it follows that $\sigma_{ap}(S) \subset \bar{D} \setminus D = S^1$.

Remark. We may directly show that $S^1 \subset \sigma_{ap}(S^*)$ as follows. Let $\lambda \in S^1$ and then set $\omega^N := (1, \lambda, \lambda^2, \dots \lambda^N, 0, 0, \dots)$. We then have $\left\|\omega^N\right\|_{\ell^2}^2 = N + 1$ while

$$S^*\omega^N - \lambda\omega^N = \lambda\omega^{N-1} - \lambda\omega^N = -\lambda^{N+1} e_{N+1}$$

and therefore,

$$(S^* - \lambda)\frac{\omega^N}{\sqrt{N+1}} = -\frac{1}{\sqrt{N+1}}\lambda^{N+1} e_{N+1} \to 0 \text{ as } N \to \infty$$

while $\left\|\omega^N/\sqrt{N+1}\right\|_{\ell^2} = 1$.

Exercise 3.2. Continuing then notation used in Example 3.24, show $\sigma_{ap}(S) = S^1$.

Exercise 3.3. Let $H = L^2([0,1], m)$, $g \in L^\infty([0,1])$, and define $T \in B(H)$ by

$$(Tf)(x) = \int_0^x g(y)\, f(y)\, dy.$$

Show;

1. $\sigma(T) = \{0\}$,
2. $\sigma_{ev}(T) \neq \emptyset$ iff $m\{\{g = 0\}\} > 0$.
3. Show $\sigma_{ap}(T) = \{0\}$.

3.1.4 Spectrum of Normal Operators

Exercise 3.4. If T is a subset of H, show $T^{\perp\perp} = \overline{\operatorname{span}(T)}$ where $\operatorname{span}(T)$ denotes all finite linear combinations of elements from T.

Lemma 3.25. *If H and K be Hilbert spaces and $A \in L(H, K)$, then;*

1. $\overline{\operatorname{Nul}(A^*)} = \operatorname{Ran}(A)^\perp$, *and*
2. $\overline{\operatorname{Ran}(A)} = \operatorname{Nul}(A^*)^\perp$,
3. *If we further assume that $K = H$, and $V \subset H$ is an A – invariant subspace (i.e. $A(V) \subset V$), then $V^\perp$ is A^* – invariant.*

Proof. 1. We have $y \in \operatorname{Nul}(A^*) \iff A^*y = 0 \iff \langle y, Ah\rangle = \langle 0, h\rangle = 0$ for all $h \in H \iff y \in \operatorname{Ran}(A)^\perp$.

2. By Exercise 3.4, $\overline{\operatorname{Ran}(A)} = \operatorname{Ran}(A)^{\perp\perp}$, and so $\overline{\operatorname{Ran}(A)} = \operatorname{Ran}(A)^{\perp\perp} = \operatorname{Nul}(A^*)^\perp$.

3. Now suppose that $K = H$ and $AV \subset V$. If $y \in V^\perp$ and $x \in V$, then

$$\langle A^*y, x\rangle = \langle y, Ax\rangle = 0 \text{ for all } x \in V \implies A^*y \in V^\perp.$$

■

For this section we always assume that H is a complex Hilbert space.

Lemma 3.26. *If $C \in B(H)$ and $\langle C\psi, \psi\rangle = 0$ for all $\psi \in H$, then $C = 0$.*

Proof. If $\psi, \varphi \in H$, then

$$\begin{aligned} 0 &= \langle C(\psi + \varphi), \psi + \varphi\rangle \\ &= \langle C\psi, \psi\rangle + \langle C\varphi, \varphi\rangle + \langle C\psi, \varphi\rangle + \langle C\varphi, \psi\rangle \\ &= \langle C\psi, \varphi\rangle + \langle C\varphi, \psi\rangle. \end{aligned}$$

Replacing ψ by $i\psi$ in this identity also shows

$$0 = i\left[\langle C\psi, \varphi\rangle - \langle C\varphi, \psi\rangle\right]$$

which combined with the previous equation easily gives, $\langle C\psi, \varphi\rangle = 0$. Since $\psi, \varphi \in H$ are arbitrary we must have $C \equiv 0$. ■

Lemma 3.27. *If $C \in B(H)$, then;*

1. $C^* = C$ *iff* $\langle C\psi, \psi\rangle \in \mathbb{R}$ *for all* $\psi \in H$ *and*
2. $C^* = -C$ *iff* $\langle C\psi, \psi\rangle \in i\mathbb{R}$ *for all* $\psi \in H$.

Proof. If $C = C^*$, then

$$\overline{\langle C\psi, \psi\rangle} = \langle \psi, C\psi\rangle = \langle C^*\psi, \psi\rangle = \langle C\psi, \psi\rangle$$

which $\langle C\psi, \psi\rangle \in \mathbb{R}$. Conversely if $\langle C\psi, \psi\rangle \in \mathbb{R}$ for all $\psi \in H$ then

$$\langle C\psi, \psi\rangle = \overline{\langle C\psi, \psi\rangle} = \langle \psi, C\psi\rangle = \langle C^*\psi, \psi\rangle$$

from which it follows that $\langle (C - C^*)\,\psi, \psi\rangle = 0$ for all $\psi \in H$. Therefore, by Lemma 3.26, $C - C^* = 0$ which completes the proof of item 1. Item 2. follows from item 1. since, $C^* = -C$ iff $(iC)^* = iC$ iff $\langle iC\psi, \psi\rangle \in \mathbb{R}$ iff $\langle C\psi, \psi\rangle \in i\mathbb{R}$. ■

Definition 3.28 (Normal operators). *An operator $A \in B(H)$ is **normal** iff $[A, A^*] = 0$, i.e. $A^*A = AA^*$.*

Lemma 3.29. *An operator $A \in B(H)$ is normal iff*

$$\|A\psi\| = \|A^*\psi\| \;\; \forall\; \psi \in H. \tag{3.5}$$

Proof. If A is normal and $\psi \in H$, then

$$\|A\psi\|^2 = \langle A^*A\psi, \psi\rangle = \langle AA^*\psi, \psi\rangle = \langle A^*\psi, A^*\psi\rangle = \|A^*\psi\|^2 .$$

Conversely if Eq. (3.5) holds and $C := [A, A^*] = AA^* - A^*A$, then the above computation shows $\langle C\psi, \psi\rangle = 0$ for all $\psi \in H$. Thus by Lemma 3.26, $0 = C = [A, A^*]$, i.e. A is normal. ■

Corollary 3.30. *If $A \in B(H)$ is a normal operator, then $\mathrm{Nul}\,(A) = \mathrm{Nul}\,(A^*)$ and $\sigma_{ev}(A^*) = \mathrm{cong}\,(\sigma_{ev}(A))$ where for any $\Omega \subset \mathbb{C}$,*

$$\mathrm{cong}\,(\Omega) = \left\{\bar{\lambda} \in \mathbb{C} : \lambda \in \Omega\right\}.$$

Proof. If $\lambda \in \mathbb{C}$, then $\mathrm{Nul}\,(A - \lambda) = \mathrm{Nul}\left(A^* - \bar{\lambda}\right)$, i.e. $Au = \lambda u$ iff $A^*u = \bar{\lambda}u$. ■

Lemma 3.31. *If $B, C \in B(H)$ are commuting self-adjoint operators, then*

$$\|(B + iC)\,\psi\|^2 = \|B\psi\|^2 + \|C\psi\|^2 \;\; \forall\; \psi \in H.$$

Proof. Simple manipulations show,

$$\begin{aligned}\|(B + iC)\,\psi\|^2 &= \|B\psi\|^2 + \|C\psi\|^2 + 2\,\mathrm{Re}\,\langle B\psi, iC\psi\rangle \\ &= \|B\psi\|^2 + \|C\psi\|^2 + 2\,\mathrm{Im}\,\langle CB\psi, \psi\rangle \\ &= \|B\psi\|^2 + \|C\psi\|^2\end{aligned}$$

where the last equality follows from Lemma 3.27 because,

$$(CB)^* = B^*C^* = BC = CB.$$

■

Remark 3.32. Here is another way to understand Lemma 3.29. If A is normal then $A = B + iC$ where

$$B = \frac{1}{2}\,(A + A^*) \text{ and } C = \frac{1}{2i}\,(A - A^*)$$

are two commuting self-adjoint operators. Therefore by Lemma 3.31,

$$\|A\psi\|^2 = \frac{1}{4}\left[\|(A + A^*)\,\psi\|^2 + \|(A - A^*)\,\psi\|^2\right]$$

which is symmetric under the interchange of A with A^*.

Remark 3.33. Suppose that a, b are commuting elements of $\mathcal{A}$, then $ab \in \mathcal{A}_{inv}$ iff $a, b \in \mathcal{A}_{inv}$. More generally if $a_i \in \mathcal{A}$ for $i = 1, 2, \dots, n$ are commuting elements then $\prod_{i=1}^n a_i \in \mathcal{A}_{inv}$ iff $a_i \in \mathcal{A}_{inv}$ for all i. To prove this suppose that $c := ab \in \mathcal{A}_{inv}$, then c commutes with both a and b and hence c^{-1} also commutes with a and b. Therefore $1 = \left(c^{-1}a\right)b = b\left(c^{-1}a\right)$ which shows that $b \in \mathcal{A}_{inv}$ and $b^{-1} = c^{-1}a$. Similarly one shows that $a \in \mathcal{A}_{inv}$ as well and $a^{-1} = c^{-1}b$. The more general version is easily proved in the same way or by induction on n.

Lemma 3.34. *Suppose that $A \in B(H)$ is a normal operator, i.e. $[A, A^*] = 0$. Then $\sigma(A) = \sigma_{ap}(A)$ and*

$$\sigma(A) = \left\{\lambda \in \mathbb{C} : 0 \in \sigma\left((A - \lambda)^*(A - \lambda)\right)\right\}. \tag{3.6}$$

[In other words, $(A - \lambda)$ is invertible iff $(A - \lambda)^*(A - \lambda)$ is invertible.]

Proof. By Proposition 3.23, $\sigma_{ap}(A) \subset \sigma(A)$. If $\lambda \notin \sigma_{ap}(A)$, then there exists $\varepsilon > 0$ so that

$$\varepsilon := \inf_{\|\psi\|=1} \|(A - \lambda 1)\psi\| > 0$$

or equivalently

$$\|(A - \lambda 1)\psi\| \geq \varepsilon\,\|\psi\| \;\; \forall\; \psi \in H.$$

As $A - \lambda I$ is normal we also know (see Lemma 3.29) that

$$\|(A - \lambda 1)^*\psi\| = \|(A - \lambda 1)\psi\| \geq \varepsilon\,\|\psi\| \;\; \forall\; \psi \in H$$

and in particular,

$$\mathrm{Nul}\,(A - \lambda I) = \{0\} = \mathrm{Nul}\left((A - \lambda I)^*\right).$$

By Corollary 2.10, $\mathrm{Ran}\,(A - \lambda I)$ is closed. Using these comments along with Lemma 3.25 allows us to conclude,

$$\mathrm{Ran}\,(A - \lambda I) = \overline{\mathrm{Ran}\,(A - \lambda I)} = \mathrm{Nul}\left((A - \lambda I)^*\right)^{\perp} = \{0\}^{\perp} = H$$

and hence $A-\lambda I$ is invertible and therefore $\lambda \notin \sigma(A)$. Thus we have shown $\sigma(A) \subset \sigma_{ap}(A)$ and hence $\sigma_{ap}(A) = \sigma(A)$.

We now prove Eq. (3.6). First note that because A is normal, $A-\lambda$ is normal and also $A-\lambda$ is invertible iff $(A-\lambda)^*$ is invertible. Therefore by Remark 3.33, $(A-\lambda)^*(A-\lambda)$ iff both $(A-\lambda)^*$ and $(A-\lambda)$ are invertible iff $A-\lambda$ is invertible. This is the contrapositive of Eq. (3.6). ∎

Example 3.35. Let S be the shift operator as in Example 3.24. Then $S^*S = I$ while $SS^* \neq I$ since

$$SS^*(\omega_1, \omega_2, \omega_3, \dots) = (0, \omega_2, \omega_3, \dots).$$

Thus S is not normal and by Example 3.24, $\sigma_{ap}(S) \subsetneq \sigma(S)$. Moreover, S^*S is invertible even though neither S nor S^* is invertible, i.e. $0 \in \sigma(S)$ while $0 \notin \sigma(S^*S)$. This example shows that we can not drop the assumption that $[a,b]=0$ in Remark 3.33.

Lemma 3.36. *If $A \in B(H)$ is self-adjoint (i.e. $A = A^*$), then $\sigma(A) \subset \mathbb{R}$. This is generalized in Lemma 4.5.*

Proof. Let $\lambda = \alpha + i\beta$ with $\alpha, \beta \in \mathbb{R}$, then

$$\begin{aligned}\|(A+\alpha+i\beta)\psi\|^2 &= \|(A+\alpha)\psi\|^2 + |\beta|^2\|\psi\|^2 + 2\operatorname{Re}\langle (A+\alpha)\psi, i\beta\psi\rangle \\ &= \|(A+\alpha)\psi\|^2 + |\beta|^2\|\psi\|^2 \geq |\beta|^2\|\psi\|^2 \end{aligned} \tag{3.7}$$

wherein we have used Lemma 3.27 to conclude, $\operatorname{Re}\langle (A+\alpha)\psi, i\beta\psi\rangle = 0$. [Equation (3.7) is a simply a special case of Lemma 3.31.] Equation (3.7) along with Lemma 3.34 shows that $\lambda \notin \sigma(A)$ if $\beta \neq 0$, i.e. $\sigma(A) \subset \mathbb{R}$. ∎

Remark 3.37. It is not true that $\sigma(A) \subset \mathbb{R}$ implies $A = A^*$. For example, let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ on $\mathbb{C}^2$, then $\sigma(A) = \{0\}$ yet $A \neq A^*$. This result is true if we require A to be normal.

3.2 Basic Properties of $\sigma(a)$

Definition 3.38. *The **resolvent** (operators) of a is the function,*

$$\rho(a) \ni \lambda \to R_\lambda = (a-\lambda)^{-1} \in \mathcal{A}_{inv}.$$

Lemma 3.39 (Resolvent Identity). *If $a \in \mathcal{A}$ and $\mu, \lambda \in \rho(a)$, then*

$$R_\lambda - R_\mu = (\lambda - \mu) R_\lambda R_\mu \tag{3.8}$$

and in particular by interchanging the roles of μ and λ it follows that $[R_\lambda, R_\mu] = 0$.

Proof. Apply Eq. (2.7) with $b = (a-\lambda)$ and $c = (a-\mu)$ to find

$$R_\lambda - R_\mu = R_\lambda[(a-\mu)-(a-\lambda)]R_\mu = R_\lambda(\lambda-\mu)R_\mu = (\lambda-\mu)R_\lambda R_\mu.$$

∎

Equation (3.8) is easily remembered by the following heuristic;

$$R_\lambda - R_\mu = \frac{1}{a-\lambda} - \frac{1}{a-\mu} = \frac{(a-\mu)-(a-\lambda)}{(a-\lambda)(a-\mu)} = (\lambda-\mu)R_\lambda R_\mu.$$

Corollary 3.40. *Let $\mathcal{A}$ be a complex Banach algebra with identity and let $a \in \mathcal{A}$. Then the function, $\rho(a) \ni \lambda \to R_\lambda \in \mathcal{A}$ is analytic with $\frac{d}{d\lambda}R_\lambda = R_\lambda^2$ and $\|R_\lambda\| \to 0$ as $\lambda \to \infty$.*

Proof. For $h \in \mathbb{C}$ small,

$$R_{\lambda+h} - R_\lambda = (\lambda + h - \lambda) R_{\lambda+h}R_\lambda = hR_{\lambda+h}R_\lambda$$

and therefore,

$$\frac{1}{h}(R_{\lambda+h} - R_\lambda) = R_{\lambda+h}R_\lambda \to R_\lambda^2 \text{ as } h \to 0$$

wherein we have used Corollary 2.16 in order to see that $R_{\lambda+h} \to R_\lambda$ as $h \to 0$. Since

$$R_\lambda = (a-\lambda)^{-1} = -\lambda^{-1}\left(1-\lambda^{-1}a\right)^{-1},$$

if $|\lambda| > \|a\|$ (i.e. $\left\|\lambda^{-1}a\right\| < 1$) it follows that

$$\|R_\lambda\| = \frac{1}{|\lambda|}\left\|\left(1-\lambda^{-1}a\right)^{-1}\right\| \leq \frac{1}{|\lambda|}\frac{1}{1-\|\lambda^{-1}a\|} = O\left(\frac{1}{|\lambda|}\right) \to 0 \text{ as } |\lambda| \to \infty.$$

∎

Corollary 3.41. *Let $\mathcal{A}$ be a complex Banach algebra with unit, $\mathbf{1} \neq 0$ (as we have assumed that $\|\mathbf{1}\| = 1$.) Then $\sigma(a) \neq \emptyset$ for every $a \in \mathcal{A}$.*

Proof. If $\sigma(a) = \emptyset$, then $R_\lambda = (a-\lambda)^{-1}$ is analytic on all of $\mathbb{C}$ and moreover $\|R_\lambda\| = O\left(\frac{1}{|\lambda|}\right)$ as $\lambda \to \infty$. Therefore by Liouville's theorem (Corollary 1.12), R_λ is constant and in fact must be 0 by letting $\lambda \to \infty$. Therefore

$$\mathbf{1} = R_\lambda(a-\lambda) = 0(a-\lambda) = 0$$

which is a contradiction and therefore $\sigma(a) \neq \emptyset$.

Remark, if we only want to use the classical Liouville's theorem, just apply it to $\lambda \to \xi(R_\lambda)$ for all $\xi \in \mathcal{A}^*$ to find $\xi(R_\lambda) = \xi(R_0)$. As this holds for all $\xi \in \mathcal{A}^*$ it follows again that $R_\lambda = R_0$.

■

Theorem 3.42 (Spectral Mapping Theorem). *If $p : \mathbb{C} \to \mathbb{C}$ is a polynomial and $a \in \mathcal{A}$ then $p(\sigma(a)) = \sigma(p(a))$.*

Proof. Let p be a non-constant polynomial (otherwise there is nothing to prove) and let $\mu \in \mathbb{C}$ be given. Then factor $p(\lambda) - \mu$ as

$$p(\lambda) - \mu = \alpha(\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$$

where $\alpha \in \mathbb{C}^\times$ and $\{\lambda_i\}_{i=1}^n \subset \mathbb{C}$ are the solutions (with multiplicity) to $p(\lambda) = \mu$. Since

$$p(a) - \mu = \alpha(a - \lambda_1)\cdots(a - \lambda_n)$$

we may conclude using Remark 3.33 that $\mu \in \sigma(p(a))$ iff $\lambda_i \in \sigma(a)$ for some i, i.e. iff $\mu = p(\lambda)$ for some $\lambda \in \sigma(a)$, i.e. iff $\mu \in p(\sigma(a))$.

■

Corollary 3.43. *If $p \in \mathbb{C}[z]$ and $a \in \mathcal{A}$, then*

$$r(p(a)) = \sup_{\lambda \in \sigma(a)} |p(\lambda)| = \|p\|_{\infty,\sigma(a)} \tag{3.9}$$

and in particular, $r(a^n) = r(a)^n$ for all $n \in \mathbb{N}$.

Proof. Using Theorem 3.42 and the definition of r,

$$r(p(a)) = \sup\{|z| : z \in \sigma(p(a))\} = \sup\{|p(\lambda)| : \lambda \in \sigma(a)\}$$

which proves Eq. (3.9). Taking $p(z) = z^n$ in this equation shows,

$$r(a^n) = \sup\{|\lambda|^n : \lambda \in \sigma(a)\} = [\sup\{|\lambda| : \lambda \in \sigma(a)\}]^n = r(a)^n.$$

■

Corollary 3.44. *The function, $\lambda \to (1 - \lambda a)^{-1}$, is analytic on $|\lambda| < 1/r(a)$ and moreover admits the power series representation,*

$$(1 - \lambda a)^{-1} = \sum_{n=0}^{\infty} \lambda^n a^n \tag{3.10}$$

which is valid for $|\lambda| < 1/r(a)$.

Proof. If $|\lambda| \|a\| = \|\lambda a\| < 1$, we know that Eq. (3.10) is valid and hence $(1 - \lambda a)^{-1}$ is analytic near 0 as well, see Remark 1.11. [Alternatively we may compute by the chain rule that

$$\frac{d}{d\lambda}(1 - \lambda a)^{-1} = (1 - \lambda a)^{-1} a (1 - \lambda a)^{-1}.]$$

For $\lambda \neq 0$,

$$(1 - \lambda a)^{-1} = \lambda^{-1}\left(\frac{1}{\lambda} - a\right)^{-1} = \lambda^{-1} R_{\lambda^{-1}}$$

which is valid provided $1/\lambda \in \rho(a)$ which will hold if $\frac{1}{|\lambda|} > r(a)$, i.e. if $0 < |\lambda| < 1/r(a)$. So we have shown $(1 - \lambda a)^{-1}$ is analytic near 0 and also, by Corollary 3.40, for $0 < |\lambda| < 1/r(a)$. Thus it follows that $(1 - \lambda a)^{-1}$ is analytic on for $|\lambda| < 1/r(a)$ and hence by Theorem 1.10, the expansion in Eq. (3.10) is valid for $|\lambda| < 1/r(a)$.

■

Corollary 3.45. *The spectral radius $r(a)$ may be computed by taking the following limit,*

$$r(a) = \lim_{n\to\infty} \|a^n\|^{1/n}.$$

Proof. By Corollary 3.43,

$$r(a)^n = r(a^n) \leq \|a^n\| \implies r(a) \leq \|a^n\|^{1/n}.$$

Passing to the limit as $n \to \infty$ in this inequality shows

$$r(a) \leq \liminf_{n\to\infty} \|a^n\|^{1/n}. \tag{3.11}$$

For the opposite we conclude from Eq. (3.10) that $\lim_{n\to\infty} \|(\lambda a)^n\| = 0$ when $|\lambda| < 1/r(a)$. This assertion then implies,

$$|\lambda| \limsup_{n\to\infty} \|a^n\|^{1/n} = \limsup_{n\to\infty} \|(\lambda a)^n\|^{1/n} \leq 1 \ \forall \ |\lambda| < 1/r(a)$$

and hence $\limsup_{n\to\infty} \|a^n\|^{1/n} \leq r(a)$ which along with Eq. (3.11) completes the proof.

■

Exercise 3.5 (Compare with Proposition 8.3). Let $\mathcal{B}$ be a complex Banach algebra with unit, then for any $a, b \in \mathcal{B}$ which commute, show;

1. $r(ab) \leq r(a) r(b)$ and
2. $r(a + b) \leq r(a) + r(b)$.

Proposition 3.46 (Optional). *If $a \in \mathcal{A}$ and $\lambda \in \rho(a)$, then*

$$\left\|(a - \lambda)^{-1}\right\| \geq r\left((a - \lambda)^{-1}\right) \geq \frac{1}{\operatorname{dist}(\lambda, \sigma(a))}.$$

Proof. If $\lambda \in \rho(a)$ and $\beta \in \mathbb{C}$, then

$$(a - (\lambda + \beta)) = (a - \lambda) - \beta = (a - \lambda)\left[I - \beta(a - \lambda)^{-1}\right]$$

is invertible if

$$\sum_{n=1}^{\infty} \left\| \left[\beta (a - \lambda)^{-1}\right]^n \right\| < \infty.$$

The latter condition is implied by requiring $\limsup_{n\to\infty} \left\| \left[\beta (a - \lambda)^{-1}\right]^n \right\|^{1/n} < 1$, i.e.

$$\begin{aligned} & |\beta| \limsup_{n\to\infty} \left\| \left[(a - \lambda)^{-1}\right]^n \right\|^{1/n} < 1 \\ \iff & |\beta| < \limsup_{n\to\infty} \left\| \left[(a - \lambda)^{-1}\right]^n \right\|^{-1/n} = \frac{1}{r\left((a - \lambda)^{-1}\right)} \end{aligned}$$

and hence

$$\operatorname{dist}(\lambda, \sigma(a)) \geq \frac{1}{r\left((a - \lambda)^{-1}\right)} \iff r\left((a - \lambda)^{-1}\right) \geq \frac{1}{\operatorname{dist}(\lambda, \sigma(a))}.$$

■

Holomorphic and Continuous Functional Calculus

In this chapter we wish to consider two methods for defining functions of a given element of a Banach algebra, $\mathcal{B}$. The first method allows us to define $f(a)$ for almost any $a \in \mathcal{B}$ provided that f is analytic on an open neighborhood of the spectrum of a. Later we will specialize to the case where $\mathcal{B}$ is a C^*-algebra and $a \in \mathcal{B}$ is Hermitian. In this case we will make sense of $f(a)$ for any bounded measurable function, $f : \sigma(a) \to \mathbb{C}$.

4.1 Holomorphic (Riesz) Functional Calculus

The material in this section was probably taken from M. Taylor [49, pages 576-578]. Let $\mathcal{B}$ be a unital Banach algebra and $a \in \mathcal{B}$. Suppose that $\sigma(a)$ is a disjoint union of sets $\{\Sigma_k\}_{k=1}^n$ which are surrounded by contours $\{C_k\}_{k=1}^n$ and Ω is an open subset of $\mathbb{C}$ which contains the contours and their interiors, see Figure 4.1.

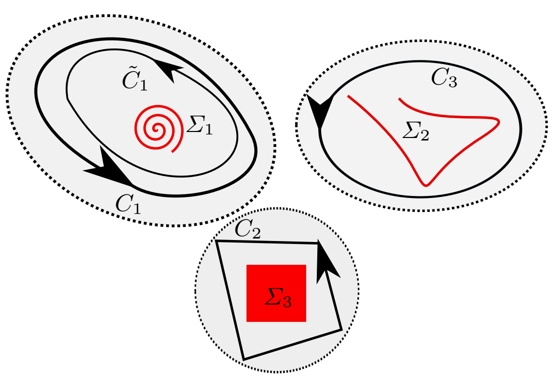

Fig. 4.1. The spectrum of a is in red, the counter clockwise contours are in black, and Ω is the union of the grey sets.

Given a holomorphic function, f, on Ω we let

$$f(a) := \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-a}dz := \sum_{k=1}^n \frac{1}{2\pi i}\oint_{C_k} \frac{f(z)}{z-a}dz,$$

where $\frac{1}{z-a} := (z-a)^{-1}$ and $C = \cup_{k=1}^n C_k$.

Let us observe that $f(a)$ is independent of the possible choices of contours C as described above. One way to prove this is to choose $\ell \in B(X)^*$ and notice that

$$\ell(f(a)) = \frac{1}{2\pi i}\oint_C f(z)\,\ell\left((z-a)^{-1}\right)dz$$

where $f(z)\,\ell\left((z-a)^{-1}\right)$ is a holomorphic function on $\Omega \setminus \sigma(a)$. Therefore $\frac{1}{2\pi i}\oint_C f(z)\,\ell\left((z-a)^{-1}\right)dz$ remains constant over deformations of C which remain in $\Omega \setminus \sigma(a)$. As ℓ is arbitrary it follows that $\frac{1}{2\pi i}\oint_C \frac{f(z)}{z-a}dz$ remains constant over such deformations as well.

Theorem 4.1. *The map $H(\Omega) \ni f \to f(a) \in \mathcal{B}$ is an algebra homomorphism satisfying the consistency criteria; if $f(z) = \sum_{m=0}^N c_m z^m$ is a polynomial then*

$$f(a) = \sum_{m=0}^N c_m a^m.$$

More generally, $\rho > 0$ is chosen so that $r(a) < \rho$ and $f \in H(D(0,\rho))$, then

$$f(a) = \sum_{m=0}^\infty \frac{f^m(0)}{m!}a^m. \tag{4.1}$$

Proof. It is clear that $H(\Omega) \ni f \to f(a) \in \mathcal{B}$ is linear in f. Now suppose that $f, g \in H(\Omega)$ and for each k let $\tilde{C}_k$ be another contour around Σ_k which is inside C_k for each k. Then

$$\begin{aligned} f(a)\,g(a) &= \left(\frac{1}{2\pi i}\right)^2 \sum_{k,l=1}^n \oint_{C_k} \frac{f(z)}{z-a}dz \oint_{\tilde{C}_l} \frac{g(\zeta)}{\zeta-a}d\zeta \\ &= \left(\frac{1}{2\pi i}\right)^2 \sum_{k,l=1}^n \oint_{C_k} dz \oint_{\tilde{C}_l} d\zeta \frac{f(z)}{z-a}\frac{g(\zeta)}{\zeta-a}. \\ &=: \left(\frac{1}{2\pi i}\right)^2 \sum_{k,l=1}^n A_{kl} \end{aligned}$$

Using the resolvent formula,

$$\frac{1}{z-a} - \frac{1}{\zeta - a} = \frac{\zeta - z}{(z-a)(\zeta - a)},$$

we find (using Fubini-Tonelli) that

$$\begin{aligned} A_{kl} &:= \oint_{C_k} dz \oint_{\tilde{C}_l} d\zeta \frac{f(z)}{z-a}\frac{g(\zeta)}{\zeta - a} \\ &= \oint_{C_k} dz \oint_{\tilde{C}_l} d\zeta f(z)\, g(\zeta)\, \frac{1}{\zeta - z}\left(\frac{1}{z-a} - \frac{1}{\zeta - a}\right) \\ &= \oint_{C_k} dz \frac{f(z)}{z-a} \oint_{\tilde{C}_l} d\zeta \frac{g(\zeta)}{\zeta - z} \\ &\quad - \oint_{\tilde{C}_l} d\zeta \frac{g(\zeta)}{\zeta - a} \oint_{C_k} dz \frac{f(z)}{\zeta - z}. \end{aligned} \tag{4.2}$$

For $z \in C_k$, $\zeta \to g(\zeta)\frac{1}{\zeta - z}$ is analytic for ζ inside $\tilde{C}_l$ no matter the l and therefore,

$$\oint_{\tilde{C}_l} d\zeta g(\zeta)\, \frac{1}{\zeta - z} = 0 \tag{4.3}$$

and Eq. (4.2) simplifies to

$$A_{kl} = -\oint_{\tilde{C}_l} d\zeta \frac{g(\zeta)}{\zeta - a}\oint_{C_k} dz \frac{f(z)}{\zeta - z}.$$

If $k \neq l$, we still have $z \to \frac{f(z)}{\zeta - z}$ is analytic inside of C_k and for each $\zeta \in \tilde{C}_l$ and so

$$\oint_{C_k} dz \frac{f(z)}{\zeta - z} = 0$$

which implies $A_{kl} = 0$. On the other hand when $k = l$

$$\oint_{C_k} dz \frac{f(z)}{\zeta - z} = -2\pi i f(\zeta) \text{ for all } \zeta \in \tilde{C}_k.$$

Hence we have shown,

$$A_{k,k} = 2\pi i \cdot \oint_{\tilde{C}_l} d\zeta \frac{g(\zeta)\, f(\zeta)}{\zeta - a}$$

and therefore,

$$\begin{aligned} f(a)\, g(a) &= \left(\frac{1}{2\pi i}\right)^2 \sum_{k=1}^{n} A_{k,k} \\ &= \sum_{k=1}^{n} \frac{1}{2\pi i} \oint_{\tilde{C}_l} d\zeta \frac{g(\zeta)\, f(\zeta)}{\zeta - a} = (f \cdot g)(a) \end{aligned}$$

which shows that $a \to f(a)$ is an algebra homomorphism.

If $f \in H(D(0,\rho))$, then for every $0 < r < \rho$, there exists $C(r) < \infty$ such that $\left|\frac{f^m(0)}{m!}\right| r^m \leq C(r)$. Therefore choosing $r(a) < r < \rho$, we have

$$\left\| \frac{f^m(0)}{m!} a^m \right\| \leq C(r)\, \frac{\|a^m\|}{r^m}$$

and hence

$$\limsup_{m\to\infty} \left\| \frac{f^m(0)}{m!} a^m \right\|^{1/m} \leq \limsup_{m\to\infty} \left[C(r)^{1/m} \left(\frac{\|a^m\|}{r^m}\right)^{1/m} \right] = r(a)/r < 1.$$

It now follows by the root test that the sum in Eq. (4.1) is absolutely convergent. [Technically we could skip this convergence argument but it is nice to verify directly that the sum is convergent.]

We now verify the equality in Eq. (4.1). Suppose that $f \in H(D(0,\rho))$ where $\rho > r(a)$. From Corollary 3.44, we know that

$$\frac{1}{1-\lambda a} = \sum_{n=0}^{\infty} \lambda^n a^n \text{ is convergent for } |\lambda| < \frac{1}{r(a)}$$

and therefore

$$\frac{1}{z-a} = \frac{1}{z}\frac{1}{1-a/z} = \frac{1}{z}\sum_{n=0}^{\infty}\left(\frac{a}{z}\right)^n = \sum_{n=0}^{\infty} a^n z^{-(n+1)} \text{ for } |z| > r(a).$$

Let $r \in (r(a), \rho)$ as above and let C be the contour, $z = re^{i\theta}$ with $-\pi \leq \theta \leq \pi$. Then

$$f(a) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-a} dz = \frac{1}{2\pi i}\oint_C \sum_{n=0}^{\infty} a^n z^{-(n+1)} f(z)\, dz = \sum_{n=0}^{\infty} c_n a^n$$

where, by the residue theorem or by differentiating the Cauchy integral formula,

$$c_n = \frac{1}{2\pi i}\oint_C z^{-(n+1)} f(z)\, dz = \frac{f^{(n)}(0)}{n!}.$$

∎

Theorem 4.2 (Spectral Mapping Theorem). *Keeping the same notation as above,* $f(\sigma(a)) = \sigma(f(a))$.

Proof. Suppose that $\mu \in \sigma(a)$ and define

$$g(z) := \begin{cases} \frac{f(z)-f(\mu)}{z-\mu} & \text{if } z \neq \mu \\ f'(\mu) & \text{if } z = \mu \end{cases}$$

so that $g \in H(U)$ and $f(z) - f(\mu) = (z-\mu) g(z)$. Therefore $f(a) - f(\mu) = (a-\mu) g(a)$ and so if $f(\mu) \notin \sigma(f(a))$ then $f(a) - f(\mu)$ is invertible and therefore $a - \mu$ would be invertible contradicting $\mu \in \sigma(a)$. Thus we have shown $f(\sigma(a)) \subset \sigma(f(a))$. Conversely if $\alpha \notin f(\sigma(a))$ then $g(z) := \frac{1}{f(z)-\alpha}$ is holomorphic on a neighborhood of $\sigma(a)$. Since $(f(z) - \alpha) g(z) = 1$ it follows that $(f(a) - \alpha) g(a) = I$ and therefore $\alpha \notin \sigma(f(a))$ and we have shown $[f(\sigma(a))]^c \subset [\sigma(f(a))]^c$, i.e. $\sigma(f(a)) \subset \sigma(f(a))$. ■

Exercise 4.1. Continue the notation used in Theorem 4.1 but now assume that $\mathcal{B}$ is a C^*-algebra or is at least equipped with a continuous involution, $*$. Show $f(a)^* = f^*(a^*)$ where where $f^*(z) := \overline{f(\bar{z})}$ is holomorphic on

$$\operatorname{cong}(\Omega) = \left\{\bar{\lambda} \in \mathbb{C} : \lambda \in \Omega\right\}.$$

Recall that $\sigma(a^*) = \operatorname{cong}(\sigma(a)) = \left\{\bar{\lambda} \in \mathbb{C} : \lambda \in \sigma(a)\right\} \subset \operatorname{cong}(\Omega)$ so that g is holomorphic on a neighborhood of $\sigma(a^*)$.

See the Section **??** (**yet to be written**) on perturbation theory for applications of this formalism.

4.2 Hermitian Continuous Functional Calculus

For the remainder of this chapter let $\mathcal{B}$ be a unital C^*-algebra.

Proposition 4.3. *If $\mathcal{B}$ is a C^*-algebra with unit, then $r(a) = \|a\|$ whenever $a \in \mathcal{B}$ is normal, i.e. $[a, a^*] = 0$. [We will give another proof of this result in Lemmas **??** and 8.11 below that $r(a) = \|a\|$ when a is any normal element of $\mathcal{B}$.]*

Proof. We start by showing, for $a \in \mathcal{B}$ which is normal and $n \in \mathbb{N}$, that

$$\left\|a^{2^n}\right\| = \|a\|^{2^n}. \tag{4.4}$$

We will prove Eq. (4.4) by induction on $n \in \mathbb{N}$. By Lemma 2.64, we know that $\|b^2\| = \|b\|^2$ whenever $b \in \mathcal{B}$ is normal. Taking $b = a$ gives Eq. (4.4) for $n = 1$ and then applying the identity with $b = a^{2^n}$ while using the induction hypothesis shows,

$$\left\|a^{2^{n+1}}\right\| = \left\|a^{2^n}\right\|^2 = \left(\|a\|^{2^n}\right)^2 = \|a\|^{2^n+1} \text{ for } n \in \mathbb{N}.$$

The statement that $r(a) = \|a\|$ now follows from Eq. (4.4) and Corollary 3.45 which allows us to compute $r(a)$ as

$$r(a) = \lim_{n\to\infty} \left\|a^{2^n}\right\|^{1/2^n} = \lim_{n\to\infty} \|a\| = \|a\|.$$

■

Example 4.4. Let N be an $n \times n$ complex matrix such that $N_{ij} = 0$ if $i \leq j$, i.e. N is upper triangular with zeros along the diagonal. Then $\sigma(N) = \{0\}$ while $\|N\| \neq 0$. Thus $r(N) = 0 < \|N\|$. On the other hand, $N^n = 0$ so $\lim_{n\to\infty} \|N^n\|^{1/n} = 0 = r(N)$.

Lemma 4.5 (Reality). *Let $\mathcal{B}$ be a unital C^*-algebra. If $a \in \mathcal{B}$ is Hermitian, then $\sigma(a) \subset \mathbb{R}$. [This generalizes Lemma 3.36 above. Also see Lemma 8.12 below for related results.]*

Proof. We must show $a - \lambda \in \mathcal{B}_{inv}$ whenever $\operatorname{Im}\lambda \neq 0$. We first consider $\lambda = i$. For sake of contradiction, suppose that $i \in \sigma(a)$.Then by the spectral mapping Theorem 3.42[1] with $p(z) = \lambda - iz$ implies

$$\lambda + 1 = p(i) \in \sigma(p(a)) = \sigma(\lambda - ia) \text{ for all } \lambda \in \mathbb{R}.$$

Therefore using the fact that $r(x) \leq \|x\|$ for all $x \in \mathcal{B}$ along with the C^*-identity shows,

$$(\lambda + 1)^2 \leq [r(\lambda - ia)]^2 \leq \|\lambda - ia\|^2$$

wherein

$$\begin{aligned} \|\lambda - ia\|^2 &\overset{C^*\text{-cond}}{=} \left\|(\lambda - ia)^* (\lambda - ia)\right\| = \|(\lambda + ia)(\lambda - ia)\| \\ &= \left\|\lambda^2 + a^2\right\| \leq \lambda^2 + \left\|a^2\right\| \overset{C^*\text{-cond}}{=} \lambda^2 + \|a\|^2. \end{aligned}$$

Combining the last two displayed equation leads to the nonsensical inequality, $2\lambda + 1 \leq \|a\|^2$ for all $\lambda \in \mathbb{R}$, and we have arrived at the desired contradiction and hence $i \notin \sigma(a)$.

For general $\lambda = x + iy$ with $y \neq 0$, we have then

$$a - \lambda = a - x - iy = y\left[y^{-1}(a - x) - i\right]$$

which is invertible by step 1. with a replaced by $y^{-1}(a - x)$ which shows $\lambda \notin \sigma(a)$. As this was valid for all λ with $\operatorname{Im}\lambda \neq 0$, we have shown $\sigma(a) \subset \mathbb{R}$. ■

[1] More directly,

$$\lambda + 1 - (\lambda - ia) = 1 + ia = i(a - i)$$

is not invertible by assumption and hence $\lambda + 1 \in (\lambda - ia)$.

Corollary 4.6. *If $a \in \mathcal{B}$ is a Hermitian element of a unital C^*-algebra, then*

$$\|p(a)\| = \sup_{x \in \sigma(a)} |p(x)| \ \forall\ p \in \mathbb{C}[x].$$

Proof. Since $p(a)$ is normal, it follows that $\|p(a)\| = r(p(a))$ which by the spectral mapping theorem may be computed as,

$$r(p(a)) = \max_{\lambda \in \sigma(p(a))} |\lambda| = \max_{\lambda \in \sigma(a)} |p(x)|.$$

■

Theorem 4.7 (Continuous Functional Calculus). *If $a \in \mathcal{B}$ is a Hermitian element of a unital C^*-algebra, then there exists a unique C^*-algebra isomorphism, $\varphi_a : C(\sigma(a)) \to C^*(a,1)$ such that $\varphi_a(x) = a$ or equivalently, $\varphi_a(p) = p(a)$ for all $p \in \mathbb{C}[x]$. [We usually write $\varphi_a(f)$ as $f(a)$.] Let us note that for general $f \in C(\sigma(a); \mathbb{C})$,*

$$f(a) = \varphi_a(f) = \lim_{n\to\infty} p_n(a)$$

where$\{p_n\}_{n=1}^\infty \subset \mathbb{C}[x]$ are any sequence of polynomials such that $p_n|_{\sigma(a)} \to f$ uniformly on $\sigma(a)$. Moreover,[2]

$$\sigma_{C^*(a,1)}(\varphi_a(f)) = f(\sigma(a)).$$

Proof. By the classical Stone–Weierstrass theorem, $\{p|_{\sigma(a)} : p \in \mathbb{C}[x]\}$ is dense in $C(\sigma(a))$ and so because of Corollary 4.6, there exists a unique linear map, $\varphi_a : C(\sigma(a)) \to C^*(a,1)$, such that $\varphi_a(p) = p(a)$ for all $p \in \mathbb{C}[x]$ and $\|\varphi_a(f)\| = \|f\|_{\ell^\infty(\sigma(a))}$. It is now easily verified that φ_a is a homomorphism with dense closed range and hence φ_a is an isomorphism. Moreover, using $p(a)^* = \bar{p}(a)$ we easily conclude by a simple limiting argument that $\varphi_a(\bar{f}) = \varphi_a(f)^*$. For the last assertion, as φ_a is a $*$-homomorphism, it follows that

$$\sigma_{C^*(a,1)}(\varphi_a(f)) = \sigma_{C(\sigma(a))}(f) = f(\sigma(a)).$$

■

Corollary 4.8 (Square Roots). *If $a \in \mathcal{B}$ is a Hermitian element of a unital C^*-algebra and $\sigma(a) \subset [0,\infty)$, then there exists a Hermitian element $b \in \mathcal{B}$ such that $\sigma(b) \subset [0,\infty)$ and $a = b^2$. Moreover, if $c \in \mathcal{B}$ is Hermitian and $c^2 = a$, then $b = |c|$. [See Corollary 9.10 for the polar decomposition.]*

[2] We will see later that in Corollary 7.34 below that $\sigma_{\mathcal{B}}(f(a)) = \sigma_{C^*(a,1)}(f(a))$ and therefore we also have $\sigma_{\mathcal{B}}(f(a)) = f(\sigma(a))$.

Proof. For existence let $b = \sqrt{a} := \varphi_a(\sqrt{\cdot})$. Now suppose that $c \in \mathcal{B}$ is Hermitian and $c^2 = a$. Then $a \in C^*(c,1)$ and c itself has its own associated functional calculus. Choose polynomials, p_n, so that $p_n \to \sqrt{\cdot}$ uniformly on $\sigma(a) = \sigma(c)^2$. If we let $q_n(x) = p_n(x^2)$, then

$$\begin{aligned}\max_{t\in\sigma(c)} |q_n(t) - |t|| &= \max_{t\in\sigma(c)} \left|p_n(t^2) - \sqrt{t^2}\right| = \max_{x\in\sigma(c)^2} \left|p_n(x) - \sqrt{x}\right| \\ &= \max_{x\in\sigma(c^2)} \left|p_n(x) - \sqrt{x}\right| = \max_{x\in\sigma(a)} \left|p_n(x) - \sqrt{x}\right| \to 0 \text{ as } n \to \infty,\end{aligned}$$

i.e. $q_n(x) \to |x|$ uniformly on $\sigma(c)$. Thus we may conclude,

$$b = \sqrt{a} = \lim_{n\to\infty} p_n(a) = \lim_{n\to\infty} p_n(c^2) = \lim_{n\to\infty} q_n(c) = |c|.$$

If we further assume that $\sigma(c) \subset [0,\infty)$ we will know that $|x| = x$ on $\sigma(c)$ and hence $b = |c| = c$ and the uniqueness or b is proved. ■

For the rest of this chapter we will explore the ramifications of having a C^*-algebra isomorphism of the form in Theorem 4.7. We will work more generally at this stage so that the results derived here will be applicable later when we have more general forms of Theorem 4.7 at our disposal.

4.3 Cyclic Vector and Subspace Decompositions

The first point we need to deal with is that understanding the structure of a C^*-subalgebra $(\mathcal{B})$ of $B(H)$ does not fully describe how $\mathcal{B}$ is embedded in $B(H)$. To understand the embedding problem we need to introduce the notation of cyclic vector and cyclic subspaces of H.

Definition 4.9 (Cyclic vectors). *If $\mathcal{A}$ is a sub-algebra of $B(H)$ a vector x in H is called a* ***cyclic****vector for $\mathcal{A}$ if $\mathcal{A}x \equiv \{Ax : A \in \mathcal{A}\}$ is dense in H. We further say that an $\mathcal{A}$ – invariant subspace, $M \subset H$, is an $\mathcal{A}$ –* ***cyclic subspace*** *of H if there exists $x \in M$ such that $\mathcal{A}x := \{Ax : A \in \mathcal{A}\}$ is dense in M.*

Lemma 4.10. *If $\mathcal{A}$ is a $*$ – sub-algebra of $B(H)$ and $M \subset H$ is an $\mathcal{A}$ – invariant subspace, then $\bar{M}$ and $M^\perp$ are $\mathcal{A}$ – invariant subspaces.*

Proof. If $m \in M$ and $m^\perp \in M^\perp$, then

$$\langle Am^\perp, m\rangle = \langle m^\perp, A^* m\rangle = 0$$

for all $A \in \mathcal{A}$ as $A^* \in \mathcal{A}$ ($\mathcal{A}$ is a $*$ – subalgebra). In other words, $\langle \mathcal{A}M^\perp, M\rangle = \{0\}$ and hence $\mathcal{A}M^\perp \subset M^\perp$. The assertion that $\bar{M}$ is also $\mathcal{A}$-invariant follows by a simple continuity argument. ■

Theorem 4.11. *Let H be a separable Hilbert space and $\mathcal{A}$ be a unital $*$ – subalgebra of $B(H)$ with identity. Then H may be decomposed into an orthogonal direct sum, $H = \oplus_{n=1}^{N} H_n$ ($N = \infty$ possible) such that H_n is a cyclic subspace of $\mathcal{A}$. [This cyclic decomposition is typically highly non-unique.]*

Proof. Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal basis for H and let

$$v_1 := e_1 \text{ and } H_1 := \overline{\mathcal{A}v_1}.$$

Then let $k_2 = \min\{k \in \mathbb{N} : e_k \notin H_1\}$ and let

$$v_2 := P_{H_1^{\perp}} e_{k_2} \text{ and } H_2 := \overline{\mathcal{A}v_2} \subset H_1^{\perp}.$$

Now let $k_3 := \min\min\{k \in \mathbb{N} : e_k \notin H_1 \oplus H_2\}$ and let

$$v_3 := P_{[H_1 \oplus H_2]^{\perp}} e_{k_3} \text{ and } H_3 := \overline{\mathcal{A}v_3}$$

and continue this way inductively forever or until $\{e_k\}_{k=1}^{\infty} \subset H_N$ for some $N < \infty$. ■

Exercise 4.2. Show (using Zorn's lemma say) that Theorem 4.11 holds without the assumption that H is separable. In this case the second item should be replaced by the statement that there exists an index set I and $\{v_\alpha\}_{\alpha \in I}$ a collection of non-zero vectors such that $H = \overset{\perp}{\oplus}_{\alpha \in I} H_\alpha$ (orthogonal direct sum) where $H_\alpha := H_{v_\alpha} = \overline{\mathcal{A}v_\alpha}^H$.

Before leaving this topic let us explore the meaning of cyclic vectors by looking at the finite dimensional case.

Proposition 4.12. *Let T be a $n \times n$-diagonal matrix, $T = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ for some $\lambda_i \in \mathbb{C}$ and set $\sigma(T) := \{\lambda_1, \dots, \lambda_n\}$. If $u \in \mathbb{C}^n$ is expressed as*

$$u = \sum_{\lambda \in \sigma(T)} e_\lambda \tag{4.5}$$

where $e_\lambda \in \operatorname{Nul}(T - \lambda I)$ for each $\lambda \in \sigma(T)$, then

$$\{p(T)u : p \in \mathbb{C}[z]\} = \operatorname{span}\{e_\lambda : \lambda \in \sigma(T)\}.$$

In particular, there is a cyclic vector for T iff $\#(\sigma(T)) = n$, i.e. all eigenvalues of T have multiplicity 1. In this case, one may take $u = \sum_{\lambda \in \sigma(T)} e_\lambda$ where $e_\lambda \in \operatorname{Nul}(T - \lambda I) \setminus \{0\}$ for all $\lambda \in \sigma(T)$. [Moral, the existence of a cyclic vector is equivalent to T having no repeated eigenvalues.]

Proof. If u is as in Eq. (4.5) and $p \in \mathbb{C}[z]$, then

$$p(T)u = \sum_{\lambda \in \sigma(T)} p(T)e_\lambda = \sum_{\lambda \in \sigma(T)} p(\lambda)e_\lambda.$$

As usual, given $\lambda_0 \in \sigma(T)$, we may choose $p \in \mathbb{C}[z]$ such that $p(\lambda) = \delta_{\lambda_0,\lambda}$ for all $\lambda \in \sigma(T)$. For this p we have $p(T)u = e_{\lambda_0}$ and hence we learn

$$\{p(T)u : p \in \mathbb{C}[z]\} = \operatorname{span}\{e_\lambda : \lambda \in \sigma(T)\}.$$

From this relation we see that maximum possible dimension of $\{p(T)u : p \in \mathbb{C}[z]\}$ is $\#(\sigma(T))$ which is equal to n iff $\#(\sigma(T)) = n$. ■

4.4 The Diagonalization Strategy

Definition 4.13 (Radon measure). *If Y is a locally compact Hausdorff space, let $\mathcal{F}_Y = \sigma(\{\text{open sets}\})$ be the Borel σ-algebra on Y. A measure μ on $(Y, \mathcal{F}_Y)$ is a **Radon measure** if it $\mu(K) < \infty$ when K is compact at it is a regular Borel measure, i.e.*

1. *μ is outer regular on Borel sets, i.e. if $A \in \mathcal{F}_Y$, then*
$$\mu(A) = \inf\{\mu(V) : A \subset V \subset_o Y\}, \text{ and}$$
2. *it is inner regular on open sets, i.e. if $V \subset_o Y$, then*
$$\mu(V) = \sup\{\mu(K) : K \subset V \text{ with } K \text{ compact}\}.$$

Proposition 4.14. *Suppose that Y is a compact Hausdorff space, H is a Hilbert space, $\mathcal{B}$ is a commutative unital C^*-subalgebra of $\mathcal{B}(H)$, and $\varphi : C(Y) \to \mathcal{B}$ is a given C^*-isomorphism of C^*-algebras. [This is in fact can always be arranged, see Theorem 8.14 below.] Then for each $v \in H \setminus \{0\}$, there exists a unique finite radon measure, μ_v, on $(Y, \mathcal{F}_Y)$ such that*

$$\langle \varphi(f)v, v\rangle = \int_Y f d\mu_v \;\forall\; f \in C(Y). \tag{4.6}$$

Proof. For $f \in C(Y)$, let $\Lambda(f) := \langle \varphi(f)v, v\rangle$ which is a linear functional on $C(Y)$. Moreover if $f \geq 0$, then $g = \sqrt{f} \in C(Y)$ and hence

$$\begin{aligned}\Lambda(f) = \Lambda(g^2) &= \langle \varphi(g^2)v, v\rangle = \langle \varphi(g)\varphi(g)v, v\rangle \\ &= \langle \varphi(g)v, \varphi(g)^* v\rangle = \langle \varphi(g)v, \varphi(\bar{g})v\rangle = \|\varphi(g)v\|^2 \geq 0.\end{aligned}$$

Thus Λ is a positive linear functional on $C(Y)$ and hence by the Riesz-Markov theorem there exists a unique (necessarily finite) Radon measure, μ_v, on $(Y, \mathcal{F}_Y)$ such that

$$\langle \varphi(f) v, v\rangle = \Lambda(f) = \int_Y f d\mu_v \ \forall\ f \in C(Y).$$

■

Proposition 4.15. *Continue the notation and assumptions in Proposition 4.14 and for each $v \in H \setminus \{0\}$, let*

$$H_v := \overline{\mathcal{B}v}^H \subset H. \tag{4.7}$$

Then there exists a unique unitary isomorphism, $U_v : L^2(\mu_v) \to H_v$ which is uniquely determined by requiring

$$U_v f = \varphi(f) v \in H_v \text{ for all } f \in C(Y). \tag{4.8}$$

Moreover, this unitary map satisfies,

$$U_v^* \varphi(f)|_{H_v} U_v = M_f \text{ on } L^2(\mu_v) \quad \forall\ f \in C(Y). \tag{4.9}$$

Proof. Since

$$\begin{aligned}\|U_v f\|^2 &= \langle \varphi(f) v, \varphi(f) v\rangle = \left\langle \varphi(f)^* \varphi(f) v, v\right\rangle \\ &= \left\langle \varphi(\bar f) \varphi(f) v, v\right\rangle = \left\langle \varphi\left(|f|^2\right) v, v\right\rangle = \int_Y |f|^2 d\mu_v = \|f\|_2,\end{aligned}$$

and $C(Y)$ is dense in $L^2(\mu_v)$, it follows that U_v extends uniquely to an isometry from $L^2(\mu_v)$ to H_v. Clearly U_v has dense range and the range is closed since U_v is isometric, therefore $\operatorname{Ran}(U_v) = H_v$ and hence U_v is unitary.

Let us further note that for $f, g \in C(Y)$,

$$U_v^* \varphi(f) U_v g = U_v^* \varphi(f) \varphi(g) v = U_v^* \varphi(fg) v = fg = M_f g. \tag{4.10}$$

If $g \in L^2(\mu_v)$, we may choose $\{g_n\} \subset C(Y)$ so that $g_n \to g$ in $L^2(\mu_v)$. So by replacing g by g_n in Eq. (4.10) and then passing to the limit as $n \to \infty$ we conclude It then follows that

$$U_v^* \varphi(f) U_v g = fg = M_f g \ \forall\ g \in L^2(\mu)$$

which proves Eq. (4.9). ■

Theorem 4.16. *Continue the notation and assumptions in Proposition 4.14. Then there exist $N \in \mathbb{N} \cup \{\infty\}$, a probability measure μ measure on*

$$\Omega := \Lambda_N \times Y = \sum_{j \in \Lambda_N} Y_j \text{ where } Y_j = \{j\} \times Y$$

equipped with the product σ – algebra (here $\Lambda_N = \{1, 2, \dots, N\} \cap \mathbb{N}$), and a unitary map $U : L^2(\mu) \to H$ such that

$$U^* \varphi(f) U = M_{f \circ \pi} \text{ on } L^2(\mu) \tag{4.11}$$

where $\pi : \Omega \to Y$ is defined by $\pi(j, w) = w$ for all $j \in \Lambda_N$ and $w \in Y$, see Figure 4.2.

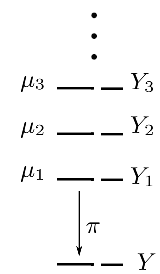

Fig. 4.2. Making disjoint copies of Y to take care of multiplicities.

Proof. By Theorem 4.11, there exists an $N \in \mathbb{N} \cup \{\infty\}$ so that we may decompose H into an orthogonal direct sum, $\oplus_{i \in \Lambda_N} H_i$, of cyclic subspaces for $\mathcal{B}$. Choose a cyclic vector, $v_i \in H_i$, for all $i \in \Lambda := \Lambda_N$ and normalize the $\{v_i\}_{i \in \Lambda}$ so that

$$\sum_{i \in \Lambda} \|v_i\|^2 = 1.$$

Let $\mu_i = \mu_{v_i}$ be the measure in Proposition 4.14 and let $\Omega := \Lambda \times Y$ which we equip with the product σ – algebra, $\mathcal{F}$, and the probability measure μ defined as follows. Every $G \in \mathcal{F}$ may be written (see Remark 4.17 below) may be uniquely written as

$$G = \sum_{i \in \Lambda} \{i\} \times G_i \text{ for some } \{G_i\}_{i \in \Lambda} \subset \mathcal{F}_Y$$

and if we let

$$\mu(G) := \sum_{i \in \Lambda} \mu_i(G_i),$$

then μ is a measure on $\mathcal{F}$. For this measure,

$$\int_\Omega g d\mu = \sum_{i \in \Lambda} \int_\Omega g 1_{\{i\} \times Y} d\mu = \sum_{i \in \Lambda} \int_\Omega g(i, \cdot) d\mu_i$$

From which it easily follows that the map,

$$L^2(\Omega,\mu) \ni g \to \{g(i,\cdot)\}_{i\in\Lambda} \in \oplus_{i\in\Lambda} L^2(Y,\mu_i)$$

is a unitary. For $g \in L^2(\Omega,\mu)$ we define,

$$Ug = \sum_{i\in\Lambda} U_{v_i} g(i,\cdot) \in \oplus_{i\in\Lambda} H_i = H,$$

where U_{v_i} is the unitary map in Proposition 4.15. Since

$$\|Ug\|_H^2 = \sum_{i\in\Lambda} \|U_{v_i} g(i,\cdot)\|_{H_i}^2 = \sum_{i\in\Lambda} \int_Y |g(i,w)|^2 \, d\mu_i(w) = \int_\Omega |g|^2 \, d\mu,$$

U is an isometry and since U has dense range it is in fact unitary. Lastly if $f \in C(Y)$ and $g \in L^2(\mu)$, we have

$$\begin{aligned} UM_{f\circ\pi} g &= \sum_{i\in\Lambda} U_{v_i}[f\circ\pi(i,\cdot)\, g(i,\cdot)] = \sum_{i\in\Lambda} U_{v_i}[f g(i,\cdot)] \\ &= \sum_{i\in\Lambda} U_{v_i}[M_f g(i,\cdot)] = \sum_{i\in\Lambda} \varphi(f) U_{v_i} g(i,\cdot) = \varphi(f) Ug. \end{aligned}$$

This completes the proof. ∎

Remark 4.17. The product σ-algebra on $\Lambda \times Y$ is given by the collection of sets

$$\mathcal{F} := \left\{ \sum_{j\in\Lambda} (\{j\} \times G_j) : \ \{G_j\}_{j=1}^\infty \subset \mathcal{F}_Y \right\}.$$

If is clear that every element in $\mathcal{F}$ is in the product σ-algebra and hence it suffices to shows $\mathcal{F}$ is a σ-algebra. The main point is to notice that if $G = \sum_{j\in\Lambda} (\{j\} \times G_j)$, then

$$(i,y) \in G^c \iff (i,y) \notin G \iff y \notin G_i \iff (i,y) \in \{i\} \times G_i^c.$$

This shows $G^c = \sum_{i\in\Lambda} (\{i\} \times G_i^c)$ which is graphically easy to understand.

To see that μ is a measure on $\mathcal{F}$, first observe that if $H = \sum_{i\in\Lambda} \{i\} \times H_i$, then

$$H \cap G = \sum_{i\in\Lambda} \{i\} \times [G_i \cap H_i]$$

and so if $\left\{G(n) = \sum_{i\in\Lambda} \{i\} \times G_i(n)\right\}_{n\in\Lambda}$ are pairwise disjoint then $\{G_i(n)\}_{n\in\Lambda}$ must be pairwise disjoint for each $i \in \Lambda$. Hence it follows that

$$\sum_{n\in\mathbb{N}} G(n) = \sum_{i\in\Lambda} \{i\} \times \left(\sum_{n\in\mathbb{N}} G_i(n) \right)$$

and therefore,

$$\begin{aligned} \mu\left(\sum_{n\in\mathbb{N}} G(n)\right) &= \sum_{i\in\Lambda} \mu_i \left(\sum_{n\in\mathbb{N}} G_i(n) \right) = \sum_{i\in\Lambda} \sum_{n\in\mathbb{N}} \mu_i(G_i(n)) \\ &= \sum_{n\in\mathbb{N}} \sum_{i\in\Lambda} \mu_i(G_i(n)) = \sum_{n\in\mathbb{N}} \mu(G(n)). \end{aligned}$$

Notation 4.18 *If $(\Omega,\mathcal{F})$ is a measurable space, let $\ell^\infty(\Omega,\mathcal{F})$ denote the bounded $\mathcal{F}/\mathcal{B}_\mathbb{C}$-measurable functions from Ω to $\mathbb{C}$.*

Let us now rewrite Eq. (4.11) as

$$\varphi(f) = UM_{f\circ\pi} U^* \text{ for } f \in C(Y). \tag{4.12}$$

From this equation we see there is a "natural" extension φ to a map, $\psi : \ell^\infty(Y,\mathcal{F}_Y) \to B(H)$ defined by

$$\psi(f) := UM_{f\circ\pi} U^* \text{ for all } f \in \ell^\infty(Y,\mathcal{F}_Y). \tag{4.13}$$

This map ψ has the following properties.

Theorem 4.19 (Measurable Functional Calculus I). *The map, $\psi : \ell^\infty(Y,\mathcal{F}_Y) \to B(H)$ in Eq. (4.13) has the following properties.*

1. *$\psi = \varphi$ on $C(Y)$.*
2. *$\|\psi(f)\| \le \|f\|_\infty$ for all $f \in \ell^\infty(Y,\mathcal{F}_Y)$.*
3. *If $f_n \in \ell^\infty(Y,\mathcal{F}_Y)$ converges to $f \in \ell^\infty(Y,\mathcal{F}_Y)$ boundedly then $\psi(f_n) \xrightarrow{s} \psi(f)$.*
4. *ψ is a C^*-algebra homomorphism.*
5. *If $f \ge 0$ then $\psi(f) \ge 0$.*

Proof. The proof of this theorem is straight forward and for the most part is left to the reader. Let me only verify items 3. and 5. here.

3. Let $u \in H$ and $g = U^* u \in L^2(\mu)$. Then

$$\begin{aligned} \|\psi(f)u - \psi(f_n)u\|^2 &= \|UM_{f\circ\pi}U^*u - UM_{f_n\circ\pi}U^*u\|^2 \\ &= \|[f\circ\pi - f_n\circ\pi]\, g\|_{L^2(\mu)}^2 \to 0 \text{ as } n \to \infty \end{aligned}$$

by DCT.

5. If $f \ge 0$, then

$$\langle \psi(f)u, u\rangle = \langle UM_{f\circ\pi}U^*u, u\rangle = \langle M_{f\circ\pi} g, g\rangle_{L^2(\mu)} = \int_\Omega f\circ\pi\, |g|^2 \, d\mu \ge 0.$$

Alternatively, simply note that $f = \left(\sqrt{f}\right)^2$ and hence

$$\psi(f) = \psi\left(\sqrt{f}\right)^2 = \psi\left(\sqrt{f}\right)^* \psi\left(\sqrt{f}\right) \ge 0.$$

∎

Definition 4.20. *If $\mathcal{B} \subset B(H)$, let $\mathcal{B}' := \{B \in B(H) : [B, \mathcal{B}] = \{0\}\}$ be the commutant of $\mathcal{B}$. Thus $A \in \mathcal{B}'$ iff $[A, B] = 0$ for all $B \in \mathcal{B}$.*

Remark 4.21. If (Y, d) is a compact metric space, then $\sigma(C(Y)) = \mathcal{F}_Y$ where $\sigma(C(Y))$ is the smallest σ-algebra on Y for which all continuous functions are measurable. Indeed we always have $\sigma(C(Y)) \subset \mathcal{F}_Y$ and so it suffices to show $V \in \sigma(C(Y))$ for all $V \subset_o Y$. However, if V is an open set, then $d_{V^c}(x) := \inf_{y \in V^c} d(x, y)$ is a continuous function on Y such that $V = \{d_{V^c} > 0\} \in \sigma(C(Y))$.

Proposition 4.22. *If Y is a compact metric space then there is precisely one map, $\psi : \ell^\infty(Y, \mathcal{F}_Y) \to B(H)$, which satisfies properties 1.-4. in Theorem 4.19. Moreover the image of this map is in $\mathcal{B}''$.*

Proof. If $\hat{\psi} : \ell^\infty(Y, \mathcal{F}_Y) \to B(H)$ also satisfies items 1.-4. of Theorem 4.19, let

$$\mathbb{H} = \left\{ f \in \ell^\infty(Y, \mathcal{F}_Y) : \psi(f) = \hat{\psi}(f) \right\}.$$

One then easily verifies that $\mathbb{H}$ is closed is a subspace of $\ell^\infty(Y, \mathcal{F}_Y)$ which is closed under conjugation and bounded convergence and hence by the multiplicative system Theorem A.9 it follows that $\mathbb{H}$ contains all bounded $\sigma(C(Y)) = \mathcal{F}_Y$-measurable functions, i.e. $\mathbb{H} = \ell^\infty(Y, \mathcal{F}_Y)$.

To prove the second assertion, let

$$\mathbb{H} = \{ f \in \ell^\infty(Y, \mathcal{F}_Y) : [\psi(f), \mathcal{B}'] = \{0\} \}.$$

Then $\mathbb{H}$ is a linear space closed under conjugation and bounded convergence and contains $C(Y)$ as the reader should verify. Thus by another application of the multiplicative system Theorem A.9, $\mathbb{H} = \ell^\infty(Y, \mathcal{F}_Y)$ and the proof is complete. ■

Corollary 4.23 (Spectral Theorem I). *Let H be a separable Hilbert space and $A \in B(H)$ be a self-adjoint operator. Then there exists a finite measure space, $(\Omega, \mathcal{F}, \mu)$, a bounded function, $a : \Omega \to \sigma(A)$, and a unitary map, $U : L^2(\mu) \to H$, such that $A = U M_a U^*$. Moreover, if $f \in \ell^\infty\left(\sigma(A), \mathcal{F}_{\sigma(A)}\right)$, then*

$$\psi_A(f) = \text{“}f(A)\text{”} = U M_{f \circ a} U^*$$

defines the unique measurable functional calculus in this setting.

Proof. Let $\mathcal{B} = C^*(A, I) \subset B(H)$ and then by Theorem 4.7, there exists C^*-isomorphism, $\varphi_A : C(\sigma(A)) \to \mathcal{B}$ such that $\varphi_A(p) = p(A)$. To complete the proof of the theorem, we apply Theorem 4.16 with $\varphi = \varphi_A$ and take $a = id \circ \pi$ where $id : \sigma(A) \to \sigma(A)$ is the identity map. as in the language of Theorem 4.16. ■

The next theorem summarizes the result we have proved for a self-adjoint element, $A \in B(H)$.

Theorem 4.24 (Measurable Functional Calculus for a Hermitian). *Let H be a separable Hilbert space and A be a self-adjoint element of $B(H)$. Then there exists a unique map $\psi_A : \ell^\infty\left(\sigma(A), \mathcal{F}_{\sigma(A)}\right) \to B(H)$ such that;*

1. *ψ_A is a $*$ – homomorphism, i.e. ψ_A is linear, $\psi_A(fg) = \psi_A(f)\psi_A(g)$ and $\psi_A(\bar{f}) = \psi_A(f)^*$ for all $f, g \in \ell^\infty(\sigma(A))$.*
2. *$\|\psi_A(f)\|_{op} \le \|f\|_\infty$ for all $f \in \ell^\infty(\sigma(A))$.*
3. *$\psi_A(p) = p(A)$ for all $p \in \mathbb{C}[x]$. [Equivalently $\varphi(1) = I$ and $\psi_A(x) = A$ where $x : \sigma(A) \to \sigma(A)$ is the identity map.]*
4. *If $f_n \in \ell^\infty(\sigma(A))$ and $f_n \to f$ pointwise and boundedly, then $\psi_A(f_n) \to \psi_A(f)$ strongly.*
Moreover this map has the following properties.
5. *If $f \ge 0$ then $\psi_A(f) \ge 0$.*
6. *If $B \in B(H)$ and $[B, A] = 0$, then $[B, \psi_A(f)] = 0$ for all $f \in \ell^\infty(\sigma(A))$.*
7. *If $Ah = \lambda h$ for some $h \in H$ and $\lambda \in \mathbb{R}$, then $\psi_A(f)h = f(\lambda) h$.*

Proof. Although there is no need to give a proof here, we do so anyway in order to solidify the above ideas in this concrete special case.

Uniqueness. Suppose that $\psi : \ell^\infty(\sigma(A)) \to B(H)$ is another map satisfying $(1) - (4)$. Let

$$\mathbb{H} := \{ f \in \ell^\infty(\sigma(A), \mathbb{C}) : \psi(f) = \psi_A(f) \}.$$

Then $\mathbb{H}$ is a vector space of bounded complex valued functions which by property 4. is closed under bounded convergence and by property 1. is closed under conjugation. Moreover $\mathbb{H}$ contains

$$\mathbb{M} = \left\{ p|_{\sigma(A)} : p \in \mathbb{C}[x] \right\}$$

and therefore also $C(\sigma(A), \mathbb{C})$ because of the Stone–Weierstrass approximation theorem. Therefore it follows from Theorem A.9 that $\mathbb{H} = \ell^\infty(\sigma(A))$, i.e. $\psi = \psi_A$.

Existence. Let $U : L^2(\Omega, \mu) \to H$ be as in Corollary 4.23 and then define

$$\psi_A(f) := U M_{f \circ a} U^* \ \forall \ f \in \ell^\infty(\sigma(A)).$$

One easily verifies that ψ_A satisfies items 1. – 4. Moreover we can easily verify items 5–7 as well.

5. If $f \ge 0$, then $f = \left(\sqrt{f}\right)^2$ and hence $\psi_A(f) = \psi_A\left(\sqrt{f}\right)^2 \ge 0$.
6. Let

$$\mathbb{H} := \{ f \in \ell^\infty(\sigma(A), \mathbb{C}) : [B, \psi_A(f)] = 0 \}$$

which is vector space closed under conjugation[3] and bounded convergence. It is easily deduced from $[B, A] = 0$ that $[B, p(A)] = 0$ for all $p \in \mathbb{C}[x]$, the result

[3] Again we use Theorem 2.68 and the fact that $\psi_A(f)$ is normal for all f.

follows by an application of the multiplicative system Theorem A.9 applied using the multiplicative system,

$$\mathbb{M} = \left\{ p|_{\sigma(A)} : p \in \mathbb{C}\left[x\right] \right\}.$$

7. If $Ah = \lambda h$ and $g := U^* h$, then $M_a g = \lambda g$ from which it follows that $(a - \lambda)\, g = 0 \;\; \mu -$ a.e. which implies $a = \lambda \; \mu -$ a.e. on $\{g \neq 0\}$. Thus it follows that $f \circ a = f(\lambda) \; \mu -$ a.e. on $\{g \neq 0\}$ and this implies $M_{f \circ a} g = f(\lambda)\, g$ which then implies,

$$\psi_A(f)h = \psi_A(f)Ug = UM_{f \circ a} g = Uf(\lambda)\, g = f(\lambda)\, h.$$

■

**More Measurable Functional Calculus

This highly optional chapter contains more details on the general construction of the measurable functional calculus.

5.1 Constructing a Measurable Functional Calculus

Assumption 1 *In this chapter we will assume that Y is a compact Hausdorff space, H is a Hilbert space, $\mathcal{B}$ is a commutative unital C^*-subalgebra of $\mathcal{B}(H)$, and $\varphi : C(Y) \to \mathcal{B}$ is a given C^* isomorphism of C^*-algebras. [This is in fact can always be arranged, see Theorem 8.14 below.]*

Let us start by recording some notation and results we introduced in Proposition 4.14 and Proposition 4.15.

Notation 5.1 *For each $v \in H$, we let*

$$H_v := \overline{\mathcal{B}v}^H \subset H \tag{5.1}$$

and μ_v be the unique (finite) Radon measure on $(Y, \mathcal{F}_Y)$ such that

$$\langle \varphi(f) v, v \rangle = \int_Y f d\mu_v \ \forall \ f \in C(Y), \tag{5.2}$$

see Proposition 4.14. Further let $U_v : L^2(\mu_v) \to H_v$ be the unique unitary isomorphism determined by

$$U_v f = \varphi(f) v \in H_v \text{ for all } f \in C(Y) \tag{5.3}$$

which satisfies

$$\varphi(f)|_{H_v} = U_v M_f U_v^* \text{ on } L^2(\mu_v) \quad \forall \ f \in C(Y) \tag{5.4}$$

as in Proposition 4.15.

Notation 5.2 *Using Theorem 4.11 when H is separable or Exercise 4.2 for general H, let us choose (and fix) $\{v_\alpha\}_{\alpha \in I} \subset H$ such that $H = \overset{\perp}{\oplus}_{\alpha \in I} H_{v_\alpha}$ and let P_α denote orthogonal projection onto H_{v_α} for each $\alpha \in I$.*

For $f \in C(Y)$ and $u \in H$, we have

$$\varphi(f) u = \varphi(f) \sum_{\alpha \in I} P_\alpha u = \sum_{\alpha \in I} \varphi(f) P_\alpha u = \sum_{\alpha \in I} U_{v_\alpha} M_f U_{v_\alpha}^* P_\alpha u \tag{5.5}$$

wherein we used Eq. (5.4) for the last equality. In light of Eq. (5.5), the following definition is a "natural" extension of φ to $f \in \ell^\infty(Y, \mathcal{F}_Y)$.

Definition 5.3 (Construction of ψ). *Continuing the notation above, let $\psi : \ell^\infty(Y, \mathcal{F}_Y) \to B(H)$ be defined by*

$$\psi(f) u := \sum_{\alpha \in I} U_{v_\alpha} M_f U_{v_\alpha}^* P_\alpha u \text{ for all } u \in H. \tag{5.6}$$

In other words, $\psi(f)$ is given in block diagonal form as

$$\psi(f) = \operatorname{diag}\left(\left\{U_{v_\alpha} M_f U_{v_\alpha}^*\right\}_{\alpha \in I}\right). \tag{5.7}$$

Theorem 5.4. *The map, $\psi : \ell^\infty(Y, \mathcal{F}_Y) \to B(H)$ in Definition 5.3 has the following properties.*

1. *$\psi = \varphi$ on $C(Y)$.*
2. *$\|\psi(f)\| \le \|f\|_\infty$ for all $f \in \ell^\infty(Y, \mathcal{F}_Y)$.*
3. *If $f_n \in \ell^\infty(Y, \mathcal{F}_Y)$ converges to $f \in \ell^\infty(Y, \mathcal{F}_Y)$ boundedly then $\psi(f_n) \xrightarrow{s} \psi(f)$.*
4. *ψ is a C^*-algebra homomorphism.*
5. *If $f \ge 0$ then $\psi(f) \ge 0$.*

Proof. Recall $I_u := \{\alpha \in I : u_\alpha := P_\alpha u \neq 0\}$ is at most countable for each $u \in H$ and so

$$u = \sum_{\alpha \in I} u_\alpha = \sum_{\alpha \in I_u} u_\alpha \text{ – a countable sum.}$$

As the reader should verify, $\psi(f) : H \to H$ is linear and $\psi(1) = I_H$. We now prove the remaining items in turn.

1. That ψ is an extension of φ follows from Eq. (5.5).
   For the rest of this proof let $u \in H$,

$$g_\alpha := U_{v_\alpha}^* P_\alpha u \in L^2(\mu_{v_\alpha}) \text{ for } \alpha \in I,$$

and m_u be the measure on $(Y, \mathcal{F}_Y)$ defined by

$$dm_u := \sum_{\alpha \in I_u} |g_\alpha|^2 \, d\mu_{v_\alpha}. \tag{5.8}$$

2. For $f \in \ell^\infty(Y, \mathcal{F}_Y)$,

$$\begin{aligned}\|\psi(f)u\|^2 &= \sum_{\alpha \in I} \left\|U_{v_\alpha} M_f U_{v_\alpha}^* P_\alpha u\right\|^2 = \sum_{\alpha \in I} \left\|M_f U_{v_\alpha}^* P_\alpha u\right\|^2_{L^2(\nu_\alpha)} \\ &= \sum_{\alpha \in I} \|f g_\alpha\|^2_{L^2(\nu_\alpha)} = \int_Y |f|^2 \, dm_u. \end{aligned} \tag{5.9}$$

Taking $f = 1$ in this equation shows

$$m_u(Y) = \|u\|^2 < \infty \tag{5.10}$$

which combined with Eq. (5.9) shows

$$\|\psi(f)u\|^2 \le \|f\|_\infty^2 \|u\|^2$$

which proves item 2.

3. Item 3. is now also easily proved since if $f_n \to f$ boundedly then

$$\begin{aligned}\|\psi(f)u - \psi(f_n)u\|^2 &= \|\psi(f - f_n)u\|^2 \\ &= \int_Y |f - f_n|^2 \, dm_u \to 0 \text{ as } n \to \infty. \end{aligned}$$

by DCT.

4. For $f, g \in \ell^\infty(Y, \mathcal{F}_Y)$ and $u \in H$,

$$\begin{aligned}\psi(f)\psi(g)u &= \sum_{\alpha \in I} \psi(f) U_{v_\alpha} M_g U_{v_\alpha}^* P_\alpha u \\ &= \sum_{\alpha \in I} U_{v_\alpha} M_f U_{v_\alpha}^* U_{v_\alpha} M_g U_{v_\alpha}^* P_\alpha u \\ &= \sum_{\alpha \in I} U_{v_\alpha} M_f M_g U_{v_\alpha}^* P_\alpha u \\ &= \sum_{\alpha \in I} U_{v_\alpha} M_{fg} U_{v_\alpha}^* P_\alpha u = \psi(fg)u \end{aligned}$$

which shows $\psi(fg) = \psi(f)\psi(g)$. Moreover for another $v \in H$,

$$\begin{aligned}\langle \psi(f)u, v\rangle &= \sum_{\alpha \in I} \left\langle U_{v_\alpha} M_f U_{v_\alpha}^* P_\alpha u, v\right\rangle = \sum_{\alpha \in I} \left\langle P_\alpha U_{v_\alpha} M_f U_{v_\alpha}^* P_\alpha u, v\right\rangle \\ &= \sum_{\alpha \in I} \left\langle M_f U_{v_\alpha}^* P_\alpha u, U_{v_\alpha}^* P_\alpha v\right\rangle_{L^2(\mu_{v_\alpha})} \\ &= \sum_{\alpha \in I} \left\langle U_{v_\alpha}^* P_\alpha u, M_{\bar f} U_{v_\alpha}^* P_\alpha v\right\rangle_{L^2(\mu_{v_\alpha})} \\ &= \sum_{\alpha \in I} \left\langle u, U_{v_\alpha} M_{\bar f} U_{v_\alpha}^* P_\alpha v\right\rangle = \left\langle u, \psi(\bar f) v\right\rangle \end{aligned} \tag{5.11}$$

which shows $\psi(f)^* = \psi(\bar f)$ and item 3. is proved.

5. Taking $v = u$ in Eq. (5.11) shows,

$$\begin{aligned}\langle \psi(f)u, u\rangle &= \sum_{\alpha \in I_u} \left\langle M_f U_{v_\alpha}^* P_\alpha u, U_{v_\alpha}^* P_\alpha u\right\rangle_{L^2(\mu_{\nu_\alpha})} \\ &= \sum_{\alpha \in I_u} \langle M_f g_\alpha u, g_\alpha u\rangle_{L^2(\mu_{\nu_\alpha})} = \sum_{\alpha \in I_u} \int_Y f |g_\alpha|^2 \, d\mu_{v_\alpha}, \end{aligned}$$

i.e.

$$\langle \psi(f)u, u\rangle = \int_Y f dm_u \text{ for all } f \in \ell^\infty(Y, \mathcal{F}_Y). \tag{5.12}$$

It clearly follows from this identity that $\psi(f) \ge 0$ if $f \ge 0$.

■

Proposition 5.5. *If we now further assume*[1] *that $\mathcal{F}_Y = \mathcal{F}_0$, then there is precisely one map, $\psi : \ell^\infty(Y, \mathcal{F}_Y) \to B(H)$ such that properties in items 1.-4. of Theorem 5.4 hold and moreover ψ is uniquely determined by*

$$\langle \psi(f)u, u\rangle = \int_Y f d\mu_u \textit{ for all } u \in H \textit{ and } f \in \ell^\infty(Y, \mathcal{F}_Y). \tag{5.13}$$

Proof. Suppose that $\hat\psi : \ell^\infty(Y, \mathcal{F}_Y) \to B(H)$ also satisfies items 1.-4. of Theorem 5.4. Then let

$$\mathbb{H} = \left\{ f \in \ell^\infty(Y, \mathcal{F}_Y) : \psi(f) = \hat\psi(f) \right\}.$$

One then easily verifies that $\mathbb{H}$ is closed is a subspace of $\ell^\infty(Y, \mathcal{F}_Y)$ which is closed under conjugation and bounded convergence and hence by the multiplicative system Theorem A.9 it follows that $\mathbb{H}$ contains all bounded Baire-measurable functions, i.e. $\ell^\infty(Y, \mathcal{F}_0) \subset \mathbb{H} \subset \ell^\infty(Y, \mathcal{F}_Y)$. Since we are assuming $\mathcal{F}_Y = \mathcal{F}_0$, it follows that $\ell^\infty(Y, \mathcal{F}_0) = \mathbb{H} = \ell^\infty(Y, \mathcal{F}_Y)$. To finish the proof it

[1] This will be the case if Y is metrizable for example, see below.

suffices to show Eq. (5.13) holds which is equivalent to showing the measure m_u in Eq. (5.12) is μ_u. However, we do know that

$$\int_Y f d\mu_u = \langle \varphi(f) u, u\rangle = \langle \psi(f) u, u\rangle = \int_Y f dm_u \ \forall\ f \in C(Y). \qquad (5.14)$$

From this identity and a simple application of the multiplicative system Theorem A.9 it follows from Eq. (5.14) that $m_u = \mu_u$ and the proof is complete. ■

Proposition 5.6. *Continuing the notation and setup in Proposition 5.5, then $\psi(\ell^\infty(Y, \mathcal{F}_Y)) \subset \mathcal{B}''$, the double commutant of $\mathcal{B}$. That is $A \in B(H)$ and $[A, \mathcal{B}] = \{0\}$, then $[A, \psi(\ell^\infty(Y, \mathcal{F}_Y))] = \{0\}$. In words, for $f \in \ell^\infty(Y, \mathcal{F}_Y)$, $\psi(f)$ commutes every $A \in B(H)$ which commutes with every $B \in \mathcal{B}$.*

Proof. Let $\mathbb{H} = \{f \in \ell^\infty(Y, \mathcal{F}_Y) : [\psi(f), \mathcal{B}'] = \{0\}\}$. Then $\mathbb{H}$ is a linear space closed under conjugation and bounded convergence and contains $C(Y)$ as the reader should verify. Thus by the multiplicative system Theorem A.9, $\ell^\infty(Y, \mathcal{F}_0) \subset \mathbb{H} \subset \ell^\infty(Y, \mathcal{F}_Y)$ and as we assume $\mathcal{F}_0 = \mathcal{F}_Y$, the proof is complete. ■

For the remainder of this chapter we are going to remove the added assumption that $\mathcal{F}_Y = \mathcal{F}_0$. Before getting down to business we need to take care some measure theoretic details.

5.2 Baire Sets and Radon Measures

Theorem 5.7 (Properties of Locally compact spaces). *Suppose (X, τ) is a locally compact Hausdorff space where τ is the collection of open subsets of X. We write $C \sqsubset X$ and $K \sqsubset\sqsubset X$ to indicate that C is a closed subset of X and K is a compact subset of X respectively.*

1. *It $K \sqsubset\sqsubset X$ and $K \subset U \cup V$ with $U, V \in \tau$, then $K = K_1 \cup K_2$ with $K_1 \sqsubset\sqsubset U$ and $K_2 \sqsubset\sqsubset V$.*
2. *If $K \sqsubset\sqsubset X$ and $F \sqsubset X$ are disjoint, then there exists $f \in C(X, [0,1])$ such that $f = 0$ on K and $f = 1$ on F.*
3. *If f is a real valued continuous function, then for all $c \in \mathbb{R}$ the sets $\{f \geq c\}$, $\{f \leq c\}$, and $\{f = c\}$ are closed $\mathcal{G}_\delta$.*
4. *If $K \sqsubset\sqsubset U \subset_o X$ then there exists $K \sqsubset\sqsubset U_0 \subset K_0 \subset U$ such that K_0 is a Baire measurable set and a compact $\mathcal{G}_\delta$ and U_0 is a σ – compact open set.*

Proof. We take each item in turn.

1. $K \backslash U$ and $K \backslash V$ are disjoint compact sets and hence there exists two disjoint open sets U' and V' such that
$$K \setminus U \subset V' \text{ and } K \setminus V \subset U'.$$
Let $K_1 := K \setminus V' \subset U$ and $K_2 = K \setminus U' \subset V$.
2. Tietze extension theorem with elementary proof in Halmos.
3. $\{f \leq c\} = \cap_{n=1}^\infty \{f < c + 1/n\}$ with similar formula for the other cases.
4. For each $x \in K$, let V_x be an open neighborhood of K such that $\bar{V}_x \sqsubset\sqsubset U$, and set $V = \cup_{x \in \Lambda} V_x$ where $\Lambda \subset\subset K$ is a finite set such that $K \subset V$. Since $\bar{V} = \cup_{x \in \Lambda} \bar{V}_x$ is compact, we may replace U by V if necessary and assume that U is pre-compact. By Urysohn's lemma, there exists $f \in C_c(X, [0,1])$ such that $f = 0$ on K and $f = 1$ on U^c. If we now defined $U_0 = \{f < 1/2\}$ and $K_0 = \{f \leq 1/2\}$ then $K \sqsubset\sqsubset U_0 \subset K_0 \subset U$, K_0 is a Baire set which is compact and a $\mathcal{G}_\delta$ by item 2. Moreover U_0 is σ-compact because
$$U_0 = \{f < 1/2\} = \cup_{n=3}^\infty \{f \leq 1/2 - 1/n\}.$$

■

Definition 5.8 (Borel and Baire σ-algebras). *Let $\mathcal{F}_X$ denote the **Borel σ-algebra** on X, i.e. the σ-algebra generated by open sets and $\mathcal{F}_0$ be the **Baire-σ-algebra**, i.e. the sigma algebra, $\sigma(C_c(X))$, generated by $C_c(X)$. A **Baire measure** is a positive measure, μ_0, on $(X, \mathcal{F}_0)$ which is finite on compact Baire sets.*

Notation 5.9 *If $(\Omega, \mathcal{F})$ is a general measurable space we let $\ell^\infty(\Omega, \mathcal{F})$ denote the bounded $\mathcal{F}/\mathcal{B}_{\mathbb{C}}$ – measurable functions, $f : \Omega \to \mathbb{C}$.*

For the rest of this section we will suppose that Y is a **compact** Hausdorff space.

Theorem 5.10 (Riesz-Markov Theorem). *Let Y be a compact Hausdorff space. There is a one to one correspondence between positive linear functionals, $\Lambda : C_c(Y, \mathbb{C}) \to \mathbb{C}$, Radon measures μ on $(Y, \mathcal{F}_Y)$, and Baire measures μ_0 on $(Y, \mathcal{F}_0)$ determined by;*

$$\Lambda(f) = \int_Y f d\mu = \int_Y f d\mu_0 \text{ for all } f \in C(Y),$$

and $\mu_0 = \mu|_{\mathcal{F}_0}$.

Proof. The main point is that if μ_0 is a Baire measure on Y, then $\Lambda(f) := \int_Y f d\mu_0$ is a positive linear functional on $C(Y, \mathbb{C})$. Therefore, by the Riesz-Markov theorem, there exists a unique Radon measure, μ, on $(Y, \mathcal{F}_Y)$ such that

$$\int_Y f d\mu = \int_Y f d\mu_0 \text{ for all } f \in C(Y). \tag{5.15}$$

It is now a simple application of the multiplicative system Theorem A.9 to show Eq. (5.15) is valid for all $f \in \ell^\infty(Y, \mathcal{F}_0)$ and hence $\mu_0 = \mu|_{\mathcal{F}_0}$. ■

Remark 5.11. In general it is not true that $\mathcal{F}_0 = \mathcal{F}_Y$, only that $\mathcal{F}_0 \subset \mathcal{F}_Y$. This is the reason one uses Radon measures on $(Y, \mathcal{F}_Y)$ rather than arbitrary measures. For the reader wishing to avoid such unpleasantries (at least on first reading) should further assume Y is metrizable, i.e. the topology on Y is induced from a metric, d, on Y. By Remark 4.21, it follows that $\mathcal{F}_0 = \mathcal{F}_Y$ and as a consequence, if Y is metrizable, then all finite measures on $(Y, \mathcal{B}_Y)$ are in fact Radon-measures, see Theorem 5.10.

Exercise 5.1. Let Y be a compact Hausdorff space. Prove the following assertions.

1. If μ is a Radon measure and $0 \le f \in L^1(Y, \mathcal{F}_Y, \mu)$, then $d\nu = f d\mu$ *is* a Radon measure.
2. If μ_1, μ_2 are two Radon measures, then so is $\mu_1 + \mu_2$.
3. Suppose that $\{\mu_j\}_{j=1}^\infty$ are finite Radon measures such that $\mu := \sum_{j=1}^\infty \mu_j$ is finite measure. Then μ is a Radon measure on $(Y, \mathcal{F}_Y)$.

5.3 Generalization to arbitrary compact Hausdorff spaces

Lemma 5.12. *If $f \in \ell^\infty(Y, \mathcal{F}_Y)$ and $v \in H$, then $\langle U_v f, v\rangle = \int_Y f d\mu_v$. In particular if $f \le g$ then $\langle U_v f, v\rangle \le \langle U_v g, v\rangle$.*

Proof. The result holds for all $f \in C(Y)$ by definition of μ_v and U_v, see Eqs. (5.2) and (5.3). Given a general $f \in \ell^\infty(Y, \mathcal{F}_Y)$ we may find $f_n \in C(Y)$ such that $f_n \to f$ in $L^2(\mu_v)$ and therefore,

$$\langle U_v f, v\rangle = \lim_{n\to\infty} \langle U_v f_n, v\rangle = \lim_{n\to\infty} \int_Y f_n d\mu_v = \int_Y f d\mu_v.$$

■

Lemma 5.13. *If $V \subset_o Y$ then*

$$\sup_{K \subset_k V} \langle U_v 1_K, v\rangle = \langle U_v 1_V, v\rangle$$

and if $E \in \mathcal{F}_Y$, then

$$\inf_{E \subset V \subset_o Y} \langle U_v 1_V, v\rangle = \langle U_v 1_E, v\rangle.$$

Proof. The proof of these statements are elementary consequences of Lemma 5.12 and the fact that μ_v is a Radon measure. ■

Theorem 5.14. *The map, $\psi : \ell^\infty(Y, \mathcal{F}_Y) \to B(H)$ in Definition 5.3 has the following additional properties over those stated in Theorem 5.4.*

1. *ψ satisfies Eq. (5.13), i.e.*
$$\langle \psi(f) u, u\rangle = \int_Y f d\mu_u \ \forall\ f \in \ell^\infty(Y, \mathcal{F}_Y) \text{ and } u \in H.$$
This identity, because of Lemma 3.26, uniquely specifies $\psi(f)$ and hence shows that ψ is independent of the choices made in the cyclic subspace decomposition of H.
2. *ψ is regular in the following sense;*
   a) *if $E \in \mathcal{F}_Y$ then*
   $$\psi(1_E) = \inf_{E \subset V \subset_o Y} \psi(1_V)$$
   or by abuse of notation
   $$\psi(E) = \inf\{\psi(V) : E \subset V \subset_o Y\}.$$
   [We abuse notation here and are writing $\psi(E)$ to mean $\psi(1_E)$ where E is a Borel set.]
   b) *If $V \subset_o Y$, then*
   $$\psi(1_V) = \sup_{K \subset_k V} \psi(1_K)$$
   or by abuse of notation, $\psi(V) = \sup\{\psi(K) : K \sqsubset\sqsubset V\}$.

Proof. Let $dm_u = \sum_{\alpha \in I_u} |g_\alpha|^2 d\mu_{\nu_\alpha}$ as in Eq. (5.8) in Theorem 5.4.

1. By item 1. of Exercise 5.1 $|g_\alpha|^2 d\mu_{\nu_\alpha}$ is a regular Radon measure and then by item 3. of the same exercise it follows that m_u is a Radon measure. Therefore from Eq. (5.14) and the uniqueness assertion in the Riesz-Markov theorem, we may conclude $\mu_u = m_u$ which coupled with Eq. (5.12) completes the proof of item 1.
2. The regularity statements follows by combining Lemmas 5.21 and Lemma 5.13.

■

Theorem 5.15. *There is exactly one C^*-homomorphism, $\psi : \ell^\infty(Y, \mathcal{F}_Y) \to B(H)$ such that Properties 1.-4. of Theorem 5.4 and the regularity property in item 2. of Theorem 5.14.*

Proof. We have already proved existence and so we now need only prove uniqueness. Let $\hat{\psi} : \ell^\infty (Y, \mathcal{F}_Y) \to B(H)$ be another C^*-homomorphism satisfying the stated properties in the statement of the theorem. Following the proof of Proposition 5.5 we already know that $\hat{\psi} = \psi$ on $\ell^\infty (Y, \mathcal{F}_0)$. We now need to use the regularity assumption to extend the identity to all of $\ell^\infty (Y, \mathcal{F}_Y)$ which we now do.

Let $V \subset_o Y$. By item 4. of Theorem 5.7, to each compact set, $K \subset V$, there exists a Baire measurable compact set K_0 such that $K \subset K_0 \subset V$. Thus by the given regularity and equality of ψ and $\hat{\psi}$ on Baire sets we may conclude that

$$\psi (1_V) = \sup \left\{ \psi (1_{K_0}) = \hat{\psi} (1_{K_0}) : V \supset K_0 \text{ compact \& Baire} \right\} = \hat{\psi} (1_V).$$

Then given any Borel measurable set $E \subset Y$ we find,

$$\psi (1_E) = \inf \left\{ \psi (1_V) = \hat{\psi} (1_V) : E \subset V \subset_o Y \right\} = \hat{\psi} (1_E).$$

By linearity, $\psi = \hat{\psi}$ on all Borel simple functions and then by taking uniform limits we conclude that $\psi = \hat{\psi}$ on $\ell^\infty (Y, \mathcal{F}_Y)$. ■

Proposition 5.16. *If $f \in \ell^\infty (Y, \mathcal{F}_Y)$, then $\psi (f) \in \mathcal{B}''$, i.e. if $[A, B] = 0$ for all $B \in \mathcal{B}$ then $[A, \psi (f)] = 0$. In words, $\psi (f)$ commutes with every operator, $A \in B(H)$, that commutes with every operator in $\mathcal{B}$.*

Proof. Suppose that $A \in \mathcal{B}'$. As $\psi (f) = \varphi (f) \in \mathcal{B}$ for all $f \in C(Y)$ it follows that $[\psi (f), A] = 0$. An application of the multiplicative system Theorem A.9. then shows that $[\psi (f), A] = 0$ for all Baire measurable bounded functions, $f : Y \to \mathbb{C}$. Now suppose that $V \subset_o Y$. By item 4. of Theorem 5.7, to each compact set, $K \subset V$, there exists a Baire measurable compact set K_0 such that $K \subset K_0 \subset V$. Therefore by the regularity of ψ as proved in Theorem 5.14, we may conclude that

$$\psi (1_V) = \sup_{K_0 \subset V} \psi (1_{K_0})$$

which along with Lemma 5.20 shows that $[\psi (1_V), A] = 0$. Then given any $E \in \mathcal{F}_Y$, we have

$$\psi (1_E) = \inf_{E \subset V \subset_0 Y} \psi (1_V)$$

also commutes with A, again by Lemma 5.20. Therefore $[\psi (f), A] = 0$ for any simple function in $\ell^\infty (Y, \mathcal{F}_Y)$ and therefore by uniformly approximating $f \in \ell^\infty (Y, \mathcal{F}_Y)$ by Borel simple functions shows $[\psi (f), A] = 0$ for all $f \in \ell^\infty (Y, \mathcal{F}_Y)$.

■

5.4 Appendix: Operator Ordering and the Lattice of Orthogonal Projections

Exercise 5.2. Suppose that $T \in B(H)$, M is a closed subspace of H, and $P = P_M$ is orthogonal projection onto M. Show $0 = [T, P] := TP - PT$ iff $TM \subset M$ and $T^* M \subset M$.

See Definition **??** and related material for operator ordering basics.

Definition 5.17. *If P and Q are two orthogonal projections on a Hilbert space H, then we write $P \le Q$ to mean* $\operatorname{Ran}(Q) \subset \operatorname{Ran}(P)$. *This defines a partial ordering on the collection of orthogonal projection on H. If $\mathcal{P}$ is an family of orthogonal projections on H then an orthogonal projection, Q, is an **upper bound (lower bound)** for $\mathcal{P}$ if $P \le Q$ ($Q \le P$) for all $P \in \mathcal{P}$.*

Remark 5.18. The notation $P \le Q$ is also consistent with the common meaning of ordering of self-adjoint operators given by $A \le B$ iff $\langle Av, v\rangle \le \langle Bv, v\rangle$ for all $v \in H$. Indeed if $\operatorname{Ran}(P) \subset \operatorname{Ran}(Q)$ and $v \in H$ then $Qv = PQv + w = Pv + w$ where $w \perp Pv$ and hence,

$$\langle Qv, v\rangle = \|Qv\|^2 \ge \|Pv\|^2 = \langle Pv, v\rangle.$$

Conversely if $\langle Pv, v\rangle \le \langle Qv, v\rangle$ for all $v \in H$, then by taking $v \in \operatorname{Ran}(Q)^\perp$ we learn that

$$\|Pv\|^2 = \langle Pv, v\rangle \le \langle Qv, v\rangle = 0$$

so that $v \in \operatorname{Ran}(P)^\perp$, i.e. $\operatorname{Ran}(Q)^\perp \subset \operatorname{Ran}(P)^\perp$. Taking orthogonal compliments then shows $\operatorname{Ran}(P) \subset \operatorname{Ran}(Q)$, i.e. $P \le Q$ as in Definition 5.17.

Lemma 5.19. *If $\mathcal{P}$ is a family of orthogonal projections on a Hilbert space H, then there exists unique orthogonal projections, $P_{\sup}$ and $P_{\inf}$, such that*

1. *$P_{\sup}$ is an upper bound for $\mathcal{P}$ and if Q is any other upper bound for $\mathcal{P}$ then $P_{\sup} \le Q$.*
2. *$P_{\inf}$ is a lower bound for $\mathcal{P}$ and if Q is any other lower bound for $\mathcal{P}$ then $Q \le P_{\inf}$.*

We write $P_{\sup} = \sup \mathcal{P}$ and $P_{\inf} = \inf \mathcal{P}$.

Proof. If Q is an upper bound for $\mathcal{P}$ (which exists, take $Q = I$) then $\operatorname{Ran}(P) \subset \operatorname{Ran}(Q)$ for all $P \in \mathcal{P}$ and hence

$$M_{\sup} := \overline{\sum_{P \in \mathcal{P}} \operatorname{Ran}(P)} \subset \operatorname{Ran}(Q).$$

It is now easy to verify that $P_{\sup}$ defined to be orthogonal projection onto $M_{\sup}$ is the desired least upper bound for $\mathcal{P}$.

If Q is an lower bound for $\mathcal{P}$ (which exists, take $Q \equiv 0$) then $\operatorname{Ran}(Q) \subset \operatorname{Ran}(P)$ for all $P \in \mathcal{P}$ and hence

$$\operatorname{Ran}(Q) \subset M_{\inf} := \cap_{P \in \mathcal{P}} \operatorname{Ran}(P).$$

It is now easy to verify that $P_{\inf}$ defined to be orthogonal projection onto $M_{\inf}$ is the desired greatest lower bound for $\mathcal{P}$. ∎

For the next result recall Lemma 9.29 which states; If $\mathcal{A}$ is a $*$ subalgebra of $B(H)$, K is a closed subspace of H, and P is the projection on K, then K is and $\mathcal{A}$ – invariant subspace iff $P \in \mathcal{A}'$

Lemma 5.20. *Let $\mathcal{P}$ be a family of orthogonal projections on a Hilbert space H. If $A \in \mathcal{P}'$, i.e. $[A, P] = 0$ for all $P \in \mathcal{P}'$ then $[A, \inf P] = 0 = [A, \sup \mathcal{P}]$.*

Proof. As $AP = PA$ for all $P \in \mathcal{P}$, by taking adjoints we also have $A^*P = PA^*$ for all $P \in \mathcal{P}$. From these equation it follows that

$$A \operatorname{Ran}(P) \subset \operatorname{Ran}(P) \text{ and } A^* \operatorname{Ran}(P) \subset \operatorname{Ran}(P) \ \forall \ P \in \mathcal{P}. \tag{5.16}$$

By Eq. (5.16),

$$\begin{aligned} A\left[\cap_{P \in \mathcal{P}} \operatorname{Ran}(P)\right] &\subset \left[\cap_{P \in \mathcal{P}} \operatorname{Ran}(P)\right] \text{ and} \\ A^*\left[\cap_{P \in \mathcal{P}} \operatorname{Ran}(P)\right] &\subset \left[\cap_{P \in \mathcal{P}} \operatorname{Ran}(P)\right] \end{aligned}$$

and therefore both A and A^* both preserve $\operatorname{Ran}(P_{\inf})$, i.e.

$$AP_{\inf} = P_{\inf} A P_{\inf} \text{ and } A^* P_{\inf} = P_{\inf} A^* P_{\inf}.$$

Taking adjoints of these equations also shows,

$$P_{\inf} A^* = P_{\inf} A^* P_{\inf} \text{ and } P_{\inf} A = P_{\inf} A P_{\inf}$$

and therefore $[A, P_{\inf}] = 0$.

Similarly by Eq. (5.16) we may conclude that

$$A \sum_{P \in \mathcal{P}} \operatorname{Ran}(P) \subset \sum_{P \in \mathcal{P}} \operatorname{Ran}(P) \text{ and } A^* \sum_{P \in \mathcal{P}} \operatorname{Ran}(P) \subset \sum_{P \in \mathcal{P}} \operatorname{Ran}(P)$$

and then by taking closures we learn that A and A^* both preserve $\operatorname{Ran}(P_{\sup})$. The same argument as above then shows $[A, P_{\sup}] = 0$. ∎

Lemma 5.21. *Let $\mathcal{P}$ be a family of orthogonal projections on a Hilbert space H.*

1. *If there exists and orthogonal projection Q such that $\langle Qv, v\rangle = \sup_{P \in \mathcal{P}} \langle Pv, v\rangle$ for all $v \in H$, then $Q = P_{\sup}$.*
2. *If there exists and orthogonal projection such Q such that $\langle Qv, v\rangle = \inf_{P \in \mathcal{P}} \langle Pv, v\rangle$ for all $v \in H$, then $Q = P_{\inf}$.*

Proof. Since $P \le P_{\sup}$ for all $P \in \mathcal{P}$, it follows by Remark 5.18 that

$$\langle Qv, v\rangle = \sup_{P \in \mathcal{P}} \langle Pv, v\rangle \le \langle P_{\sup} v, v\rangle \ \forall \ v \in H$$

which then implies $P \le Q \le P_{\sup}$ for all $P \in \mathcal{P}$ and hence $Q = P_{\sup}$.

Similarly, since $P_{\inf} \le P$ for all $P \in \mathcal{P}$, it follows by Remark 5.18 that

$$\langle Qv, v\rangle = \inf_{P \in \mathcal{P}} \langle Pv, v\rangle \ge \langle P_{\inf} v, v\rangle \ \forall \ v \in H$$

which then implies $P_{\inf} \le Q \le P$ for all $P \in \mathcal{P}$ and hence $Q = P_{\inf}$. ∎

Part III

Structure Theory of Commutative C^*-algebras

Throughout this part, $\mathcal{B}$ will be a complex unital commutative Banach algebra. So far we have been considering a single operator and its spectral properties and functional calculus. What we would like to do now is to simultaneously diagonalize a collection of commuting operators. The goals of this part are;

1. study the structure of $\mathcal{B}$,
2. show that when $\mathcal{B}$ is a C^*-algebra, that $\mathcal{B}$ is isomorphic to $C(X)$ for some compact Hausdorff space, X,
3. develop the continuous functional calculus for commutative C^*-algebras,
4. and simultaneously diagonalize all of the operators in commutative unital C^*-subalgebra of $B(H)$ where H is a separable Hilbert space.

The following two notions will play a key role in our discussions below.

Definition 5.22 (Characters and Spectrum). *A **character** of $\mathcal{B}$ is a **nonzero** multiplicative linear functional on $\mathcal{B}$, i.e. $\alpha : \mathcal{B} \to \mathbb{C}$ is an algebra homomorphism so in particular $\alpha(ab) = \alpha(a)\,\alpha(b)$.*

*The **spectrum** of $\mathcal{B}$ is the set $\widetilde{\mathcal{B}}$ (or denoted by $\operatorname{spec}(\mathcal{B})$)[2] of all characters of $\mathcal{B}$.*

Please note that we do not assume α to be bounded (i.e. continuous). However, as shown in Proposition 7.2 below the continuity is automatic. If $\alpha \in \widetilde{\mathcal{B}}$ then $\alpha(\mathbf{1}) = 1$ because $\alpha(\mathbf{1}^2) = \alpha(\mathbf{1})^2$ so $\alpha(\mathbf{1}) = 0$ or $\alpha(\mathbf{1}) = 1$. If $\alpha(\mathbf{1}) = 0$ then $\alpha \equiv 0$ so $\alpha(\mathbf{1}) = 1$. Given this information,

$$\tilde{\mathcal{B}} := \{\alpha \in \mathcal{B}^* : \alpha(\mathbf{1}) = 1 \text{ and } \alpha(AB) = \alpha(A)\,\alpha(B)\}\,. \tag{5.17}$$

For the next definition, let $\operatorname{Func}\left(\widetilde{\mathcal{B}} \to \mathbb{C}\right)$ denote the space of functions from $\tilde{\mathcal{B}}$ to $\mathbb{C}$.

Definition 5.23 (Gelfand Map). *For $a \in \mathcal{B}$ let $\hat{a} : \mathcal{B} \to \mathbb{C}$ be the function defined by $\widehat{a}(\alpha) = \alpha(a)$ for all $\alpha \in \tilde{\mathcal{B}}$. The map*

$$\mathcal{B} \ni a \to \widehat{a} \in \operatorname{Func}\left(\widetilde{\mathcal{B}} \to \mathbb{C}\right)$$

*is called the **canonical mapping** or **Gelfand mapping** of $\mathcal{B}$ into* $\operatorname{Func}\left(\widetilde{\mathcal{B}} \to \mathbb{C}\right)$. *[This definition will be refined in Definition 7.17 below.]*

Before getting down to business, we will pause to motivate the theory by first working in a finite dimensional linear algebra setting. This is the content of the first chapter of this part.

---

[2] We will see shortly that $\tilde{\mathcal{B}} \neq \emptyset$, see Lemmas 7.7 and 7.11.

Finite Dimensional Matrix Algebra Spectrum

For the purposes of this motivational chapter, let V be a finite dimensional inner product space and suppose that $\mathcal{B}$ is a unital commutative sub-algebra of $\mathrm{End}_{\mathbb{C}}(V)$.

6.1 Gelfand Theory Warm-up

Proposition 6.1. *If $\mathcal{B}$ is a commutative sub-algebra of $\mathrm{End}_{\mathbb{C}}(V)$ with $I \in \mathcal{B}$, then there exists $v \in V \setminus \{0\}$ which is a simultaneous eigenvector of B for all $B \in \mathcal{B}$. Moreover, there exists a character, $\alpha \in \tilde{\mathcal{B}}$, such that $Bv = \alpha(B) v$ for all $B \in \mathcal{B}$.*

Proof. Let $\{B_j\}_{j=1}^k$ be a basis for $\mathcal{B}$. Using the theory of characteristic polynomials along with the fact that $\mathbb{C}$ is algebraically closed, there exists $\lambda_1 \in \mathbb{C}$ which is an eigenvalue of B_1, i.e. $\mathrm{Nul}(B_1 - \lambda_1) \neq \{0\}$. Since $B_2 \mathrm{Nul}(B_1 - \lambda_1) \subset \mathrm{Nul}(B_1 - \lambda_1)$ it follows in the same way that there exists a $\lambda_2 \in \mathbb{C}$ so that $\mathrm{Nul}(B_2 - \lambda_2) \cap \mathrm{Nul}(B_1 - \lambda_1) \neq \{0\}$. Again one verifies that B_3 leaves the joint eigenspace, $\mathrm{Nul}(B_2 - \lambda_2) \cap \mathrm{Nul}(B_1 - \lambda_1)$, invariant and hence there exists $\lambda_3 \in \mathbb{C}$ such that

$$\mathrm{Nul}(B_3 - \lambda_3) \cap \mathrm{Nul}(B_2 - \lambda_2) \cap \mathrm{Nul}(B_1 - \lambda_1) \neq \{0\}.$$

Continuing this process inductively allows us to find $\{\lambda_j\}_{j=1}^k \subset \mathbb{C}$ so that $\cap_{j=1}^k \mathrm{Nul}(B_j - \lambda_j) \neq \{0\}$. Let v be a non-zero element of $\cap_{j=1}^k \mathrm{Nul}(B_j - \lambda_j)$. As the general element $B \in \mathcal{B}$ is of the form $B = \sum_{j=1}^k b_j B_j$, it follows that

$$Bv = \sum_{j=1}^k b_j B_j v = \left(\sum_{j=1}^k b_j \lambda_j\right) v \tag{6.1}$$

showing that v is a joint eigenvector for all $B \in \mathcal{B}$.

For the second assertion, let $\alpha : \mathcal{B} \to \mathbb{C}$ be defined by requiring $Bv = \alpha(B) v$ for all $B \in \mathcal{B}$. Then for $A, B \in \mathcal{B}$ and $\lambda \in \mathbb{C}$ we have $\alpha(I) v = Iv = v$,

$$\alpha(A + \lambda B) v = (A + \lambda B) v = Av + \lambda Bv = [\alpha(A) + \lambda \alpha(B)] v,$$

and

$$\alpha(AB) v = ABv = A[\alpha(B) v] = \alpha(B) Av = \alpha(B) \alpha(A) v$$

which altogether shows α is linear, multiplicative, and $\alpha(I) = 1$. ∎

Corollary 6.2. *For every $B \in \mathcal{B}$ and $\sigma(B) = \left\{\alpha(B) : \alpha \in \tilde{\mathcal{B}}\right\}$, where $\sigma(B) \subset \mathbb{C}$ is now precisely the set of eigenvalues of B.*

Proof. If $\alpha \in \tilde{\mathcal{B}}$ is given and $b := \alpha(B)$, then $\alpha(B - b) = 0$ which implies $B - bI$ can have no inverse in $\mathcal{B}$ which according to Lemma 3.9 implies that $B - bI$ has no inverse in $\mathrm{End}(V)$ and hence $b \in \sigma(B)$. Conversely, if $b \in \sigma(B)$ is given, in the proof of Proposition 6.1, choose $B_1 = B$ and and $\lambda_1 = b$. Then the proof of Proposition 6.1 produces a $\alpha \in \tilde{\mathcal{B}}$ so that $\alpha(B) = \lambda_1 = b$. ∎

Definition 6.3 (Joint Spectrum). *For $\{B_j\}_{j=1}^n \subset \mathcal{B}$, the set,*

$$\sigma(B_1, \dots, B_n) \subset \sigma(B_1) \times \dots \times \sigma(B_n) \subset \mathbb{C}^n,$$

defined by

$$\sigma(B_1, \dots, B_n) := \left\{(\alpha(B_1), \dots, \alpha(B_n)) : \alpha \in \tilde{\mathcal{B}}\right\}$$

*will be called the **joint spectrum** of $(B_1, \dots, B_n)$.*

Corollary 6.4. *Under the assumptions of this chapter, $\tilde{\mathcal{B}}$, is a non-empty finite set.*

Proof. Suppose that $\{B_j\}_{j=1}^k$ is a basis for $\mathcal{B}$ (or at least a generating set). Then the map,

$$\tilde{\mathcal{B}} \ni \alpha \to (\alpha(B_1), \dots, \alpha(B_n)) \in \sigma(B_1, \dots, B_n)$$

is easily seen to be a bijection. As $\sigma(B_1, \dots, B_n) \subset \sigma(B_1) \times \dots \times \sigma(B_n)$ and the latter set is a finite set, it follows that $\#\left(\tilde{\mathcal{B}}\right) < \infty$. The fact that $\tilde{\mathcal{B}}$ is not empty is the part of the content of Proposition 6.1. ∎

The general converse of the second assertion in Proposition 6.1 holds. The full proof of this Proposition is left to the appendix. Here we will prove an easier special case. Another, even slightly easier case (and all that we really need) of the next proposition may be found in Proposition 6.9 where we further restrict to $\mathcal{B}$ being a commutative C^*-subalgebra of $\mathrm{End}(V)$ where in that proposition V is an inner product space.

Proposition 6.5. *Let $\mathcal{B}$ be a commutative sub-algebra of $\operatorname{End}_{\mathbb{C}}(V)$ with $I \in \mathcal{B}$ and suppose that $\alpha : \mathcal{B} \to \mathbb{C}$ is a homomorphism. [We require $\alpha(I) = 1$.] Then there exists $v \in V \setminus \{0\}$ such that $Bv = \alpha(B) v$ for all $B \in \mathcal{B}$.*

Proof. We give a proof here under the added assumption that every element, $A \in \mathcal{B}$, is diagonalizable. Let $\{A_j\}_{j=1}^n$ be a subset of $\mathcal{B}$ which generates $\mathcal{B}$[1] and then let $a_j := \alpha(A_j)$ for $j \in [n]$. For each $j \in [n]$ let us define the polynomial,

$$p_j(z) := \prod_{\lambda \in \sigma(A_j) \setminus \{a_j\}} \frac{z - a_j}{\lambda - a_j}.$$

This polynomial has the property that $p_j(\lambda) = 0$ for all $\lambda \in \sigma(A_j) \setminus \{a_j\}$ and $p_j(a_j) = 1$. Since A_j is assumed to be diagonalizable we know that

$$V = \oplus_{\lambda \in \sigma(A_j)} \operatorname{Nul}(A_j - \lambda)$$

and (you prove) $p_j(A_j)$ is projection onto $\operatorname{Nul}(A_j - a_j)$ in this decomposition. Next let $Q := \prod_{j=1}^n p_j(A_j)$, order does not matter as $\mathcal{B}$ is commutative. Since

$$\alpha(Q) = \prod_{j=1}^n \alpha(p_j(A_j)) = \prod_{j=1}^n p_j(\alpha(A_j)) = \prod_{j=1}^n p_j(\alpha(a_j)) = 1,$$

we know $Q \neq 0$ and so there exists $w \in V$ so that $v := Qw \neq 0$. Again, since$\{p_j(A_j)\}_{j=1}^n$ all commute with one another it follows that $v \in \operatorname{Ran}(p_j(A_j)) = \operatorname{Nul}(A_j - a_j)$ for each $j \in [n]$ and this implies $A_j v = a_j v$ for all $j \in [n]$. Since the general element $A \in \mathcal{B}$ is of the form, $A = p(A_1, \dots, A_n)$, for some polynomial p, we conclude that

$$Av = p(A_1, \dots, A_n) v = p(a_1, \dots, a_n) v = \alpha(A) v.$$

■

Remark 6.6 (Joint Spectrum Characterization). Altogether Propositions 6.1 and 6.5 shows the following characterization of the joint spectrum from Definition 6.3. If $\mathcal{B} \subset \operatorname{End}(V)$ is a commutative sub-algebra generated by $\{B_j\}_{j=1}^k$, then

$$\sigma(B_1, \dots, B_n) = \left\{(\lambda_1, \dots, \lambda_k) \in \mathbb{C}^k : \cap_{j=1}^k \operatorname{Nul}(B_j - \lambda_j) \neq \{0\}\right\}.$$

That is $(\lambda_1, \dots, \lambda_k)$ is an element of $\sigma(B_1, \dots, B_n)$ iff there exist $v \in V \setminus \{0\}$ such that $B_j v = \lambda_j v$ for all $j \in [k]$.

[1] One could simply let $\{A_j\}_{j=1}^n$ be a basis for $\mathcal{B}$.

Lemma 6.7. *The Gelfand map in Definition 5.23 is an algebra homomorphism. The range, $\hat{\mathcal{B}} := \left\{\hat{B} : B \in \mathcal{B}\right\}$, is a sub-algebra which separates points but need not be closed under conjugation. The Gelfand map need not be injective.*

Proof. The homomorphism property is straightforward to verify, $\hat{I}(\alpha) = \alpha(I) = 1$ so that $\hat{I} = \mathbf{1}$,

$$\begin{aligned}(B_1 + \lambda B_2)^{\hat{}}(\alpha) &= \alpha(B_1 + \lambda B_2) = \alpha(B_1) + \lambda \alpha(B_2) \\ &= \left(\hat{B}_1 + \lambda \hat{B}_2\right)(\alpha)\end{aligned}$$

and

$$(B_1 B_2)^{\hat{}}(\alpha) = \alpha(B_1 B_2) = \alpha(B_1)\alpha(B_2) = \left(\hat{B}_1 \hat{B}_2\right)(\alpha).$$

If $\alpha_1 \neq \alpha_2$ are two distinct elements of $\tilde{\mathcal{B}}$ then by definition there exists $B \in \mathcal{B}$ so that $\alpha_1(B) \neq \alpha_2(B)$, i.e. $\hat{B}(\alpha_1) \neq \hat{B}(\alpha_2)$. This shows $\hat{\mathcal{B}}$ separates points.

Lastly if $\hat{B} \equiv 0$, then $0 = \hat{B}(\alpha) = \alpha(B)$ for all $\alpha \in \tilde{\mathcal{B}}$ which implies $\sigma(B) = \{0\}$ and hence B must be nilpotent. This certainly indicates that the Gelfand map need not be injective. For an explicit example, let

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

so that $A^2 = 0$ and hence

$$\mathcal{B} := \langle A \rangle = \operatorname{span}\{I, A\} = \left\{\begin{pmatrix} a & b \\ 0 & a \end{pmatrix} : a, b \in \mathbb{C}\right\}.$$

In this case we must have

$$\alpha\left(\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}\right) = a$$

and hence $\tilde{\mathcal{B}}$ consists of this single α. If

$$B := \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$$

then $\hat{B}(\alpha) = \alpha(B) = a$ and hence $\hat{B} \equiv 0$ iff $B = \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}$ for some $b \in \mathbb{C}$. The kernel of the Gelfand map is called the radical of $\mathcal{B}$ and in this case we have shown, $rad(\mathcal{B}) = \mathbb{C} \cdot A$. ■

6.2 Restricting to the C^*-case

In the this and then next section, we now restrict to the finite dimensional C^*-algebra setting.

Notation 6.8 *Let $(V = H, \langle\cdot,\cdot\rangle)$ is a complex finite dimensional inner product space that $\mathcal{B} \subset \operatorname{End}(H)$ is unital commutative $*$-algebra (i.e. $A \in \mathcal{B}$ implies $A^* \in \mathcal{B}$). Further (as $\dim \mathcal{B} < \infty$), let $\{B_j\}_{j=1}^k$ which is a basis for $\mathcal{B}$.*

We could construct $\mathcal{B}$ by choosing commuting normal operators, $\{B_j\}_{j=1}^k$, and then letting $\mathcal{B}$ be the C^*-subalgebra of $\operatorname{End}(H)$ generated by these operators. According to the Fuglede-Putnam Theorem 2.68, it is automatic that the collection of operators, $\left\{B_j, B_j^*\right\}_{j=1}^k$, all commute with one another and hence $\mathcal{B}$ consists of all elements of the form $p(B_1,\dots,B_k,B_1^*,\dots,B_k^*)$ where p is a polynomial in $2k$ -complex variables.

From Proposition 6.5 above, if $\alpha : \mathcal{B} \to \mathbb{C}$ is an algebra homomorphism, there exists a vector $v \in H \setminus \{0\}$ such that $Bv = \alpha(B)v$ for all $B \in \mathcal{B}$. Let us pause to give another proof of this statement in current C^*-algebra context.

Proposition 6.9. *Let $\mathcal{B}$ be a unital commutative $*$-subalgebra of $\operatorname{End}(H)$. If $\alpha \in \tilde{\mathcal{B}}$, there exists $v \in H \setminus \{0\}$ such that*

$$Bv = \alpha(B)v \text{ for all } B \in \mathcal{B}. \tag{6.2}$$

Proof. This is of course a special case of Proposition 6.5. Nevertheless, as the proof of this special case is a fair bit easier we will give another proof here. Moreover, the idea of this proof will be used again later.

Let $\{B_1,\dots,B_k\} \subset \mathcal{B}$ be a generating set for $\mathcal{B}$, let $\lambda_j := \alpha(B_j)$ for $1 \le j \le k$, and define

$$Q := \sum_{j=1}^k (B_j - \lambda_j)^* (B_j - \lambda_j) \in \mathcal{B}.$$

We then have

$$\begin{aligned}\alpha(Q) &= \sum_{j=1}^k \alpha\left((B_j - \lambda_j)^*(B_j - \lambda_j)\right) \\ &= \sum_{j=1}^k \alpha\left((B_j - \lambda_j)^*\right)\alpha(B_j - \lambda_j) = \sum_{j=1}^k |\alpha(B_j - \lambda_j)|^2 = 0.\end{aligned}$$

If Q^{-1} were to exist, Lemma 3.9 would imply that $Q^{-1} \in \mathcal{B}$ and therefore,

$$1 = \alpha(I) = \alpha\left(QQ^{-1}\right) = \alpha(Q)\,\alpha\left(Q^{-1}\right)$$

which would contradict the assertion that $\alpha(Q) = 0$. Thus we conclude Q is not invertible and therefore there exists $v \in H \setminus \{0\}$ so that $Qv = 0$. Since

$$\begin{aligned}0 = \langle 0, v\rangle = \langle Qv, v\rangle &= \sum_{j=1}^k \left\langle (B_j - \lambda_j)^*(B_j - \lambda_j)v, v\right\rangle \\ &= \sum_{j=1}^k \langle (B_j - \lambda_j)v, (B_j - \lambda_j)v\rangle = \sum_{j=1}^k \|(B_j - \lambda_j)v\|^2,\end{aligned}$$

it follows that $B_j v = \lambda_j v$ for all j. As the general element $B \in \mathcal{B}$ may be written as, $B = P(B_1,\dots,B_k)$ for some polynomial[2], P, it follows that

$$Bv = P(\lambda_1,\dots,\lambda_k)v = \alpha(P(B_1,\dots,B_k))v = \alpha(B)v \text{ for all } B \in \mathcal{B}.$$

■

Corollary 6.10. *If $\alpha \in \tilde{\mathcal{B}}$, then $\alpha : \mathcal{B} \to \mathbb{C}$ is a $*$-homomorphism, i.e. $\alpha(B^*) = \overline{\alpha(B)}$ for all $B \in \mathcal{B}$.*

Proof. We choose a unit vector $v \in H$ so that $Bv = \alpha(B)v$ for all $B \in \mathcal{B}$ in which case we have

$$\alpha(B) = \langle Bv, v\rangle \ \forall\ B \in \mathcal{B}. \tag{6.3}$$

It then follows from this equation that

$$\alpha(B^*) = \langle B^*v, v\rangle = \langle v, Bv\rangle = \overline{\langle Bv, v\rangle} = \overline{\alpha(B)}.$$

■

Proposition 6.11. *If X is a finite set and $\mathcal{A}$ is a sub-algebra of $C(X)$ that separates points and contains 1, then $\mathcal{A} = C(X)$. [We do not need to assume that $\mathcal{A}$ is closed under conjugation, this comes for free in this finite dimensional setting!.]*

Proof. By assumption for each $x, y \in X$ there exists $f \in \mathcal{A}$ so that $f(x) \neq f(y)$. We then let

$$f_y := \frac{1}{f(x) - f(y)}\left[f - f(y)1\right] \in \mathcal{A}$$

where now $f_y(x) = 1$ and $f_y(y) = 0$. Thus it follows that

$$\delta_x := \prod_{y \neq x} f_y \in \mathcal{A}$$

where $\delta_x(y) = 1_{x=y}$ for all $y \in X$. As $\{\delta_x\}_{x\in X}$ is a basis for $C(X)$, the proof is complete. ■

[2] If $\mathcal{B}$ were non-commutative, we would have to take P to be a non-commutative polynomial.

Theorem 6.12. *If $\mathcal{B}$ is a unital commutative C^*-subalgebra of* $\operatorname{End}(H)$, *then the Gelfand map,*

$$\mathcal{B} \ni A \to \hat{A} \in C\left(\tilde{\mathcal{B}}\right)$$

is an isometric C^*-isomorphism.

Proof. Let $\mathcal{A} := \left\{\hat{B} : B \in \mathcal{B}\right\} \subset C\left(\tilde{\mathcal{B}}\right)$ be the range of the Gelfand map. Then $\mathcal{A}$ is a sub-algebra of $C\left(\tilde{\mathcal{B}}\right)$ which contains 1. If α_1, α_2 are two points in $\tilde{\mathcal{B}}$ such that $\hat{B}(\alpha_1) = \hat{B}(\alpha_2)$ for all $B \in \mathcal{B}$ then

$$\alpha_1(B) = \hat{B}(\alpha_1) = \hat{B}(\alpha_2) = \alpha_2(B) \text{ for all } B \in \mathcal{B}$$

from which it follows that $\alpha_1 = \alpha_2$. This shows that $\mathcal{A}$ separates points and hence by the finite set version of the Stone-Wierstrass theorem, see Proposition 6.11, $\mathcal{A} = C\left(\tilde{\mathcal{B}}\right)$ and so the Gelfand map is surjective. Lastly

$$\begin{aligned}\|B\| &= r(B) = \max\{|\lambda| : \lambda \in \sigma(B)\} \\ &= \max\left\{|\alpha(B)| = \left|\hat{B}(\alpha)\right| : \alpha \in \tilde{\mathcal{B}}\right\} = \left\|\hat{B}\right\|_\infty,\end{aligned}$$

which shows the Gelfand map is isometric which of course implies that it is injective. ∎

The previous results illustrate well the key new result we are going to prove in the next chapter for general commutative C^*-algebras. The rest of this section is optional at this point.

6.3 Toward's Spectral Projections

Notation 6.13 *For each $\alpha \in \tilde{\mathcal{B}}$, let*

$$H_\alpha := \{v \in H : Bv = \alpha(B)v \text{ for all } B \in \mathcal{B}\}.$$

Lemma 6.14. *Let $\mathcal{B}$ be a commutative $*$-subalgebra of* $\operatorname{End}(H)$ *with unit. The inner product space, H, admits the orthogonal direct sum decomposition;*

$$H = \oplus_{\alpha \in \tilde{\mathcal{B}}} H_\alpha.$$

Proof. If α_1 and α_2 are distinct elements of $\tilde{\mathcal{B}}$, then there exists $B \in \mathcal{B}$ so that $\lambda_1 := \alpha_1(B) \neq \alpha_2(B) =: \lambda_2$. Thus if $v_j \in H_{\alpha_j}$, then

$$\begin{aligned}\lambda_1 \langle v_1, v_2\rangle &= \langle Bv_1, v_2\rangle = \langle v_1, B^* v_2\rangle = \langle v_1, \alpha_2(B^*) v_2\rangle \\ &= \langle v_1, \bar{\lambda}_2 v_2\rangle = \lambda_2 \langle v_1, v_2\rangle\end{aligned}$$

from which it follows that $\langle v_1, v_2\rangle = 0$. This shows $H_{\alpha_1} \perp H_{\alpha_2}$.

Let $H_0 := \oplus_{\alpha \in \tilde{\mathcal{B}}} H_\alpha$ and $H_1 := H_0^\perp$. Since H_0 is a $\mathcal{B}$-invariant subset of H it follows that H_1 is also $\mathcal{B}$-invariant. Indeed,

$$\langle BH_1, H_0\rangle = \langle H_1, B^* H_0\rangle \subset \langle H_1, H_0\rangle = \{0\} \text{ for all } B \in \mathcal{B}.$$

If $H_1 \neq \{0\}$, we may restrict $\mathcal{B}$ to H_1 and use Proposition 6.1 to find[3] a simultaneous eigenvector $v_1 \in H_1 \setminus \{0\}$ of $\mathcal{B}$. Associated to this vector is the character, α, of $\mathcal{B}$ such that $Bv_1 = \alpha(B) v_1$ for all $B \in \mathcal{B}$. But this then leads to the contradiction that $v_1 \in H_\alpha \subset H_0$. ∎

Notation 6.15 *For each $\alpha \in \tilde{\mathcal{B}}$, let $P_\alpha : H \to H$ be orthogonal projection onto H_α.*

For $v \in H$, we have, with $v_\alpha = P_\alpha v$, that $v = \sum_{\alpha \in \tilde{\mathcal{B}}} v_\alpha$ and so for $B \in \mathcal{B}$,

$$Bv = \sum_{\alpha \in \tilde{\mathcal{B}}} Bv_\alpha = \sum_{\alpha \in \tilde{\mathcal{B}}} \alpha(B) v_\alpha = \sum_{\alpha \in \tilde{\mathcal{B}}} \alpha(B) P_\alpha v.$$

Thus we have shown that

$$B = \sum_{\alpha \in \tilde{\mathcal{B}}} \alpha(B) P_\alpha \text{ for all } B \in \mathcal{B}. \tag{6.4}$$

Corollary 6.16. *For each $\alpha \in \tilde{\mathcal{B}}$, $P_\alpha \in \mathcal{B}$.*

Proof. Let $Q_\alpha \in \tilde{\mathcal{B}}$ be the unique element such $\hat{Q}_\alpha = \delta_\alpha$, i.e. $\alpha'(Q_\alpha) = 1_{\alpha = \alpha'}$. Then by Eq. (6.4)

$$Q_\alpha = \sum_{\alpha' \in \tilde{\mathcal{B}}} \alpha'(Q_\alpha) P_{\alpha'} = P_\alpha.$$

∎

Corollary 6.17. *For $f \in C\left(\tilde{\mathcal{B}}\right)$, let*

$$f^\vee = \sum_{\alpha \in \tilde{\mathcal{B}}} f(\alpha) \cdot P_\alpha \in \mathcal{B}.$$

Then $C\left(\tilde{\mathcal{B}}\right) \ni f \to f^\vee \in \mathcal{B}$ is the inverse to the Gelfand map.

Proof. We have

$$\left(f^\vee\right)\hat{} = \sum_{\alpha \in \tilde{\mathcal{B}}} f(\alpha) \cdot \hat{P}_\alpha = \sum_{\alpha \in \tilde{\mathcal{B}}} f(\alpha) \cdot \delta_\alpha = f.$$

∎

[3] Here is where we use the assumption that $\mathcal{B}$ is commutative.

6.4 Appendix: Full Proof of Proposition 6.5

Proof of Proposition 6.5. Suppose that $\mathcal{B}$ is generated by $\{A_j\}_{j=1}^k$ and let $a_j = \alpha(A_j)$. Let us further choose an $N \in \mathbb{N}$ sufficiently large so that

$$\operatorname{Nul}\left([A_j - \lambda]^{N+1}\right) = \operatorname{Nul}\left([A_j - \lambda]^N\right) \ \forall\, 1 \le j \le k \text{ and } \lambda \in \sigma(A_j).$$

Thus $\operatorname{Nul}\left([A_j - \lambda]^N\right)$ is the generalized λ-eigenspace of A_j for each j and $\lambda \in \sigma(A_j)$ and recall that

$$V = \oplus_{\lambda \in \sigma(A_j)} \operatorname{Nul}\left([A_j - \lambda]^N\right)$$

for each j. We then let

$$p_j(z) = \prod_{\lambda \in \sigma(A_j) \setminus \{a_j\}} \left(\frac{z - \lambda}{a_j - \lambda}\right)^N$$

so that $\alpha(p_j(A_j)) = p_j(a_j) = 1$, $p_j(A_j)$ annihilates $\operatorname{Nul}\left([A_j - \lambda]^N\right)$ for ever $\lambda \in \sigma(A_j) \setminus \{a_j\}$, and

$$\operatorname{Ran}[p_j(A_j)] = p_j(A_j) \operatorname{Nul}\left([A_j - a_j]^N\right) \subset \operatorname{Nul}\left([A_j - a_j]^N\right)$$

for each j. From this it follows that $1 = \alpha\left(\prod_{j=1}^k p_j(A_j)\right)$ and hence there exists $v \neq 0$ in V so that

$$\prod_{j=1}^k p_j(A_j)\, v = v. \tag{6.5}$$

As the $\{p_j(A_j)\}_{j=1}^k$ commute along with the above remarks we learn that $v \in \cap_{j=1}^k \operatorname{Nul}\left((A_j - a_j)^N\right)$. We now have to modify v a bit to produce an non-zero element of $\cap_{j=1}^k \operatorname{Nul}(A_j - a_j)$ which suffices to complete the proof of the proposition.

Start by choosing $0 \le \ell_1 < N$ so that

$$v_1 = (A_1 - a_1)^{\ell_1} v \in \operatorname{Nul}(A_1 - a_1) \setminus \{0\}.$$

Applying $(A_1 - a_1)^{\ell_1}$ to Eq. (6.5) (while using $p_1(A_1) v_1 = p_1(a) v_1 = v_1$) shows,

$$\prod_{j=2}^k p_j(A_j)\, v_1 = v_1. \tag{6.6}$$

Next we choose ℓ_2 so that

$$v_2 = (A_2 - a_2)^{\ell_2} v_1 \in \operatorname{Nul}(A_2 - a_2) \setminus \{0\}.$$

Applying $(A_2 - a_2)^{\ell_2}$ to Eq. (6.6) shows,

$$\prod_{j=3}^k p_j(A_j)\, v_2 = v_2. \tag{6.7}$$

Let us note that $A_2 v_2 = a_2 v_2$ and

$$A_1 v_2 = A_1 (A_2 - a_2)^{\ell_2} v_1 = (A_2 - a_2)^{\ell_2} A_1 v_1 = a_1 (A_2 - a_2)^{\ell_2} v_1 = a_1 v_2$$

and so

$$v_2 \in \operatorname{Nul}(A_2 - a_2) \cap \operatorname{Nul}(A_1 - a_1) \cap \left[\cap_{j=3}^k \operatorname{Nul}\left([A_j - \lambda]^N\right)\right].$$

Again choosing $0 \le \ell_3 < N$ so that

$$v_3 = (A_3 - a_3)^{\ell_3} v_2 \in \operatorname{Nul}(A_3 - a_3) \setminus \{0\}.$$

Applying $(A_3 - a_3)^{\ell_3}$ to Eq. (6.6) shows,

$$\prod_{j=4}^k p_j(A_j)\, v_3 = v_3. \tag{6.8}$$

Working as above it not follows that $v_3 \neq 0$ and

$$v_3 \in \left[\cap_{j=1}^3 \operatorname{Nul}(A_j - a_j)\right] \cap \left[\cap_{j=4}^k \operatorname{Nul}\left([A_j - \lambda]^N\right)\right].$$

Continuing this way inductively eventually produces $0 \neq v_k \in \cap_{j=1}^k \operatorname{Nul}(A_j - a_j)$.

■

6.5 *Appendix: Why not characters for non-commutative $\mathcal{B}$

Question: why don't we use characters when $\mathcal{B}$ is non-commutative?

Answer: they may vary well not exists. For example if $\mathcal{B}$ is all 2×2 matrices and α is a character, then $\alpha([A, B]) = 0$ for all $A, B \in \mathcal{B}$. When

$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \text{and } A' = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

we find

$$[A,B] = \begin{bmatrix} -b & 0 \\ a-d & b \end{bmatrix} \text{ and } [A',B] = \begin{bmatrix} c & d-a \\ 0 & -c \end{bmatrix}.$$

Taking $b = 0$ in the first case and $c = 0$ in the second case we see that $\operatorname{span}\{[A,B] : A, B \in \mathcal{B}\}$ contains A, A', and then it also follows that it contains

$$\begin{bmatrix} c & 0 \\ 0 & -c \end{bmatrix} \text{ for all } c \in \mathbb{C}.$$

In other words, $\operatorname{span}\{[A,B] : A, B \in \mathcal{B}\}$ is precisely the set of trace free matrices. Thus it follows that $\alpha(A) = 0$ whenever $\operatorname{tr} A = 0$. For a general matrix, A, we then have $A - \frac{1}{2}\operatorname{tr}(A) I$ is trace free and therefore,

$$0 = \alpha\left(A - \frac{1}{2}\operatorname{tr}(A) I\right) = \alpha(A) - \frac{1}{2}\operatorname{tr}(A).$$

Thus the only possible choice for α is $\alpha(A) = \frac{1}{2}\operatorname{tr}(A)$. However, this functional is not multiplicative.

A point to keep in mind below. When $\mathcal{B}$ is a non-commutative Banach algebra and if $M \subset \mathcal{B}$ is a proper two sided-ideal, then M can not contain any element, $b \in \mathcal{B}$, which have either a right or a left inverse. Whereas when $\mathcal{B}$ is commutative, this condition reduces to the statement that M can not contain any invertible elements, i.e. $\mathcal{B} \subset \mathcal{S}$ where $\mathcal{S}$ is the collections of non-invertible elements. In particular if we are expecting to use characters to find the spectrum of operators, $b \in \mathcal{B}$, as $\{\alpha(b) : \alpha \text{ runs through characters of } \mathcal{B}\}$ we are going to be sorely disappointed as we see even in the finite 2×2 matrix algebra.

As another such example, let $S : \ell^2 \to \ell^2$ be the shift operator and S^* be it's adjoint;,

$$S(x_1, x_2, \dots) = (0, x_1, x_2, \dots), \text{ and } S^*(x_1, x_2, \dots) = (x_2, x_3, \dots)$$

and suppose that α is a character on some algebra containing $\{S, S^*\}$. Since $SS^* = I \neq S^*S$, it follows that $1 = \alpha(I) = \alpha(S)\alpha(S^*)$ even though neither S nor S^* are invertible.

Commutative Banach Algebras with Identity

Henceforth $\mathcal{B}$ will denote a **unital commutative** Banach algebra over $\mathbb{C}$. (A good reference is Vol II of Dunford and Schwartz.) Recall from Definition 8.1 that a $\operatorname{spec}(\mathcal{B}) = \tilde{\mathcal{B}}$ is the set of characters, $\alpha : \mathcal{B} \to \mathbb{C}$, where α is a character if it is, non-zero, linear, and multiplicative. [See Corollary 2.63 for more motivation for the terminology.]

7.1 General Commutative Banach Algebra Spectral Properties

Lemma 7.1. *If $\alpha \in \tilde{\mathcal{B}} = \operatorname{spec}(\mathcal{B})$, then $\alpha(a) \in \sigma(a)$ for all $a \in \mathcal{B}$.*

Proof. Let $\lambda = \alpha(a)$ and $b = a - \lambda 1$ so that $\alpha(b) = 0$. If b^{-1} existed in $\mathcal{B}$ we would have

$$1 = \alpha(1) = \alpha\left(b^{-1}b\right) = \alpha\left(b^{-1}\right)\alpha(b)$$

which would imply $\alpha(b) \neq 0$. Thus b is not invertible and hence $\lambda = \alpha(a) \in \sigma(a)$. ■

Proposition 7.2 (Continuity of characters). *Every character α of $\mathcal{B}$ is continuous and moreover $\|\alpha\| \leq 1$ with equality if $\|1\| = 1$ which we always assume here.*

Proof. By Lemma 7.1, $\alpha(a) \in \sigma(a)$ for all $a \in \mathcal{B}$ and therefore,

$$|\alpha(a)| \leq r(a) \leq \|a\|.$$

■

Definition 7.3 (Maximal Ideals). *An ideal $J \subset \mathcal{B}$ is a **maximal ideal** if $J \neq \mathcal{B}$ and there is no proper ideal in $\mathcal{B}$ containing J.*

Example 7.4. If $\alpha \in \tilde{\mathcal{B}}$, then $J_\alpha := \operatorname{Nul}(\alpha)$ is a maximal ideal. Indeed, it is easily verified that is proper ideal. To see that it is maximal, suppose that $b \in \mathcal{B} \setminus J_\alpha$ and let $\lambda = \alpha(b)$ so that $\alpha(b - \lambda) = 0$, i.e. $b - \lambda \in J_\alpha$. This shows that $\mathcal{B} = J_\alpha \oplus \mathbb{C}\mathbf{1}$ and therefore J_α is maximal.

Notation 7.5 *Let $\mathcal{S} := \mathcal{B} \setminus \mathcal{B}_{inv}$ be the **singular elements** of $\mathcal{B}$. [Notice that $\mathcal{S}$ is a closed subset of $\mathcal{B}$.]*

Lemma 7.6. *If J is a proper ideal of $\mathcal{B}$, then $J \subset \mathcal{B} \setminus \mathcal{B}_{inv}$. Moreover, the closure $\left(\bar{J}\right)$ of J is also a proper ideal of $\mathcal{B}$. In particular if $J \subset \mathcal{B}$ is a maximal ideal, then J is necessarily closed.*

Proof. If J is any ideal in $\mathcal{B}$ that contains an element, b, of $\mathcal{B}_{inv}$, then J contains $b^{-1}b = 1$ and hence $J = \mathcal{B}$. Thus if $J \subsetneq \mathcal{B}$ is any proper ideal then $J \subset \mathcal{B} \backslash \mathcal{B}_{inv}$. As $\mathcal{B} \backslash \mathcal{B}_{inv}$ is a closed, $\bar{J} \subset \mathcal{B} \backslash \mathcal{B}_{inv}$. Moreover if $b = \lim_{n\to\infty} b_n \in \bar{J}$ with $b_n \in J$ and $x \in \mathcal{B}$, then $xb = \lim_{n\to\infty} xb_n \in \bar{J}$ as $xb_n \in J$ for all n. Lastly if J is a maximal ideal, then $J \subset \bar{J} \subsetneq \mathcal{B}$ and hence by maximality of J we have $J = \bar{J}$. ■

Lemma 7.7. *If $\mathcal{B}$ is a commutative Banach algebra with identity, then;*

1. *Every proper ideal $J_0 \subset \mathcal{B}$ is contained in a (not necessarily unique) maximal ideal.*
2. *An element $a \in \mathcal{B}$ is invertible iff a does not belong to any maximal ideal. In other words,*

$$\mathcal{S} := \mathcal{B} \setminus \mathcal{B}_{inv} = \cup\left(\text{maximal ideals}\right). \tag{7.1}$$

Proof. We take each item in turn.

1. Let $\mathcal{F}$ denote the collection of proper ideals of $\mathcal{B}$ which contain J_0. Order $\mathcal{F}$ by set inclusion and notice that if $\{J_\alpha\}_{\alpha\in A}$ is a totally ordered subset of $\mathcal{F}$ then $J := \cup_{\alpha\in A} J_\alpha \subset \mathcal{S}$ is a proper ideal ($1 \notin J_\alpha$ for all α) containing J_0, i.e. $J \in \mathcal{F}$. So by Zorn's Lemma, $\mathcal{F}$ contains a maximal element J which is the desired maximal ideal.
2. If $a \in \mathcal{S}$, then the ideal, (a), generated by a is a proper ideal for otherwise $1 \in (a)$ and there would exists $b \in \mathcal{B}$ such that $ba = 1$, i.e. a^{-1} would exist. By item 1. we can find a maximal ideal, J, which contains (a) and hence a. Conversely if a is in some maximal ideal, J, then a^{-1} can not exists since otherwise $1 = a^{-1}a \in J$. This verifies the identity in Eq. (7.1).

■

As is well known from basic algebra, the point of ideals are that they are precisely the subspaces which are the possible null spaces of algebra isomorphisms.

Exercise 7.1. Suppose $\mathcal{B}$ is a Banach algebra (not necessarily commutative) and $K \subset \mathcal{B}$ is a closed proper two sided ideal in $\mathcal{B}$. Show (items 1. and 3. are the most important);

1. $\mathcal{B}/K$ is a Banach algebra.
2. The bijection of closed subspaces in the factor Theorem A.24 given by,
$$\left\{\begin{array}{c}\text{closed subspaces}\\ \text{of } \mathcal{B} \text{ containing } K\end{array}\right\} \ni N \to \pi(N) \in \left\{\begin{array}{c}\text{closed subspaces}\\ \text{of } \pi(\mathcal{B}/K)\end{array}\right\},$$
restricts to a bijection of two sided closed ideals in $\mathcal{B}$ containing K to two sided closed ideals in $\mathcal{B}/K$.
3. If $\|\mathbf{1}\|_{\mathcal{B}} = 1$, then $\|\pi(\mathbf{1})\|_{\mathcal{B}/K} = 1$.

Proposition 7.8. *If $J \subset \mathcal{B}$ is a maximal ideal, then $\mathcal{B} = J \oplus \mathbb{C}\mathbf{1}$, where $\mathbf{1} = 1_{\mathcal{B}}$.*

Proof. Let $a \in \mathcal{B}$ and $\bar{a} := \pi(a) \in \mathcal{B}/J$ and $\lambda \in \sigma(\bar{a})$ and set $b := a - \lambda$. Then $\bar{b} = \pi(b) = \bar{a} - \lambda$ is not invertible in $\mathcal{B}/J$ and therefore $(\bar{b})$ is a proper ideal in $\mathcal{B}/J$. If $\bar{b} \neq 0$, then $\pi^{-1}((\bar{b}))$ would be a proper ideal in $\mathcal{B}$ which was strictly bigger than J contradicting the maximality of J. Therefore we conclude $0 = \bar{b} = \pi(b) = \pi(a - \lambda)$ which implies $a - \lambda \in J$. Thus we have shown $a = \lambda 1 \bmod J$, i.e. $\mathcal{B} = J + \mathbb{C}\mathbf{1}$. Since $\mathbf{1} \notin J$ as J is a proper ideal the proof is complete.

■

The next two result are optional at this point and the reader may safely skip to Lemma 7.11.

Theorem 7.9 (Gelfand – Mazur). *If $\mathcal{A}$ is a complex Banach algebra ($\mathcal{A}$) with unit which is a division algebra*[1]*, then $\mathcal{A}$ is isomorphic to $\mathbb{C}$. In more detail we have $\mathcal{A} = \mathbb{C} \cdot 1_{\mathcal{A}}$.*

Proof. Let $x \in \mathcal{A}$ and $\lambda \in \sigma(x)$. Then $x - \lambda 1$ is not invertible. Thus $x - \lambda 1 = 0$ so $x = \lambda 1$. Therefore every element of $\mathcal{A}$ is a complex multiple of 1, i.e. $\mathcal{A} = \mathbb{C} \cdot 1$. ■

Proposition 7.10 (Optional). *If $\mathcal{B}$ is a commutative Banach algebra with identity, then;*

1. *If $\{0\}$ is the only proper ideal in $\mathcal{B}$ then $\mathcal{B} = \mathbb{C} \cdot \mathbf{1}$.*
2. *If J is a maximal ideal in $\mathcal{B}$ then $\mathcal{B}/J = \mathbb{C} \cdot 1_{\mathcal{B}/J}$ is a field.*

Proof. 1. If $a \in \mathcal{B}$ let (a) denote the ideal generated by a. If $a \neq 0$ we must have $(a) = \mathcal{B}$ and in particular a must be invertible. Moreover, because we are working over $\mathbb{C}$, $\mathcal{B} = \mathbb{C} \cdot \mathbf{1}$ by the Gelfand – Mazur Theorem 7.9.

[1] Recall that $\mathcal{A}$ is a division algebra iff every non-zero element is invertible.

2. Since the ideals of $\mathcal{B}/J$ are in one to one correspondence with ideals $J \subset \mathcal{B}$ such that $J \subset J$, it follows that J is a maximal ideal in $\mathcal{B}$ iff (0) is the only proper ideal in $\mathcal{B}/J$. The result now follows from item 1. ■

Lemma 7.11. *The map*
$$\tilde{\mathcal{B}} \ni \alpha \to \mathrm{Nul}(\alpha) \in \{\textit{maximal ideals in } \mathcal{B}\}$$
is a bijection. In particular, $\tilde{\mathcal{B}} \neq \emptyset$ because of Lemma 7.7.

Proof. If α is a character then $\mathrm{Nul}(\alpha)$ is a maximal ideal of $\mathcal{B}$ by Example 7.4. Conversely if $J \subset \mathcal{B}$ is a maximal ideal, then by Proposition 7.8, $\mathcal{B} = \mathbb{C}\cdot\mathbf{1} \oplus J$ and we may define $\alpha : \mathcal{B} \to \mathbb{C}$ by
$$\alpha(\lambda\mathbf{1} + a) = \lambda \ \forall\ \lambda \in \mathbb{C} \text{ and } a \in J.$$
It is now easily verified that $\alpha \in \tilde{\mathcal{B}}$ and clearly we have $\mathrm{Nul}(\alpha) = J$.

Finally if $\alpha, \beta \in \tilde{\mathcal{B}}$ and $J = \mathrm{Nul}\,\alpha = \mathrm{Nul}\,\beta$, then for $\lambda \in \mathbb{C}$ and $a \in J$ we must have,
$$\alpha(\lambda\mathbf{1} + a) = \lambda = \beta(\lambda\mathbf{1} + a)$$
which shows (as $\mathcal{B} = \mathbb{C} \cdot \mathbf{1} \oplus J$) that $\alpha = \beta$. ■

Corollary 7.12. *If $\mathcal{B}$ is a commutative Banach algebra with identity, then $b \in \mathcal{S} := \mathcal{B} \setminus \mathcal{B}_{inv}$ iff $0 \in \sigma(b)$ iff there exists $\alpha \in \tilde{\mathcal{B}}$ such that $\alpha(b) = 0$. More generally,*
$$\sigma(a) = \left\{\alpha(a) : \alpha \in \tilde{\mathcal{B}}\right\}.$$

Proof. The first assertion follows from Lemmas 7.7 and 7.11. It can also be seen by Proposition 8.3 below. For the second we have $\lambda \in \sigma(a)$ iff $b = a - \lambda \in \mathcal{S}$ iff $0 = \alpha(b) = \alpha(a) - \lambda$ for some $\alpha \in \tilde{\mathcal{B}}$. ■

Notation 7.13 *Because of Lemma 7.11, $\widetilde{\mathcal{B}}$ is sometimes referred to as the **maximal ideal space** of $\mathcal{B}$.*

Let us recall Alaoglu's Theorem **??**.

Theorem 7.14 (Alaoglu's Theorem). *If X is a normed space the closed unit ball,*
$$C^* := \{f \in X^* : \|f\| \leq 1\} \subset X^*,$$
is weak-$*$ compact. [Recall that the weak-$*$ topology is the smallest topology on X^* such that $\pi_x = \hat{x} : X^* \to \mathbb{C}$ is continuous for all $x \in X$, where $\hat{x}(\ell) = \ell(x)$, see Definition A.18.]

Corollary 7.15 ($\tilde{\mathcal{B}}$ is a compact Hausdorff space). *$\widetilde{\mathcal{B}}$ is a w^*–closed subset of the unit ball in $\mathcal{B}^*$. In particular, $\widetilde{\mathcal{B}}$ is a compact Hausdorff space in the w^* – topology. [Here w^* is short for weak-$*$.]*

Proof. Since $\{\alpha \in \mathcal{B}^* : \alpha(ab) = \alpha(a)\alpha(b)\}$ for $a, b \in \mathcal{B}$ fixed and $\{\alpha \in \mathcal{B}^* : \alpha(\mathbf{1}) = 1\}$ are closed in the w^*–topology,

$$\widetilde{\mathcal{B}} = \{\alpha \in \mathcal{B}^* : \alpha(\mathbf{1}) = 1\} \cap \bigcap_{a,b\in\mathcal{B}} \{\alpha \in \mathcal{B}^* : \alpha(ab) = \alpha(a)\alpha(b)\}$$

is w^* – closed – being the intersection of closed sets. Since $\tilde{\mathcal{B}}$ is a closed subset of a compact Hausdorff space (namely the unit ball in $\mathcal{B}^*$ with the w^* – topology), $\tilde{\mathcal{B}}$ is a compact Hausdorff space as well. ∎

Remark 7.16. If $\mathcal{B}$ is a commutative Banach algebra *without* identity and we define a character as a continuous nonzero homomorphism $\alpha : \mathcal{B} \to \mathbb{C}$. Then the preceding arguments shows that $\widetilde{\mathcal{B}} \subset$ (unit ball of $\mathcal{B}^*$) but may not be closed because 0 is a limit point of $\widetilde{\mathcal{B}}$. In this case $\widetilde{\mathcal{B}}$ is locally compact.

We now recall and refine the definition of the Gelfand map given in Definition 5.23.

Definition 7.17 (Gelfand Map). *For $a \in \mathcal{B}$, let $\hat{a} \in C\left(\widetilde{\mathcal{B}}\right)$ be the function defined by $\widehat{a}(\alpha) = \alpha(a)$ for all $\alpha \in \tilde{\mathcal{B}}$. The map*

$$\mathcal{B} \ni a \to \widehat{a} \in C\left(\widetilde{\mathcal{B}}\right)$$

*is called the **canonical mapping** or **Gelfand mapping** of $\mathcal{B}$ into $C\left(\widetilde{\mathcal{B}}\right)$.*

Proposition 7.18. *If $\mathcal{B}$ is a commutative Banach algebra with identity, then*

1. *$\widehat{\mathbf{1}}$ is the constant function 1 in $1 \in C\left(\tilde{\mathcal{B}}\right)$.*
2. *For $a \in \mathcal{B}$,*
$$\sigma(a) = \mathrm{Ran}(\hat{a}) = \{\alpha(a) : \alpha \in \tilde{\mathcal{B}}\}$$
3. *The spectral mapping Theorem 3.42 is a consequence of the previous assertion.*
4. *The spectral radius of $a \in \mathcal{B}$ satisfies (compare with Exercise 3.5),*
$$r(a) = \|\hat{a}\|_\infty \le \|a\|, \ r(a+b) \le r(a) + r(b), \ \text{and } r(ab) \le r(a)\, r(b).$$

Proof. We take each item in turn.

1. $\widehat{\mathbf{1}}(\alpha) = \alpha(\mathbf{1}) = 1$ for all $\alpha \in \widetilde{\mathcal{B}}$, so $\hat{1}$ is the constant function $1 \in C\left(\tilde{\mathcal{B}}\right)$.
2. This was proved in Corollary 7.12.
3. If $p \in \mathbb{C}[z]$ is a polynomial, $a \in \mathcal{B}$, and $\alpha \in \tilde{\mathcal{B}}$, then
$$\widehat{p(a)}(\alpha) = \alpha(p(a)) = p(\alpha(a)) = p(\hat{a}(\alpha)) = (p \circ \hat{a})(\alpha)$$
and therefore
$$\sigma(p(a)) = \mathrm{Ran}(\widehat{p(a)}) = \mathrm{Ran}(p \circ \hat{a}) = p(\mathrm{Ran}(\hat{a})) = p(\sigma(a)).$$
4. This is an easy direct consequence of the spectral mapping theorem of item 3. Indeed we always know $r(a) \le \|a\|$ and
$$\begin{aligned} r(a) &= \sup\{|\lambda| : \lambda \in \sigma(a)\} = \sup\left\{|\alpha(a)| : \alpha \in \tilde{\mathcal{B}}\right\} \\ &= \sup\left\{|\hat{a}(\alpha)| : \alpha \in \tilde{\mathcal{B}}\right\} = \|\hat{a}\|_\infty. \end{aligned}$$
The remaining inequalities are now easily proved as follows;
$$r(ab) = \left\|\widehat{ab}\right\|_\infty = \left\|\hat{a} \cdot \hat{b}\right\|_\infty \le \|\hat{a}\|_\infty \left\|\hat{b}\right\|_\infty = r(a)\, r(b)$$
and similarly,
$$\begin{aligned} r(a+b) &= \left\|\widehat{a+b}\right\|_\infty = \left\|\hat{a} + \hat{b}\right\|_\infty \\ &\le \|\hat{a}\|_\infty + \left\|\hat{b}\right\|_\infty = r(a) + r(b). \end{aligned}$$

∎

For more on the general Gelfand-homomorphism theory, see the optional Appendix 8.1 below. For our immediate purposes we are going to now restrict to the C^*-algebra setting.

7.2 Commutative C^*-algebras

For this section, $\mathcal{B}$ is a commutative C^*-algebra[2] with identity.

Lemma 7.19. *If $\alpha \in \tilde{\mathcal{B}}$, then $\alpha(b^*) = \overline{\alpha(b)}$ for all $b \in \mathcal{B}$. Equivalently, the Gelfand homomorphism is a $*$-homomorphism, i.e. $\widehat{b^*} = \bar{\hat{b}}$ for all $b \in \mathcal{B}$.*

Proof. If $b \in \mathcal{B}$ is decomposed as $b = x + iy$ with x, y are Hermitian, then

$$\alpha(b) = \alpha(x) + i\alpha(y)$$

[2] Recall the C^* – definition requires that $\|a^*a\| = \|a\|^2$ for all $a \in \mathcal{B}$, see Definition 2.50.

where $\alpha(x) \in \sigma(x)$ and $\alpha(y) \in \sigma(y)$. By Lemma 4.5 we know that $\sigma(x)$ and $\sigma(y)$ are contained in $\mathbb{R}$ and therefore,

$$\alpha(b^*) = \alpha(x - iy) = \alpha(x) - i\alpha(y) = \overline{\alpha(b)}.$$

■

Remark 7.20 (Second proof of Lemma 4.5). If $a \in \mathcal{B}$ is Hermitian, then (by Example 2.66) $u_t := e^{ita}$ is unitary for all $t \in \mathbb{R}$ and so by Lemma 2.65, $\|u_t\| = 1$ for all $t \in \mathbb{R}$. Now for $\alpha \in \tilde{\mathcal{B}}$, $\alpha(u_t) = e^{it\alpha(a)}$ and hence

$$e^{-t\operatorname{Im}[\alpha(a)]} = \left|e^{it\alpha(a)}\right| = |\alpha(u_t)| \le \|\alpha\| \|u_t\| = 1 \ \forall\ t \in \mathbb{R}.$$

This last inequality can only hold if $\operatorname{Im}\alpha(a) = 0$ and hence $\sigma(a) = \left\{\alpha(a) : \alpha \in \tilde{\mathcal{B}}\right\} \subset \mathbb{R}$.

Let us recall the Stone-Weierstrass theorem.

Theorem 7.21 (Complex Stone-Weierstrass Theorem). *Let X be a locally compact Hausdorff space. Suppose $\mathcal{A}$ is a subalgebra of $C_0(X, \mathbb{C})$ which is closed in the uniform topology, separates points, and is closed under complex conjugation. Then either $\mathcal{A} = C_0(X, \mathbb{C})$ (which happens if $\mathbf{1} \in \mathcal{A}$) or*

$$\mathcal{A} = \mathcal{I}_{x_0}^{\mathbb{C}} := \{f \in C_0(X, \mathbb{C}) : f(x_0) = 0\}$$

for some $x_0 \in X$.

Theorem 7.22 (Commutative C^*-algebra classification). *If $\mathcal{B}$ is a commutative C^*-algebra with identity, then the Gelfand map,*

$$\mathcal{B} \ni b \to \hat{b} \in C\left(\tilde{\mathcal{B}}\right),$$

*is an isometric $*-$isomorphism **onto** $C\left(\tilde{\mathcal{B}}\right)$.*

Proof. Since, for $b \in \mathcal{B}$,

$$\operatorname{Ran}\left(\hat{b}\right) = \left\{\alpha(b) : \alpha \in \tilde{\mathcal{B}}\right\} = \sigma(b)$$

and $r(b) = \|b\|$ as b is normal (see Proposition 4.3), it follows that

$$\left\|\hat{b}\right\|_\infty = r(b) = \|b\|.$$

This shows the Gelfand map is isometric and in particular injective. From this we find that the range, $\hat{\mathcal{B}}$, of the Gelfand map is closed under uniform limits and moreover, $\hat{\mathcal{B}}$ is an algebra closed under complex conjugation because the Gelfand map is a $*-$homomorphism. Also note that $1 = \hat{\mathbf{1}} \in \hat{\mathcal{B}}$ and $\hat{\mathcal{B}}$ separates points. Indeed, if $\alpha, \beta \in \tilde{\mathcal{B}}$ are such that $\hat{b}(\alpha) = \hat{b}(\beta)$ for all $b \in \mathcal{B}$ then $\alpha(b) = \beta(b)$ for all $b \in \mathcal{B}$, i.e. iff $\alpha = \beta$. Given all of this, an application of the Stone-Weierstrass theorem implies $\hat{\mathcal{B}} = C\left(\tilde{\mathcal{B}}\right)$.

■

Corollary 7.23. *A commutative C^*-algebra with identity is isometrically isomorphic to the algebra of complex valued continuous functions on a compact Hausdorff space.*

Notation 7.24 *If $\mathcal{B}$ is a unital commutative C^*-algebra, let $\varphi_{\mathcal{B}} : C\left(\tilde{\mathcal{B}}\right) \to \mathcal{B}$ be the inverse of Gelfand isomorphism, $\mathcal{B} \ni A \to \hat{A} \in C\left(\tilde{\mathcal{B}}\right)$, in Theorem 7.22. That is $\varphi_{\mathcal{B}}(f) = A$ iff $\hat{A} = f$, i.e. $\varphi_{\mathcal{B}}(f)$ is the unique element of $\mathcal{B}$ such that*

$$\alpha(\varphi_{\mathcal{B}}(f)) = f(\alpha) \ \text{ for all } \alpha \in \tilde{\mathcal{B}}.$$

[We might also write $f^\vee$ for $\varphi_{\mathcal{B}}(f)$ so that $f^\vee$ is the unique element of $\mathcal{B}$ such that $(f^\vee)^\wedge = f$.]

Theorem 7.25 (Spectral Theorem). *Let H be a separable Hilbert space and $\mathcal{B}$ be a unital C^*-subalgebra of $B(H)$. Then there exist[3] $\Lambda_N = \{1, 2, \dots, N\} \cap \mathbb{N}$ (for some $N \in \mathbb{N} \cup \{\infty\}$), a probability measure μ measure on $\Omega := \Lambda_N \times \tilde{\mathcal{B}}$ equipped with the product σ – algebra, and a unitary map $U : L^2(\mu) \to H$ such that*

$$U^* A U = M_{\hat{A} \circ \pi} \text{ on } L^2(\mu) \text{ for all } A \in \mathcal{B}, \tag{7.2}$$

where $\pi : \Omega \to \tilde{\mathcal{B}}$ is the second factor projection map, i.e. $\pi(j, \alpha) = \alpha$ for $(j, \alpha) \in \Omega$.

Proof. Let $\varphi_{\mathcal{B}} : C\left(\tilde{\mathcal{B}}\right) \to \mathcal{B}$ be the inverse to the Gelfand isomorphism as in Notation 7.24. Then by Theorem 4.16, there exists an $N \in \mathbb{N} \cup \{\infty\}$, a probability measure μ on Ω as in the statement of the theorem, and a unitary map $U : L^2(\mu) \to H$ such that

$$U^* \varphi_{\mathcal{B}}(f) U = M_{f \circ \pi} \ \forall\ f \in C\left(\tilde{\mathcal{B}}\right).$$

The result in Eq. (7.2) now follows by taking $f = \hat{A}$ while using $\varphi_{\mathcal{B}}\left(\hat{A}\right) = A$.
■

Corollary 7.26. *If $\{T_i\}_{i=1}^N \subset B(H)$ is a collection of commuting normal operators, then there exists a probability space $(\Omega, \mathcal{F}, \nu)$, a unitary map, $U : L^2(\nu) \to H$, and functions, $\{f_i\}_{i=1}^N \subset L^\infty(\nu)$ such that $U^* T_i U = M_{f_i}$ for all i.*

[3] If there is a cyclic vector, $v \in H$, for $\mathcal{B}$, then we can take $N = 1$.

7.2.1 Spectral Theory of Compact Hausdorff Spaces

Exercise 7.2. If $\mathcal{A}$ is an n–dimensional commutative C^*-algebra with identity show that the spectrum of $\mathcal{A}$ consists of exactly n points ($n < \infty$).

Notation 7.27 *For $x \in X$ (X is a compact Hausdorff space), let $\alpha_x : C(X) \to \mathbb{C}$ be the evaluation map,*

$$\alpha_x(f) = f(x) \text{ for all } f \in C(X).$$

Theorem 7.28. *If X is a compact Hausdorff space, then*

$$\widetilde{C(X)} = \operatorname{spec}(C(X)) = \{\alpha_x : x \in X\}$$

and moreover, the map

$$X \ni x \to \alpha_x \in \widetilde{C(X)} \tag{7.3}$$

is a homeomorphism of compact Hausdorff spaces.

Proof. It is easily verified that α_x is a character with corresponding maximal ideal being

$$\mathcal{I}_x = \operatorname{Nul}(\alpha_x) = \{f \in C(X) : f(x) = 0\}.$$

To finish the proof it suffices to show that every maximal ideal of $C(X)$ is of the form $\mathcal{I}_x$ for some $x \in X$.

Let $\mathcal{I} \subset C(X)$ be a maximal ideal. If $\mathcal{I}$ did not separate points there would exist $x \neq y$ in X such that $f(x) = f(y)$ for all $f \in C(X)$. Since $\mathcal{I}$ is an ideal we could use Uryshon's lemma to find $\varphi \in C(X)$ such that $\varphi(x) = 1$ while $\varphi(y) = 0$ and hence we learn that

$$f(x) = \varphi(x) f(x) = (\varphi f)(x) = (\varphi f)(y) = \varphi(y) f(y) = 0$$

for all $f \in \mathcal{I}$. Thus it follows that $f \in \mathcal{I}_x \cap \mathcal{I}_y$ and $\mathcal{I}$ would not be maximal. Thus we know that $\mathcal{I}$ separates points and therefore by the Stone-Weierstrass theorem we must have $\mathcal{I} \subset \mathcal{I}_x$ for some $x \in X$. ■

Exercise 7.3. Prove the second assertion in Theorem 7.28 stating $X \ni x \to \alpha_x \in \widetilde{C(X)}$ is a homeomorphism.

Exercise 7.4. If X and Y are compact Hausdorff spaces and $\varphi : C(X) \to C(Y)$ is a C^*-isomorphism, show there exists a unique homeomorphism $T : Y \to X$ such that $\varphi = T^*$, where $T^* : C(X) \to C(Y)$ is defined by $T^* f = f \circ T$ for all $f \in C(X)$.

The remainder of this section is optional and has not been fully edited as of yet.

Notation 7.29 (Optional) *Suppose that X is a compact Hausdorff space and $\mathcal{A} := C(X, \mathbb{C})$ be the algebra of continuous function on X. To an set $E \subset X$, let*

$$I(E) := \{f \in \mathcal{A} : f|_E \equiv 0\}$$

be the closed ideal in $\mathcal{A}$ of functions vanishing on E. To any subset $T \subset \mathcal{A}$, let

$$Z(T) := \{x \in X : f(x) = 0 \text{ for all } f \in T\}$$

denote the subset of X consisting of the common zeros of functions from T. When $E = \{x\}$ with $x \in X$, we will write $m_x := I(\{x\})$.

Proposition 7.30 (Optional). *Suppose that X is a compact Hausdorff space and $\mathcal{A} := C(X, \mathbb{C})$. Then*

1. *For any subset $E \subset X$, $Z(I(E)) = \bar{E}$.*
2. *For any $T \subset \mathcal{A}$, $I(Z(T)) = \overline{(T)}$ – the closed ideal in $\mathcal{A}$ generated by T. (Items 1. and 2. implies that closed subsets $E \subset X$ are in one to one correspondence with closed ideals in $\mathcal{A}$ via $E \to I(E)$ and $J \to Z(J)$.)*
3. *For each $x \in X$, $m_x := I(\{x\})$ is a maximal (necessarily closed) ideal in $\mathcal{A}$.*
4. *Let $\mathfrak{m}$ denote the collection of maximal ideals in A, then the map $\psi : X \to \mathfrak{m}$ defined by $\psi(x) = m_x$ is bijective.*
5. *If we view $\mathfrak{m}$ as a topological space by transferring the topology on X to $\mathfrak{m}$ using ψ, the closed sets in $\mathfrak{m}$ consist precisely of the sets*

$$C_J := \{m \in \mathfrak{m} : J \subset m\}$$

where J is a closed ideal in $\mathcal{A}$.

Proof. We take each item in turn.

1. Since $Z(T) \subset X$ is closed for any $T \subset \mathcal{A}$ and $E \subset Z(I(E))$, $\bar{E} \subset Z(I(E))$. If $x \notin \bar{E}$, then by Uryhson's lemma, there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$ while $f|_{\bar{E}} = 0$, i.e. $f \in I(E)$. This shows $x \notin Z(I(E))$ and we have proved the first assertion.
2. Since $I(E)$ is a closed ideal for any subset $E \subset X$ and (T) is easily seen to be a subset of $I(Z(T))$, it follows that $\overline{(T)} \subset I(Z(T))$. – the closed ideal in $\mathcal{A}$ generated by T. Let $X_0 := X \setminus Z(T)$, a locally compact space. If $f \in I(Z(T))$ then $f|_{X_0} \in C_0(X_0, \mathbb{C})$ and if $f|_{X_0} \equiv 0$ then $f \equiv 0$ since by assumption $f = 0$ on $Z(T)$. So using this identification we have

$$\overline{(T)} \subset I(Z(T)) \subset C_0(X_0, \mathbb{C}) \tag{7.4}$$

and in particular $\overline{(T)}$ is a closed ideal in $C_0(X_0, \mathbb{C})$. Suppose there exists $x \neq y$ in X_0 such that $f(x) = f(y)$ for all $f \in \overline{(T)}$. Let $\psi \in C_0(X_0, \mathbb{C})$ be chosen so that $\psi(x) = 1$ while $\psi(y) = 0$, then for $f \in \overline{(T)}$, $\psi f \in \overline{(T)}$ and so

$$f(x) = (\psi f)(x) = (\psi f)(0) = 0$$

which shows $f(x) = 0$ for all $f \in T$. But this is impossible because of the definition of $X_0 = X \setminus Z(T)$. So the locally compact form of the Stone-Weierstrass theorem is applicable and implies $\overline{(T)} = C_0(X_0, \mathbb{C})$. Hence by Eq. (7.4), $\overline{(T)} = I(Z(T)) = C_0(X_0, \mathbb{C})$.

3. Suppose $x \in X$ and $f \in \mathcal{A} \setminus m_x$ and let I be the closed ideal generated by f and m_x. It is easily checked that I separates points and $Z(I) = \emptyset$ and hence by the Stone-Weierstrass theorem $I = \mathcal{A}$. This shows that m_x is a maximal ideal which is necessarily closed by the comments at the start of the proof.
4. Clearly the map $\psi : X \to \mathfrak{m}$ is injective. To prove surjectivity, suppose $m \in \mathfrak{m}$ is a maximal ideal. Using the same sort of argument to in the proof of item 2. above, it follows that m separates points. Since m is a closed proper subalgebra of $\mathcal{A}$, the Stone-Weierstrass theorem implies $m = m_x$ for some $x \in X$.
5. For a closed subset $E \subset X$,

$$\psi(E) = \{m_x \in \mathfrak{m} : x \in E\} = \left\{m \in \mathfrak{m} : I(\bar{E}) = I(E) \subset m\right\}.$$

Therefore the closed subsets of $\mathfrak{m}$ are precisely sets of the form

$$C_J := \{m \in \mathfrak{m} : J \subset m\}$$

where J is a closed ideal in $\mathcal{A}$.

■

7.3 Some More Spectral Theory

Proposition 7.31 (Continuous Functional Calculus II). *Suppose that $\mathcal{B}$ is a commutative unital C^*-algebra generated by $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{B}^n$. Let $g_{\mathbf{a}} = \hat{a}_1 \times \dots \times \hat{a}_n : \tilde{\mathcal{B}} \to \mathbb{C}^n$ and $\sigma(\mathbf{a}) \subset \mathbb{C}^n$ be defined by*

$$g_{\mathbf{a}}(\alpha) = (\hat{a}_1(\alpha), \dots, \hat{a}_n(\alpha)) = (\alpha(a_1), \dots, \alpha(a_n)) \tag{7.5}$$

and

$$\sigma(\mathbf{a}) = g_{\mathbf{a}}\left(\tilde{\mathcal{B}}\right) = \left\{\alpha(\mathbf{a}) : \alpha \in \tilde{\mathcal{B}}\right\} \tag{7.6}$$

be the image of $g_{\mathbf{a}}$. Then;

1. *$\sigma(\mathbf{a})$ is compact and $g_{\mathbf{a}} : \tilde{\mathcal{B}} \to \sigma(\mathbf{a})$ is a homeomorphism of compact Hausdorff spaces.*
2. *There exists a C^*-isomorphism, $\varphi_{\mathbf{a}} : C(\sigma(\mathbf{a})) \to \mathcal{B}$, uniquely determined by*
$$\varphi_{\mathbf{a}}(z_i) = a_i \text{ for } i \in [n] = \{1, 2, \dots, n\}. \tag{7.7}$$
In the sequel we will denote $\varphi_{\mathbf{a}}(f)$ by $f(\mathbf{a})$ for any $f \in C(\sigma(\mathbf{a}))$.
3. *If $f \in C(\sigma(\mathbf{a}))$, then $f(\mathbf{a})$ is the unique element[4] of $\mathcal{B}$ such that*
$$\alpha(f(\mathbf{a})) = \alpha(\varphi_{\mathbf{a}}(f)) = f(\alpha(\mathbf{a})) \ \forall\ \alpha \in \tilde{\mathcal{B}}, \text{ where} \tag{7.8}$$
$$\alpha(\mathbf{a}) := g_{\mathbf{a}}(\alpha) = (\alpha(a_1), \dots, \alpha(a_n)),$$
i.e.
$$\alpha(f(a_1, \dots, a_n)) = f(\alpha(a_1), \dots, \alpha(a_n)) \text{ for all } \alpha \in \tilde{\mathcal{B}}. \tag{7.9}$$
4. *The spectral mapping theorem holds in the form, $\sigma(f(\mathbf{a})) = f(\sigma(\mathbf{a}))$.*

Proof. 1. First off $g_{\mathbf{a}}$ is continuous because, by the definition of the weak*-topology, each of the components, $\hat{a}_j$, of $g_{\mathbf{a}}$ are continuous. Since the continuous image of compact sets are compact, it follows that $\sigma(\mathbf{a}) = g_{\mathbf{a}}\left(\tilde{\mathcal{B}}\right)$. If $g_{\mathbf{a}}(\alpha) = g_{\mathbf{a}}(\beta)$ for some $\alpha, \beta \in \tilde{\mathcal{B}}$, then since α and β are $*$–homomorphism it follows that $\alpha = \beta$ on the $*$-algebra generated by $\mathbf{a}$ and then by continuity on all of $\mathcal{B}$. Thus $g_{\mathbf{a}}(\alpha) = g_{\mathbf{a}}(\beta)$ implies $\alpha = \beta$ which means $g_{\mathbf{a}}$ is injective and therefore $g_{\mathbf{a}} : \tilde{\mathcal{B}} \to \sigma(\mathbf{a})$ is continuous bijection. Since $\tilde{\mathcal{B}}$ and $\sigma(\mathbf{a})$ are compact Hausdorff spaces it follows automatically that $g_{\mathbf{a}}$ has a continuous inverse and hence $g_{\mathbf{a}}$ is a homeomorphism.

2. and 3. Let $\varphi_{\mathcal{B}} : C\left(\tilde{\mathcal{B}}\right) \to \mathcal{B}$ be the inverse to the Gelfand isomorphism as in Notation 7.24 and note that $C(K) \ni f \to f \circ g_{\mathbf{a}} \in C\left(\bar{\mathcal{B}}\right)$ is also a C^*-isomorphism. Therefore $\varphi_{\mathbf{a}} : C(K) \to \mathcal{B}$ defined by the composition isomorphism,

$$\varphi_{\mathbf{a}}(f) := \varphi_{\mathcal{B}}(f \circ g_{\mathbf{a}}) \text{ for all } f \in C(K),$$

is again a C^*-isomorphism. Moreover, $\varphi_{\mathbf{a}}$ is uniquely determined by the equation

$$\widehat{\varphi_{\mathbf{a}}(f)} = \varphi_{\mathcal{B}}\widehat{(f \circ g_{\mathbf{a}})} = f \circ g_{\mathbf{a}}$$

which is equivalent to

$$\alpha(f(\mathbf{a})) = \alpha(\varphi_{\mathbf{a}}(f)) = \widehat{\varphi_{\mathbf{a}}(f)}(\alpha) = f \circ g_{\mathbf{a}}(\alpha) = f(\alpha(\mathbf{a})) \ \forall\ \alpha \in \tilde{\mathcal{B}}.$$

Taking $f(z) = z_i$ in Eq. (7.9) implies

[4] This is the analogue of the statement for matrices that if $a_i v = \lambda_i v$, then

$$f(a_1, \dots, a_n) v = f(\lambda_1, \dots, \lambda_n) v.$$

Also note that if f is a polynomial function as above then we would clearly have Eq. (7.9) holding as α is a $*$-algebra homomorphism.

$$\alpha(z_i(\mathbf{a})) = \alpha(\varphi_{\mathbf{a}}(z_i)) = z_i(\alpha(\mathbf{a})) = \alpha(a_i) \text{ for all } \alpha \in \tilde{\mathcal{B}}$$

which proves Eq. (7.7), i.e. $z_i(\mathbf{a}) = a_i$.

4. Item 4. follows directly from item 3.,

$$\sigma(f(\mathbf{a})) = \left\{\alpha(f(\mathbf{a})) : \alpha \in \tilde{\mathcal{B}}\right\} = \left\{f(\alpha(\mathbf{a})) : \alpha \in \tilde{\mathcal{B}}\right\} = f(\sigma(\mathbf{a})).$$

■

Notation 7.32 *Let $\mathcal{B}$ be a C^*-algebra with identity and $\mathbf{a} = (a_1, \ldots, a_n) \in \mathcal{B}^n$.*

1. *Let $\mathcal{S}_{\mathcal{B}} = \mathcal{B} \setminus \mathcal{B}_{inv}$ denote the non-invertible (singular) elements of $\mathcal{B}$.*
2. *For $\lambda \in \mathbb{C}^n$ let*

$$b_\lambda := (\mathbf{a} - \lambda)^* \cdot (\mathbf{a} - \lambda) := \sum_{j=1}^n \left(a_j^* - \bar{\lambda}_j\right)(a_j - \lambda_j).$$

Proposition 7.33. *Continuing the notation and assumptions in Proposition 7.31, we have*

$$\sigma(\mathbf{a}) := \{\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n : b_\lambda \in \mathcal{S}\}.$$

Proof. If $\lambda = \alpha(\mathbf{a}) \in \sigma(\mathbf{a})$ for some $\alpha \in \tilde{\mathcal{B}}$, then

$$\alpha(b_\lambda) = \sum_{j=1}^n |\alpha(a_j) - \lambda_j|^2 = 0 \implies b_\lambda \in \mathcal{S}.$$

Conversely if $\lambda \in \mathbb{C}^n$ is chosen so that $b_\lambda \in \mathcal{S}$, then there exists $\alpha \in \tilde{\mathcal{B}}$ such that

$$0 = \alpha(b_\lambda) = \sum_{j=1}^n |\alpha(a_j) - \lambda_j|^2.$$

For this α we have $\lambda = \alpha(\mathbf{a})$. ■

Corollary 7.34. *If $\mathcal{B}$ is a unital C^*-algebra (not necessarily commutative) and $b \in \mathcal{B}$ is a Hermitian element, such that b^{-1} exists in $\mathcal{B}$, then $b^{-1} \in C^*(b, 1)$.*

Proof. Since b^{-1} is still Hermitian and commutes with b, we may conclude that $\mathcal{A} := C^*\left(b, b^{-1}, 1\right)$ is a commutative C^*-subalgebra of $\mathcal{B}$ with $b^{-1} \in \mathcal{A}$. By Corollary 7.23 we may view b and b^{-1} to be continuous functions on the compact Hausdorff space, $Y = C^*\widetilde{(b, b^{-1}}, 1)$. As $b^{-1} \in C(Y)$, it follows that $\operatorname{Ran}(b)$ is a compact subset $\mathbb{R} \setminus \{0\}$ and so by the Weierstrass approximation theorem we may find $p_n \in \mathbb{C}[x]$ such that

$$\lim_{n\to\infty} \sup_{z \in \operatorname{Ran}(b)} \left|p_n(x) - \frac{1}{x}\right| = 0.$$

Therefore b^{-1} is the uniform limit of $p_n(b) \in C^*(b, 1) \subset C(Y)$ and hence $b^{-1} \in C^*(b, 1)$. ■

Theorem 7.35 ($\sigma_{\mathcal{A}}(a) = \sigma_{\mathcal{B}}(a)$). *Suppose that $\mathcal{B}$ is a unital C^*-algebra and $\mathcal{A} \subset \mathcal{B}$ is a unital C^*—sub-algebra with no commutativity assumptions on $\mathcal{A}$ or $\mathcal{B}$. Then for every $a \in \mathcal{A}$, $\sigma_{\mathcal{A}}(a) = \sigma_{\mathcal{B}}(a)$. In particular, for any $a \in \mathcal{B}$ we have $\sigma_{C^*(a,1)}(a) = \sigma_{\mathcal{B}}(a)$.*

Proof. Recall that we always have,

$$\sigma_{\mathcal{B}}(a) \subset \sigma_{\mathcal{A}}(a) \subset \sigma_{C^*(a,1)}(a). \tag{7.10}$$

If $\lambda \notin \sigma_{\mathcal{B}}(a)$, then $a - \lambda \in \mathcal{B}_{inv}$ and therefore $b := (a - \lambda)^*(a - \lambda)$ is a Hermitian invertible element of $\mathcal{B}$. By Corollary 7.34, $b^{-1} \in C^*(b, 1) \subset C^*(a, 1)$ and therefore

$$(a - \lambda)^{-1} = b^{-1}(a - \lambda)^* \in C^*(a, 1),$$

i.e. $\lambda \notin \sigma_{C^*(a,1)}(a)$. Thus we conclude that $\sigma_{\mathcal{B}}(a)^c \subset \sigma_{C^*(a,1)}(a)^c$ or equivalently, $\sigma_{C^*(a,1)}(a) \subset \sigma_{\mathcal{B}}(a)$ which along with Eq. (7.10) completes the proof. ■

Corollary 7.36 (Positivity). *Let a be a Hermitian element of a unital C^*-algebra, $\mathcal{B}$, then the following are equivalent;*

1. *$a = b^*b$ for some normal*[5] *element $b \in \mathcal{B}$,*
2. *$\sigma(a) \subset [0, \infty)$,*
3. *$a = b^2$ for a unique Hermitian element $b \in \mathcal{B}$ with $\sigma(b) \subset [0, \infty)$. [We will denote this b by $\sqrt{a}$.]*

Proof. 1. $\implies$ 2. If $a = b^*b$ with b normal. If we let $\mathcal{B}_0 = C^*(b, 1)$ (a commutative C^*-algebra), then

$$\sigma(a) = \sigma_{\mathcal{B}_0}(a) = \left\{\alpha(a) : \alpha \in \tilde{\mathcal{B}}_0\right\} = \left\{|\alpha(b)|^2 : \alpha \in \tilde{\mathcal{B}}_0\right\} \subset [0, \infty).$$

2. $\implies$ 3. This was proved in Corollary 4.8. For completeness we repeat a proof here. For existence let $b = \sqrt{a} := \varphi_a\left(\sqrt{\cdot}\right)$. For uniqueness suppose that c is a Hermitian element of $\mathcal{B}$ such that $a = c^2$. Then working in $\mathcal{B}_0 = C^*(c, 1)$, we have $\hat{a} = \hat{c}^2$ which implies $|\hat{c}| = \sqrt{\hat{a}}$. Choose $p_n \in \mathbb{C}[x]$ such that $p_n(x) \to \sqrt{x}$ uniformly for $x \in \sigma(a) = \operatorname{Ran}(\hat{a})$. If we now let $q_n(x) = p_n\left(x^2\right)$, then

$$\begin{aligned}\max_{t\in\sigma(c)} |q_n(t) - |t|| &= \max_{t\in\sigma(c)} \left|p_n\left(t^2\right) - \sqrt{t^2}\right| = \max_{x\in\sigma(c)^2} \left|p_n(x) - \sqrt{x}\right| \\ &= \max_{x\in\sigma(c^2)} \left|p_n(x) - \sqrt{x}\right| = \max_{x\in\sigma(a)} \left|p_n(x) - \sqrt{x}\right| \to 0 \text{ as } n \to \infty,\end{aligned}$$

i.e. $q_n(t) \to |t|$ uniformly on $t \in \sigma(c)$. Thus we may conclude,

[5] The condition that b is normal may be omitted from this statement, see Lemma 7.38 when $\mathcal{B} = B(H)$ and Theorem 9.9 for the general case.

$$b = \sqrt{a} = \lim_{n\to\infty} p_n(a) = \lim_{n\to\infty} p_n(c^2) = \lim_{n\to\infty} q_n(c) = |c|.$$

If we further assume that $\sigma(c) \subset [0,\infty)$ we will know that $|x| = x$ on $\sigma(c)$ and hence $b = |c| = c$ and the uniqueness or b is proved.

3. $\implies$ 1. This is obvious. ■

Definition 7.37. *We say $A \in B(H)$ is non-negative (and write $A \geq 0$) if $\langle Av, v\rangle \geq 0$ for all $v \in H$. [Recall from Lemma 3.27 that $\langle Av, v\rangle \in \mathbb{R}$ for all $v \in H$ implies $A = A^*$.] Moreover, if $A, B \in B(H)$ are self-adjoint operators then we say $A \geq B$ iff $A - B \geq 0$.*

Lemma 7.38. *Suppose that H is a separable Hilbert space and $A \in B(H)$ is a self-adjoint operator, then the following are equivalent;*

1. *$A \geq 0$*
2. *$\sigma(A) \subset [0,\infty)$ and*
3. *$A = B^2$ for some $B \geq 0$.*
4. *$A = B^*B$ for some $B \in B(H)$.*

Proof. (1) $\implies$ (2). **Proof 1.** Suppose that $A \geq 0$. By Eq. (3.7) with $\beta = 0$ and $\alpha > 0$,

$$\|(A+\alpha)\psi\|^2 = \|A\psi\|^2 + 2\alpha(A\psi,\psi) + |\alpha|^2\|\psi\|^2 \geq |\alpha|^2\|\psi\|^2$$

which implies by Lemma 3.34 $-\alpha \notin \sigma(A)$. That is to say $\sigma(A) \subset [0,\infty)$.

Proof 2. By the spectral theorem, we may assume there exists a finite measure space $(\Omega, \mathcal{F}, \mu)$ and a bounded measurable function, $f : \Omega \to \mathbb{R}$, such that $A = M_f$ on $H = L^2(\mu)$. The condition $A \geq 0$ is then equivalent to

$$0 \leq \langle Ag, g\rangle = \int_\Omega f|g|^2 d\mu \;\forall\; g \in L^2(\mu).$$

Taking $g = 1_E$ for $E \in \mathcal{F}$ shows $\int_\Omega f 1_E d\mu \geq 0$ and this is sufficient to show $f \geq 0$ a.e.. Since $\sigma(A) = \text{essran}_\mu(f) \subset [0,\infty)$, the proof is complete.

(2) $\implies$ (3). Take $B = \sqrt{A}$ which exists by the functional calculus or in the model above, take $B = M_{\sqrt{f}}$.

(3) $\implies$ (4) is obvious and (4) $\implies$ (1) is easy since

$$\langle Ax, x\rangle = \|Bx\|^2 \geq 0 \text{ for all } x \in H.$$

■

Exercise 7.5. Suppose that H is a separable Hilbert space and $A \in B(H)$ and $A \geq 0$. Show A^{-1} exists iff there exists $\varepsilon > 0$ so that $A \geq \varepsilon I$, i.e. iff

$$\varepsilon := \inf_{\|x\|=1} \langle Ax, x\rangle > 0.$$

Corollary 7.39 (Joint approximate eigensequences). *Suppose H is a separable Hilbert space, $\{T_j\}_{j=1}^n \subset B(H)$ are commuting normal random variables, and $\mathcal{B} = C^*\left(\{T_j\}_{j=1}^n, I\right)$. Then $\lambda \in \sigma(T_1,\dots,T_n)$ iff there exists $\{x_k\}_{k=1}^\infty \subset H$ such that $\|x_k\| = 1$ and $\lim_{k\to\infty}(T_j - \lambda_j)x_k = 0$ for all $j \in [n]$.*

Proof. Recall that $\lambda \in \sigma(T_1,\dots,T_n)$ iff $b_\lambda = \sum_{j=1}^n (T_j - \lambda_j)^*(T_j - \lambda_j)$ is not invertible. Since $b_\lambda \geq 0$ the following statements are equivalent;

1. b_λ is not invertible,
2. $\inf_{\|x\|=1}\langle b_\lambda x, x\rangle = 0$,
3. there exists $\{x_k\}_{k=1}^\infty \subset H$ such that
$$0 = \lim_{k\to\infty}\langle b_\lambda x_k, x_k\rangle = \lim_{k\to\infty}\sum_{j=1}^n \left\langle (T_j-\lambda_j)^*(T_j-\lambda_j)x_k, x_k\right\rangle$$
$$= \lim_{k\to\infty}\sum_{j=1}^n \|(T_j-\lambda_j)x_k\|^2,$$
4. there exists $\{x_k\}_{k=1}^\infty \subset H$ such that $\lim_{k\to\infty}(T_j-\lambda_j)x_k = 0$ for all $j \in [n]$.

■

Exercise 7.6. Let $(\Omega, \mathcal{F}, \mu)$ be a σ − finite measure space, $H = L^2(\Omega, \mathcal{F}, \mu)$, $f_j : \Omega \to \mathbb{C}$ be bounded measurable functions for $1 \leq j \leq n$, and let $a_j := M_{f_j} \in B(H)$. Letting $\mathbf{a} = (a_1,\dots,a_n)$ and $\mathbf{f} = (f_1,\dots,f_n) : \Omega \to \mathbb{C}^n$, show $\sigma(\mathbf{a}) = \text{essran}_\mu(\mathbf{f})$.

Corollary 7.40 (Spectral Theorem III). *If $\{T_i\}_{i=1}^n \subset B(H)$ is a collection of commuting normal operators and $K = \sigma(T_1,\dots,T_n) \subset \mathbb{C}^n$. Then there exits $N \in \mathbb{N}\cup\{\infty\}$, a probability measure, μ on $\Omega := K \times ([N]\cap\mathbb{N})$, and a unitary map, $U : L^2(\mu) \to H$ so that $U^*T_jU = M_{z_j\circ\pi}$ for $j \in [n]$, where $z_j\circ\pi(\lambda, i) = \lambda_j$ for all $(\lambda, i) \in \Omega$.*

Proof. Let $\mathcal{B} := C^*(T_1,\dots,T_n,I)$ and $\varphi : C(\sigma(T_1,\dots,T_n)) \to \mathcal{B}$ be the unique C^*-isomorphism such that $\varphi(z_i) = T_i$ for $i \in [n]$ as developed in Proposition 7.31 with $a_i = T_i$. The result now follows as a direct application of Theorem 4.16.

■

Remark 7.41. For multiplicity theory for normal operators, see Conway [7], p. 293 where invariants are assigned to normal operators which can be used to classify normal operators up to unitary equivalence. The finial theorem in Theorem 10.21 on p. 301.Given a measurable set $K \subset \mathbb{C}^n$, let $\mathcal{B}^\infty(K)$ denote the bounded complex valued Borel measurable functions on K and let $\mathcal{B}^\infty(K,\mathbb{R})$ denote the subspace of real valued functions. The following theorem is Theorem VII.2 on p.225 of Reed and Simon.

Theorem 7.42 (Functional Calculus). *Let* $\mathbf{T} = (T_1, T_2, \ldots, T_n) \in B(H)^n$ *be a collection of commuting bounded normal operators on a separable Hilbert space* H. *Then there exists a unique map* $\varphi : \mathcal{B}^\infty(\sigma(\mathbf{T})) \to B(H)$ *such that;*

1. φ *is a* $*$ – *homomorphism, i.e.* φ *is linear,* $\varphi(fg) = \varphi(f)\varphi(g)$ *and* $\varphi(\bar{f}) = \varphi(f)^*$ *for all* $f, g \in \mathcal{B}^\infty(\sigma(\mathbf{T}))$.
2. $\|\varphi(f)\|_{op} \leq \|f\|_\infty$ *for all* $f \in \mathcal{B}^\infty(\sigma(\mathbf{T}))$.
3. $\varphi(1) = I$ *and* $\varphi(z_i) = T_i$ *for all* $1 \leq i \leq n$ *where* $z_i : \mathbb{C}^n \to \mathbb{C}$ *is projection onto the* i^{th} – *coordinate. Alternatively stated, if* $p \in \mathbb{C}[z_1, \ldots, z_n, \bar{z}_1, \ldots, \bar{z}_n]$, *then* $\varphi(\tilde{p}) = p(\mathbf{T}, \mathbf{T}^*)$ *where* $\tilde{p}(w) := p(w, \bar{w})$.
4. *If* $f_n \in \mathcal{B}^\infty(\sigma(\mathbf{T}))$ *and* $f_n \to f$ *pointwise and boundedly, then* $\varphi(f_n) \to \varphi(f)$ *strongly.*
   Moreover this map has the following properties
5. *If* $f \geq 0$ *then* $\varphi(f) \geq 0$.
6. *If* $T_i h = \lambda_i h$ *for* $i = 1, \ldots, n$ *then* $\varphi(f)h = f(\lambda)h$ *where* $\lambda = (\lambda_1, \ldots, \lambda_n)$.
7. *If* $B \in B(H)$ *and* $[B, T_i] = 0$ *for* $i = 1, \ldots, n$ *then* $[B, \varphi(f)] = 0$ *for all* $f \in \mathcal{B}^\infty(\sigma(\mathbf{T}))$.

Proof. Uniqueness. Suppose that $\psi : \mathcal{B}^\infty(\sigma(\mathbf{T})) \to B(H)$ is another map satisfying (1) – (4). Let

$$\mathbb{H} := \{f \in \mathcal{B}^\infty(\sigma(\mathbf{T}), \mathbb{C}) : \psi(f) = \varphi(f)\}.$$

Then $\mathbb{H}$ is a vector space of bounded complex valued functions which by property 4. is closed under bounded convergence and by property 1. is closed under conjugation. Moreover $\mathbb{H}$ contains

$$\mathbb{M} = \{\tilde{p} : p \in \mathbb{C}[z_1, \ldots, z_n, \bar{z}_1, \ldots, \bar{z}_n]\}$$

and therefore also $C(\sigma(A), \mathbb{C})$ because of the Stone – Weierstrass approximation theorem. Therefore it follows from Theorem A.9 that $\mathbb{H} = \mathcal{B}^\infty(\sigma(A))$, i.e. $\psi = \varphi$.

Existence. Let $\Omega = \Lambda_N \times \sigma(\mathbf{T})$, μ, and $U : L^2(\Omega, \mu) \to H$ be as in Corollary 7.40 and let $\pi : \Omega \to \sigma(\mathbf{T})$ be projection onto the second factor so that $U^* \mathbf{T} U = M_\pi$ For $f \in \mathcal{B}^\infty(\sigma(\mathbf{T}))$, define

$$\varphi(f) := U M_{f \circ \pi} U^*.$$

One easily verifies that φ satisfies items 1. – 4. Moreover we can easily verify items 5–7 as well.

5. If $f \geq 0$, then $f = \left(\sqrt{f}\right)^2$ and hence $\varphi(f) = \varphi\left(\sqrt{f}\right)^2 \geq 0$.

6. If $\mathbf{T}h = \lambda h$ and $g := U^* h$, then $M_\pi g = \lambda g$ from which it follows that $(\pi_j - \lambda_j) g = 0$ μ – a.e. which implies $\pi_j = \lambda_j$ μ – a.e. on $\{g \neq 0\}$. Thus it follows that $f \circ \pi = f(\lambda)$ μ – a.e. on $\{g \neq 0\}$ and this implies $M_{f \circ \pi} g = f(\lambda) g$ which then implies,

$$\varphi(f)h = \varphi(f)Ug = U M_{f \circ \pi} g = U f(\lambda) g = f(\lambda) h.$$

7. First recall from Theorem 2.68 that we also know that $[B, T_i^*] = 0$ for $1 \leq i \leq n$ and therefore $[B, p(\mathbf{T}, \mathbf{T}^*)] = 0$ for all $p \in \mathbb{C}[z, \bar{z}]$. We now let

$$\mathbb{H} := \{f \in \mathcal{B}^\infty(\sigma(\mathbf{T}), \mathbb{C}) : [B, \varphi(f)] = 0\}.$$

Then $\mathbb{H}$ is a vector space closed under conjugation (again by Theorem 2.68) and bounded convergence. Thus applying the multiplicative system Theorem A.9 with $\mathbb{M} = \{p(\mathbf{T}, \mathbf{T}^*) : p \in \mathbb{C}[z, \bar{z}]\}$ completes the proof.

■

Example 7.43 ([43, Example 10.3]). If $B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ and $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, then $0 \leq A \leq B$ while

$$\det\left(B^2 - A^2\right) = \det\begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix} = -1 < 0$$

and therefore $A^2 \nleq B^2$.

Theorem 7.44 (Löwner-Heinz inequality). *Suppose that* A *and* B *are non-negative bounded operators on a Hilbert space,* H. *If* $0 \leq A \leq B$, *then* $0 \leq A^x \leq B^x$ *for all* $x \in [0, 1]$.

The first result of this form was for matrices in Löwner [25] and then later in the Hilbert space setting by Heinz [22]. The result as stated can be found in Theorem 2?? of Kato [24, Theorem 2?]. For a short proof in the bounded operator setting see [31] and also see [30, Theorem 18]. For a general result for $x > 1$ when B is bounded above and below, see [12] and for a variant of this theme see Ando and Hiai [2]. We will give a (new??) proof of Theorem **??** based on complex interpolation. For the case of unbounded self-adjoint operators, see [43, Proposition 10.14]. [A useful reference for the material here is [43], see Chapter 10 in particular.]

Proof. First let us assume that B^{-1} exists as a bounded operator. (The general case follows by perturbing B and by truncating A and then passing to the limits. Below we will use the following are equivalent characterizations of $0 \leq A \leq B$;

1. $0 \leq A \leq B$,
2. $0 \leq \langle A\varphi, \varphi\rangle \leq \langle B\varphi, B\varphi\rangle$
3. $0 \leq \left\langle AB^{-1/2}\varphi, B^{-1/2}\varphi\right\rangle \leq \langle\varphi, \varphi\rangle = \|\varphi\|^2$,

4. $0 \leq B^{-1/2}AB^{-1/2} \leq I$, and
5. $\left\| A^{\frac{1}{2}} B^{-1/2} \varphi \right\| \leq \|\varphi\|$ for all φ wherein we have used

$$\left\langle AB^{-1/2}\varphi, B^{-1/2}\varphi \right\rangle = \left\langle A^{\frac{1}{2}} B^{-1/2}\varphi, A^{\frac{1}{2}} B^{-1/2}\varphi \right\rangle .$$

Given $\varphi \in \mathcal{K}$ and $z \in \mathbb{C}$ let

$$f(z) := \left\langle B^{-z/2}A^z B^{-z/2}\varphi, \varphi \right\rangle = \left\langle A^{z/2}B^{-z/2}\varphi, A^{\bar{z}/2}B^{-\bar{z}/2}\varphi \right\rangle$$

so that f is a holomorphic function of which is bounded on the strip, $0 \leq \operatorname{Re} z \leq 1$. Moreover if $z = iy$ with $y \in \mathbb{R}$ then

$$|f(iy)| = \left| \left\langle A^{iy/2}B^{-iy/2}\varphi, A^{-iy/2}B^{iy/2}\varphi \right\rangle \right| \leq \|\varphi\|^2$$

and for $z = 1 + iy$,

$$\begin{aligned}
|f(1+iy)| &= \left| \left\langle A^{(1+iy)/2}B^{-(1+iy)/2}\varphi, A^{(1-iy)/2}B^{-(1-iy)/2}\varphi \right\rangle \right| \\
&= \left| \left\langle A^{iy/2}A^{1/2}B^{-1/2}B^{-iy/2}\varphi, A^{-iy/2}A^{1/2}B^{-1/2}B^{iy/2}\varphi \right\rangle \right| \\
&\leq \left\| A^{iy/2}A^{1/2}B^{-1/2}B^{-iy/2}\varphi \right\| \cdot \left\| A^{-iy/2}A^{1/2}B^{-1/2}B^{iy/2}\varphi \right\| \\
&= \left\| A^{1/2}B^{-1/2}B^{-iy/2}\varphi \right\| \cdot \left\| A^{1/2}B^{-1/2}B^{iy/2}\varphi \right\| \\
&\leq \left\| B^{-iy/2}\varphi \right\| \cdot \left\| B^{iy/2}\varphi \right\| = \|\varphi\|^2 .
\end{aligned}$$

Therefore by Haddamard's three line lemma we may conclude that $|f(z)| \leq \|\varphi\|^2$ for all $0 \leq \operatorname{Re} z \leq 1$. Taking $z = x \in [0,1]$ then implies,

$$\left\| A^{x/2}B^{-x/2}\varphi \right\|^2 = \left\langle A^{x/2}B^{-x/2}\varphi, A^{x/2}B^{-x/2}\varphi \right\rangle = f(x) = |f(x)| \leq \|\varphi\|^2$$

for all $\varphi \in \mathcal{K}$ and thus we may conclude that $0 \leq A^x \leq B^x$ for all $x \in [0,1]$. ∎

7.4 Exercises: Spectral Theorem (Multiplication Form)

Exercise 7.7. Suppose that H is a separable Hilbert space and $T \in B(H)$ is a normal operator. Show $T = T^*$ iff $\sigma(T) \subset \mathbb{R}$.

Exercise 7.8. Suppose that H is a separable Hilbert space and $T \in B(H)$ is a normal operator. Show T is unitary iff $\sigma(T) \subset S^1 := \{\lambda \in \mathbb{C} : |\lambda| = 1\}$.

Exercise 7.9. Let A be a self-adjoint operator on an n–dimensional Hilbert space ($n < \infty$) V. Show that the general spectral theorem of Theorem 7.25 or Corollary 7.40 implies that A has an orthonormal basis of eigenvectors. **Hints:** you may assume from the outset that $V = L^2(\Omega, \mathcal{F}, \mu)$ and $A = M_f$ where $(\Omega, \mathcal{F}, \mu)$ is a finite measure space such that $\dim L^2(\mu) = n$ and $f : \Omega \to \mathbb{R}$ is a bounded measurable function. [A preliminary result you might want to first prove is; if $\dim L^2(\Omega, \mathcal{F}, \mu) = n$, then there exists a partition $\Pi = \{\Omega_1, \dots, \Omega_n\} \subset \mathcal{F}$ of Ω so that $\mu(\Omega_i) > 0$ and (for any $A \in \mathcal{F}$) either $\mu(A \cap \Omega_i) = \mu(\Omega_i)$ or $\mu(A \cap \Omega_i) = 0$ for $1 \leq i \leq n$.]

Exercise 7.10. Suppose that $\mathbf{T} = (T_1, T_2, \dots, T_n) \in B(H)^n$ is a collection of commuting bounded normal operators on a separable Hilbert space H. Show; if $D \in B(H)$ is an operator such that $[D, T_j] = 0$ for all $1 \leq j \leq n$, then $[D, f(\mathbf{T})] = 0$ for all bounded measurable functions, $f : \sigma(\mathbf{T}) \to \mathbb{C}$. [Note: by Theorem 2.68, the assumption that $[D, T_j] = 0$ automatically implies $\left[D, T_j^*\right] = 0$.]

Exercise 7.11. Suppose that h is an strictly increasing bounded continuous positive function on $\mathbb{R}$ and $Tf = hf$ for $f \in L^2(\mathbb{R}, m)$. Show if $\Omega(x) > 0$ and $\Omega \in L^2(m)$, then Ω is a cyclic vector for $C^*(T, I)$. Further find the unitary map, $U : L^2(\sigma(T), \mu_\Omega) \to L^2(m)$ in the spectral theorem and show by direct computation that

$$U^*TU = M_z \text{ on } L^2(\sigma(T), \mu_\Omega).$$

Hint: use the multiplicative system theorem to show if $\langle g, h^n\Omega \rangle = 0$ for all $n \in \mathbb{N}_0$, then $g = 0$ a.e.

Exercise 7.12. Let H be a Hilbert space with O. N. basis $e_1, e_2, \dots$. Let θ_j be a sequence of real numbers in $(0, \pi/2)$. Let

$$x_j = (\cos\theta_j)e_{2j} + (\sin\theta_j)e_{2j-1} \quad j = 1, 2, \dots$$

and

$$y_j = -(\cos\theta_j)e_{2j} + (\sin\theta_j)e_{2j-1} \quad j = 1, 2 \dots .$$

Let

$$M_1 = \text{closedspan } \{x_j\}_{j=1}^\infty \text{ and}$$
$$M_2 = \text{closedspan } \{y_j\}_{j=1}^\infty.$$

1. Show that the closed span of M_1 and M_2 (i.e., the closure of $M_1 + M_2$) is all of H.
2. Show that if $\theta_j = 1/j$ then the vector

$$z = \sum_{j=1}^\infty j^{-1} e_{2j-1}$$

is not in $M_1 + M_2$, so that $M_1 + M_2 \neq H$.

Exercise 7.13. Define f on $[0,1]$ by

$$f(x) = \begin{cases} 2 & \text{if } x \text{ is rational} \\ x & \text{if } x \text{ is irrational} \end{cases}.$$

Find the spectrum of M_f as an operator on $L^2(0,1)$.

Exercise 7.14. Let

$$H = \ell^2(\mathbb{Z}) = \{a = \{a_j\}_{j=-\infty}^{\infty} : \|a\|^2 := \sum_{j=-\infty}^{\infty} |a_j|^2 < \infty\}.$$

Define $U : H \to L^2(-\pi,\pi)$ by

$$(Ua)(\theta) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} a_n e^{in\theta}.$$

It is well known that U is unitary (see Theorem **??**). For f in $\ell^1(\mathbb{Z})$ define

$$(C_f a)_n = \sum_{k=-\infty}^{\infty} f(n-k) a_k.$$

1. Show that C_f is a bounded operator on H and that $\|C_f\|_{op} \le \|f\|_1$.
2. Find C_f^* explicitly and show that C_f is normal for any f in $\ell^1(\mathbb{Z})$.
3. Show that UC_fU^{-1} is a multiplication operator.
4. Find the spectrum of C_f, where

$$f(j) = \begin{cases} 1 & \text{if} \quad |j| = 1 \\ 0 & \text{otherwise} \end{cases}.$$

Exercise 7.15. Find a bounded self-adjoint operator, A, with both of the following properties:

1. A has no eigenvectors, and
2. $\sigma(A)$ is set of Lebesgue measure zero in $\mathbb{R}$.

Hint 1: Such an operator is said to have singular continuous spectrum.
Hint 2: Consider the Cantor set, see [40, Section 7.16.].

*Gelfand Theory Expanded

This chapter is highly optional and the material here will not be used later.

8.1 More on the Gelfand Map

Definition 8.1. *Given a commutative Banach algebra ($\mathcal{B}$) with identity we define;*

1. *The **radical** of $\mathcal{B}$ is the intersection of all the maximal ideals in $\mathcal{B}$,*
$$\operatorname{rad}(\mathcal{B}) = \cap\{J : J \text{ is a maximal ideal in } \mathcal{B}\}.$$
[The radical of $\mathcal{B}$ is the intersection of closed ideals and therefore it is also a closed ideal. Let us further note that $a \in \operatorname{rad}(\mathcal{B})$ iff $\alpha(a) = 0$ for all $\alpha \in \tilde{\mathcal{B}}$.]
2. *$\mathcal{B}$ is called **semi-simple** if $\operatorname{rad}(\mathcal{B}) = \{0\}$. [In our finite dimensional examples in 6.5 are semi-simple.]*

Theorem 8.2 (Gelfand). *Let $\mathcal{B}$ be a unital commutative Banach algebra. Then the canonical mapping, $\mathcal{B} \ni a \to \hat{a} \in C\left(\widetilde{\mathcal{B}}\right)$, is a contractive homomorphism from $\mathcal{B}$ into $C\left(\widetilde{\mathcal{B}}\right)$ with $\operatorname{rad}(\mathcal{B})$ being its null-space. In particular $\widehat{(\cdot)}$ is injective iff $\operatorname{rad}(\mathcal{B}) = \{0\}$ i.e. iff $\mathcal{B}$ is semi-simple.*

Proof. Let $a, b \in \mathcal{B}$ and $\alpha \in \tilde{\mathcal{B}}$. Since

$$\widehat{ab}(\alpha) = \alpha(ab) = \alpha(a)\,\alpha(b) = \widehat{a}(\alpha)\,\widehat{b}(\alpha),$$

$\mathcal{B} \ni a \to \hat{a} \in C\left(\widetilde{\mathcal{B}}\right)$ is a homomorphism. Moreover,

$$|\widehat{a}(\alpha)| = |\alpha(a)| \le \|a\| \text{ for all } \alpha \in \tilde{\mathcal{B}}.$$

Hence $\|\hat{a}\|_\infty \le \|a\|$, i.e. canonical mapping is a contraction. Finally, $\widehat{a} = 0$ iff $\alpha(a) = 0$ for all $\alpha \in \tilde{\mathcal{B}}$ iff a is in every maximal ideal, i.e. iff $a \in \operatorname{rad}(\mathcal{B})$. ■

Proposition 8.3. *If $\mathcal{B}$ is a commutative Banach algebra with identity, then*

1. *The radical of $\mathcal{B}$ is given by*
$$\operatorname{rad}(\mathcal{B}) = \{a \in \mathcal{B} : r(a) = 0\}.$$
2. *The canonical map $\hat{}: \mathcal{B} \to C\left(\tilde{\mathcal{B}}\right)$ is an isometry (i.e. $\|\hat{a}\|_\infty = \|a\|$ for all $a \in \mathcal{B}$) iff $\|a^2\| = \|a\|^2$ for all $a \in \mathcal{B}$.*
3. *If $\left\|a^2\right\| = \|a\|^2$ for all $a \in \mathcal{B}$, then $\mathcal{B}$ is semi-simple.*

Proof. We prove each item in turn.

1. Using Theorem 8.2 and item 2. of Proposition 7.18, we have $a \in \operatorname{rad}(\mathcal{B})$ iff $\widehat{a} = 0$ iff $\|\hat{a}\|_\infty = 0$ iff $r(a) = 0$.
2. By item 4. of Proposition , $\|\widehat{a}\|_\infty = \|a\|$ iff $r(a) = \|a\|$. If $r(a) = \|a\|$ for all $a \in \mathcal{B}$ then (by the spectral mapping Theorem 3.42)
$$\left\|a^2\right\| = r(a^2) = r(a)^2 = \|a\|^2.$$
Conversely if $\left\|a^2\right\| = \|a\|^2$ for all $a \in \mathcal{B}$, then by induction, for all $a \in \mathcal{B}$ we also have
$$\|a^{2^n}\| = \|a\|^{2^n} \iff \|a\| = \|a^{2^n}\|^{1/2^n} \text{ for all } n \in \mathbb{N}.$$
The last equality along with Corollary 3.45 gives,
$$\|a\| = \lim_{n\to\infty} \|a^{2^n}\|^{1/2^n} = r(a) \ \forall\, a \in \mathcal{B}.$$
3. By item 2. the map $a \to \hat{a}$ is isometric and hence its null-space, $\operatorname{rad}(\mathcal{B})$, must be $\{0\}$. Alternatively, item 2. gives $\|a\| = \|\widehat{a}\|_\infty = r(a)$ and therefore,
$$\operatorname{rad}(\mathcal{B}) = \{a \in \mathcal{B} : r(a) = 0\} = \{a \in \mathcal{B} : \|a\| = 0\} = \{0\}.$$

■

Remark 8.4. If $\mathcal{B}$ does not have a unit then a similar theory can be developed in which $\tilde{\mathcal{B}}$ is locally compact.

For the rest of this section we will assume that $\mathcal{B}$ is a commutative unital Banach algebra with an involution, $(*)$. The main goal of this section is to prove Theorem 7.22 which asserts that the Gelfand map, $\mathcal{B} \ni b \to \hat{b} \in C\left(\tilde{\mathcal{B}}\right)$, is an isometric isomorphism of C^*-algebras.

Definition 8.5. *An element a is* ***Hermitian*** *if $a = a^*$,* ***strongly positive*** *if $a = b^*b$ for some b,* ***positive*** *if $\sigma(a) \subset [0,\infty)$ and* ***real*** *if $\sigma(a) \subset \mathbb{R}$ is real.*

Definition 8.6. *An involution $*$ in a Banach algebra $\mathcal{B}$ with unit*

1. *is* ***real*** *if $\sigma(a) \subset \mathbb{R}$ when $a = a^*$*
2. *is* ***symmetric*** *if $1 + a^*a$ is invertible for all $a \in \mathcal{B}$. [This is a repeat of Definition 2.73.]*

A Banach algebra $\mathcal{B}$ equipped with a (real) symmetric involution will be called a ***(real) symmetric Banach algebra.*** *[We will see that every C^*-algebra with unit is real, symmetric, and the notion of strongly positive and positivity agree.]*

Example 8.7. Let H be a Hilbert space, then $\mathcal{B} = B(H)$ is a real and symmetric Banach $*$-Algebra where A^* is the adjoint of A for all $A \in B(H)$. Any C^*-subalgebra of $B(H)$ is also a real and symmetric Banach $*$-algebra, see Lemma 2.74 and the next proposition.

Proposition 8.8. *Let $\mathcal{B}$ be a symmetric Banach algebra and $a \in \mathcal{B}$.*

1. *If a is Hermitian then a is real ($\sigma(a) \subset \mathbb{R}$ if $a = a^*$), i.e. $\mathcal{B}$ is real.*
2. *If a is strongly positive then a is positive, i.e. $\sigma(x^*x) \subset [0,\infty)$ for all $x \in \mathcal{B}$.*

Proof. We take each item in turn.

1. If a is Hermitian ($a^* = a$) and $\lambda = \alpha + \beta i \in \mathbb{C}$ with $\beta \neq 0$, then
$$\beta^{-1}(a - \lambda) = (a - \beta^{-1}\alpha) - i = x - i$$
where $x := (a - \beta^{-1}\alpha)$ is still Hermitian. Since
$$(x-i)^*(x-i) = x^*x + 1 \text{ and}$$
$$(x-i)(x-i)^* = xx^* + 1$$
we discover that
$$(x^*x+1)^{-1}(x-i)^*(x-i) = 1 \text{ and}$$
$$(x-i)(x-i)^*(xx^*+1)^{-1} = 1.$$
These equations shows $x - i$ has both a right and a left inverse, $x - i$ is invertible and therefore so is $a - \lambda$. This shows $\lambda \in \sigma(a)$ implies $\operatorname{Im}\lambda = 0$, i.e. $\sigma(a) \subset \mathbb{R}$.
2. Suppose that a is strongly positive, $a = b^*b$. Then $a^* = b^*b = a$ showing that a is Hermitian and hence by (1) that $\sigma(a) \subset \mathbb{R}$. If $\alpha > 0$, then
$$b^*b - (-\alpha) = b^*b + \alpha = \alpha\left(\frac{b^*b}{\alpha} + 1\right)$$
$$= \alpha\left(\left(\frac{b}{\sqrt{\alpha}}\right)^*\left(\frac{b}{\sqrt{\alpha}}\right) + 1\right)$$
which is invertible showing $\sigma(b^*b) \subset [0,\infty)$.

■

Our first order of business towards proving this theorem is to give conditions on $(\mathcal{B}, *)$ so that the Gelfand-map is a $*$-homomorphism.

Proposition 8.9. *Let $\mathcal{B}$ be a commutative, unital, and $*$-algebra. The following are equivalent:*

1. *$\mathcal{B}$ is symmetric, i.e. $a^*a + 1$ is invertible for all $a \in \mathcal{B}$.*
2. *Every Hermitian element, $a \in \mathcal{B}$, is real, i.e. if $a^* = a$, then $\sigma(a) \subset \mathbb{R}$.*
3. *If $\alpha \in \tilde{\mathcal{B}}$ then $\alpha(a^*) = \overline{\alpha(a)}$ for all $a \in \mathcal{B}$. [Alternatively put the Gelfand map, $\mathcal{B} \ni a \to \hat{a} \in C\left(\tilde{\mathcal{B}}\right)$ is a $*$ – homomorphism of Banach algebras, i.e. $\widehat{a^*} = \bar{\hat{a}}$ for all $a \in \mathcal{B}$.]*
4. *Every maximal ideal, J, of $\mathcal{B}$ is a $*$-ideal, i.e. if $a \in J$ then $a^* \in J$.*

Proof. 1) $\Rightarrow$ 2) This is Proposition 8.8.

2) $\Rightarrow$ 3) Let $a \in \mathcal{B}$,
$$b = \operatorname{Re} a := \frac{1}{2}(a + a^*) \text{ and } c = \operatorname{Im} a := \frac{1}{2i}(a - a^*).$$
Then b and c are Hermitian and so by Proposition 8.3, $\alpha(b) \in \sigma(b) \subset \mathbb{R}$ and $\alpha(c) \in \sigma(c) \subset \mathbb{R}$ for all $\alpha \in \tilde{\mathcal{B}}$. Since $a = b + ic$ it follows that
$$\alpha(a^*) = \alpha(b - ic) = \alpha(b) - i\alpha(c) = \overline{\alpha(b) + i\alpha(c)} = \overline{\alpha(a)}.$$

3) $\Rightarrow$ 1). For any $a \in \mathcal{B}$ and $\alpha \in \tilde{\mathcal{B}}$, we now have,
$$\alpha(a^*a) = \alpha(a^*)\alpha(a) = \overline{\alpha(a)}\alpha(a) = |\alpha(a)|^2$$
and therefore
$$\alpha(1 + a^*a) = 1 + |\alpha(a)|^2 \neq 0.$$
As this is true for all $\alpha \in \tilde{\mathcal{B}}$ we conclude that $0 \notin \sigma(1 + a^*a)$ by Proposition 8.3, i.e. $1 + a^*a$ is invertible.

3) $\Rightarrow$ 4) Let J be a maximal ideal and let $\alpha \in \tilde{\mathcal{B}}$ be the unique character such that $\operatorname{Nul}(\alpha) = J$, see Lemma 7.11. Since $\alpha(a) = 0$ iff $0 = \overline{\alpha(a)} = \alpha(a^*)$ and $J = \operatorname{Nul}(\alpha)$, it follows that $a \in J$ iff $a^* \in J$.

4) $\Rightarrow$ 3) Given $a \in \mathcal{B}$ and $\alpha \in \tilde{\mathcal{B}}$, let $b = a - \alpha(a) \in \operatorname{Nul}(\alpha) =: J$. By assumption we have $a^* - \overline{\alpha(a)} = b^* \in J = \operatorname{Nul}(\alpha)$ and therefore,

$$0=\alpha\left(b^{*}\right)=\alpha\left(a^{*}\right)-\overline{\alpha\left(a\right)}\implies\alpha\left(a^{*}\right)=\overline{\alpha\left(a\right)}.$$

■

Theorem 8.10 (A Dense Range Condition). *If $\mathcal{B}$ is a commutative Banach $*$ -algebra with unit which is symmetric (or equivalently real), then the image,*

$$\hat{\mathcal{B}}=\left\{\hat{b}\in C\left(\tilde{\mathcal{B}}\right):b\in\mathcal{B}\right\}, \tag{8.1}$$

of the Gelfand map is dense in $C\left(\widetilde{\mathcal{B}}\right)$.

Proof. From Proposition 8.9, the Gelfand map is a $*-$ homomorphism and therefore $\hat{\mathcal{B}}$ is closed under conjugation.[1] Hence, by the Stone-Weierstrass theorem 7.21, it suffices to observe; 1) $1=\hat{\mathbf{1}}\in\hat{\mathcal{B}}$ and 2) $\hat{\mathcal{B}}$ separates points. Indeed, if $\alpha_1,\alpha_2\in\tilde{\mathcal{B}}$ such that $\hat{a}\left(\alpha_1\right)=\hat{a}\left(\alpha_2\right)$ for all $a\in\mathcal{B}$ then

$$\alpha_1\left(a\right)=\hat{a}\left(\alpha_1\right)=\hat{a}\left(\alpha_2\right)=\alpha_2\left(a\right)\ \forall\ a\in\mathcal{B},$$

i.e. $\alpha_1=\alpha_2$.

■

Lemma 8.11 (An Isometry Condition). *If $\mathcal{B}$ is a unital commutative $*$-multiplicative Banach algebra [i.e. $\|a^*a\|=\|a^*\|\,\|a\|$ as in Definition* **??***], then the Gelfand map is isometric, i.e.*

$$\|a\|=\|\hat{a}\|_\infty=r\left(a\right)\ \forall\ a\in\mathcal{B}. \tag{8.2}$$

In particular, $\mathcal{B}$ is semi-simple, i.e. $\operatorname{rad}\left(\mathcal{B}\right)=\{0\}$.

Proof. If b is Hermitian, then

$$\left\|b^2\right\|=\|b^*b\|=\|b^*\|\,\|b\|=\|b\|^2$$

and by induction, $\left\|b^{2^n}\right\|=\|b\|^{2^n}$. It then follows from Corollary 3.45 that

$$r\left(b\right)=\lim_{n\to\infty}\left\|b^{2^n}\right\|^{2^{-n}}=\|b\|.$$

If $a\in\mathcal{B}$ is now arbitrary, then a^*a is Hermitian and therefore

$$r\left(a^*a\right)=\|a^*a\|=\|a^*\|\,\|a\|.$$

On the other hand by Proposition 8.3,

[1] If $a\in\mathcal{B}$, then $cong\left(\hat{a}\right)=\widehat{a^*}\in\hat{\mathcal{B}}$.

$$\|a^*\|\,\|a\|=r(a^*a)\le r(a^*)r\left(a\right)\le\|a^*\|\,r\left(a\right)$$

from which it follows that $r\left(a\right)\ge\|a\|$ or $\|a^*\|=0$. (If $\|a^*\|=0$ then $a^*=0$ and hence $a=a^{**}=0$ and we will have $r\left(a\right)=\|a\|$.) Since $r\left(a\right)\le\|a\|$ by Proposition 8.3, we have now shown $\|a\|=r\left(a\right)$.

For the semi–simplicity of $\mathcal{B}$ we have by Item 5 of Proposition 8.3 that

$$\operatorname{rad}\left(\mathcal{B}\right)=\left\{a\in\tilde{\mathcal{B}}:r\left(a\right)=0\right\}$$

while from Lemma 8.11 we know $r\left(a\right)=\|a\|$ and thus $\operatorname{rad}\left(\mathcal{B}\right)=\{0\}$, i.e. $\mathcal{B}$ is semi-simple. ■

We are now going to apply the previous results when $\mathcal{B}$ is a C^*-algebra. As we have claimed in Remark 2.52, every C^*-algebra can be viewed as a C^*-sub-algebra of $B\left(H\right)$ for some Hilbert space H. This comment along with Lemma 2.74 then implies that every C^*-algebra is symmetric whether it is commutative or not. As we have not proved the claim in Remark 2.52, for completeness we will prove directly the symmetry condition for commutative C^*-algebras.

Lemma 8.12 (Commutative C^*- algebras are symmetric). *A commutative C^*-algebra, $\mathcal{B}$, with identity is symmetric. [This is equivalent to every $\alpha\in\tilde{\mathcal{B}}$ being a $*$-homomorphism.]*

Proof. By Proposition 8.9, to show $\mathcal{B}$ is symmetric it suffices to show $\mathcal{B}$ is real, i.e. we must show $\sigma\left(a\right)\subset\mathbb{R}$ if $a\in\mathcal{B}$ is Hermitian. However, this has already been done in Lemma 4.5. Alternatively we may appeal to Lemma 7.19 which asserts that every $\alpha\in\tilde{\mathcal{B}}$ is a $*$-homomorphism. ■

Remark 8.13. The Shirali-Ford Theorem asserts that a Banach algebra with involution is symmetric iff it is real. We will prove a special case of this below for commutative Banach algebras in Proposition 8.9. In fact almost all of the algebras we will consider here are going to be symmetric. [For example Lemma 8.12 shows every commutative C^*-algebra is symmetric.] For some examples of non-symmetric Banach algebras, see Tenna Nielsen Bachelor's Thesis, "Hermitian and Symmetric Banach Algebras" where it is shown that $\ell^1\left(\mathbb{F}_n\right)$ is not a Hermitian (hence not symmetric) Banach algebra if $\mathbb{F}_n$ is the free group on n – generators with $n\ge2$. The reader can also find a proof of the Shirali-Ford Theorem stated on p. 20 of this reference, also see [3].

Theorem 8.14. *If $\mathcal{B}$ is a commutative Banach $*$ -algebra with unit which is symmetric and $*$–multiplicative [i.e. $\|a^*a\|=\|a^*\|\,\|a\|$ as in Definition 2.47], then the Gelfand map, $\mathcal{B}\ni b\to\hat{b}\in C\left(\tilde{\mathcal{B}}\right)$, is an isometric $*-$isomorphism* ***onto*** *$C\left(\widetilde{\mathcal{B}}\right)$. In particular, it follows that $\mathcal{B}$ is a C^*-algebra.*

Proof. By Proposition 8.9, the Gelfand map,

$$\mathcal{B} \ni b \to \hat{b} \in \hat{\mathcal{B}} \subset C\left(\tilde{\mathcal{B}}, \mathbb{C}\right),$$

is a $*-$ algebra homomorphism. Lemma 8.11 may be applied to show that Gelfand map is an isometry which in turn implies that the image $\left(\hat{\mathcal{B}}\right)$ of the Gelfand map is complete and therefore closed. By Theorem 8.10, $\hat{\mathcal{B}}$ is dense in $C\left(\tilde{\mathcal{B}}\right)$ and therefore (being closed) is equal to $C\left(\tilde{\mathcal{B}}\right)$ and the proof is complete. ■

8.2 Examples of spec ($\mathcal{B}$)

Example 8.15. Let $n \in \mathbb{N}, T \in B(\mathbb{F}^n)$, and $\mathcal{B} = \{p(T) : p \in \mathbb{C}[z]\}$. From Corollary 3.10 we know that $\sigma_{\mathcal{B}}(T) = \sigma(T)$ – the eigenvalues of T. Thus if $\alpha \in \tilde{\mathcal{B}}$, then $\alpha(T) = \lambda \in \sigma(T)$ and hence $\alpha(p(T)) = p(\alpha(T)) = p(\lambda)$. Conversely if $\lambda \in \sigma(T)$ then $T - \lambda$ is not invertible and there exists $\alpha \in \tilde{\mathcal{B}}$ such that $\alpha(T - \lambda) = 0$, i.e. $\alpha(T) = \lambda$ and we denote this α by α_λ. We have shown that the map

$$\sigma(T) \ni \lambda \to \alpha_\lambda \in \tilde{\mathcal{B}}$$

gives a one to one correspondence between $\sigma(T)$ and $\tilde{\mathcal{B}}$. In short spec$(\mathcal{B}) \cong \sigma(T)$.

Example 8.16. Continuing the notation of the previous example,we have

$$\mathrm{rad}(\mathcal{B}) = \cap_{\lambda \in \sigma(T)} \mathrm{Nul}(\alpha_\lambda) = \{p(T) : p(\sigma(T)) = \{0\}\}.$$

Thus if $\#(\sigma(T)) < n$, it follows that $p_{\min}(z) := \prod_{\lambda \in \sigma(T)} (z - \lambda)$ is a polynomial such that $p_{\min}(T) \in \mathrm{rad}(\mathcal{B})$ and in fact

$$\mathrm{rad}(\mathcal{B}) = \{q(T) : q \in (p_{\min})\}.$$

For example if T is nilpotent so that $\sigma(T) = \{0\}$ we have $p_{\min}(z) = z$ and so

$$\mathrm{rad}(\mathcal{B}) = \{p(T) : p \in \mathbb{F}[z] \text{ with } p(0) = 0\}.$$

On the other hand if $T \in B(\mathbb{C}^n)$ is normal $(TT^* = T^*T)$, then $p_{\min}(T) = 0$ and we learn that

$$\mathrm{rad}(\mathcal{B}) = \{0\}.$$

Example 8.17. As another example, suppose that $T \in B(\mathbb{F}^n)$ is the block diagonal with blocks of the form $\lambda_j I + N_j$ where $\lambda_j \in \mathbb{F}$ and N_j is nilpotent. Then $\sigma(T) = \{\lambda_j\}_{j=1}^k$ and

$$\begin{aligned} p_{\min}(\lambda_j I + N_j) &= \prod_{l=1}^k [(\lambda_j - \lambda_l) I + N_j] = N_j \cdot \prod_{l \neq j} [(\lambda_j - \lambda_l) I + N_j] \\ &= N_j [c_j I + q_j(N_j)]. \end{aligned}$$

where $q(0) = 0$ and $c_j \neq 0$. Take N_j to be the matrix of 1's just above the diagonal (Jordan canonical form) so that $N_j q_j(N_j)$ has no ones on the entries just above the diagonal hence shown $p_{\min}(\lambda_j I + N_j) = c_j N_j + O\left(N_j^2\right) \neq 0$ for all j. In this case $\mathrm{rad}(\mathcal{B}) = (p_{\min}(T)) \neq \{0\}$.

Example 8.18 (Continuation of Example **??**). Let us continue the notation in Proposition 3.14. Our goal is to work out $\tilde{\mathcal{B}} = \mathrm{spec}(\mathcal{B})$ where $\mathcal{B} = \overline{\{p(z) : p \in \mathbb{C}[z]\}}^{\mathcal{A}}$ and $\mathcal{A} = C\left(S^1\right)$. [Please note that $\mathcal{B}$ is not a C^*-algebra as it is not closed under the involution, $f \to \bar{f}$.] Here are the salient features.

1. By the maximum principle if $p_n \in \mathbb{C}[z]$ and $p_n \to f$ on S^1, then $p_n \to F \in C(\bar{D}) \cap H(D)$ on D. Thus to each $f \in \mathcal{B}$ we have an uniquely determined $F_f \in C(\bar{D}) \cap H(D)$ such that $F_f|_{S^1} = f$. Notice that $F_{p|_{S^1}} = p$ for all $p \in \mathbb{C}[z]$.
2. If $f, g \in \mathcal{B}$ then $F_f F_g = F_{fg}$. In particular if $f, g \in \mathcal{B}$ with $fg = 1$, then $F_f F_g = 1$. In particular, this shows if $f \in \mathcal{B}$ is invertible in $\mathcal{B}$ then $F_f(\lambda) \neq 0$ for all $x \in \bar{D}$.
3. Similarly, if $f \in \mathcal{B}$ and $p \in \mathbb{C}[z]$, then $pF_f = F_{pf}$.
4. If $\lambda \in \bar{D}$, we may define $\alpha_\lambda(f) := F_f(\lambda)$. Then by item 2 we will have that $\alpha_\lambda \in \tilde{\mathcal{B}}$.
5. Conversely if $\alpha \in \tilde{\mathcal{B}}$, then $\lambda = \alpha(z) \in \sigma_{\mathcal{B}}(z) \subset \bar{D}$ (as $\|z\|_\infty = 1$). If $\lambda \in S^1$, then $\lambda \in \sigma_{\mathcal{A}}(z)$. While if $|\lambda| < 1$, then we have
$$\alpha(p) = p(\lambda) = F_p(\lambda).$$
Since α is continuous it follows that in fact $\alpha(f) = F_f(\lambda)$ for all $f \in \mathcal{B}$. Thus we have shown.
6. There is a one to one correspondence between $\tilde{\mathcal{B}}$ and $\bar{D}$ given by
$$\tilde{\mathcal{B}} \ni \alpha \to \alpha(z) \in \bar{D}.$$
The inverse map is given by
$$\bar{D} \ni \lambda \to \alpha_\lambda(f) := F_f(\lambda).$$
7. As a consequence, if $f \in \mathcal{B}$ we have $\sigma_{\mathcal{B}}(f) = \left\{F_f(\lambda) : \lambda \in \bar{D}\right\}$ while $\sigma_{\mathcal{A}}(f) = \left\{f(\lambda) : \lambda \in S^1\right\}$ and in particular typically,
$$\sigma_{\mathcal{A}}(f) \subset \sigma_{\mathcal{B}}(f).$$

8. Note if $f \in \mathcal{B}$, I think we should be able to prove that $f\left(e^{i\theta}\right) = \sum_{n=0}^{\infty} c_n e^{in\theta}$ and therefore
$$F_f(\lambda) = \sum_{n=0}^{\infty} c_n \lambda^n \text{ for } |\lambda| < 1.$$
The point is that these assertions are true when $f(z) = p(z)$ with $p \in \mathbb{C}[z]$ and then the general result follows by taking limits. Thus we learn more explicitly that
$$\sigma_{\mathcal{B}}(f) = \left\{\sum_{n=0}^{\infty} c_n \lambda^n : |\lambda| \le 1\right\}$$
where one has to interpret $\sum_{n=0}^{\infty} c_n \lambda^n$ as $f(\lambda)$ for $\lambda \in S^1$.

Example 8.19. Let $\mathbf{T} = M_{\mathbf{q}}$, then $\operatorname{spec}(C^*(\mathbf{T})) \cong \sigma(\mathbf{T})$.

Example 8.20. Let $\mathcal{B} = \ell^1(\mathbb{Z})$. If $\alpha \in \tilde{\mathcal{B}}$ and $z := \alpha(\delta_1) \in \mathbb{C}$ then $|z| \le \|\delta_1\| = 1$ and $\left|z^{-1}\right| = |\alpha(\delta_{-1})| \le \|\delta_{-1}\| = 1$. This shows that $z \in S^1$ and for $f \in \mathcal{B}$ we have,
$$f = \sum_n f(n)\,\delta_n \implies \alpha(f) = \sum_n f(n)\,\alpha(\delta_n) = \sum_{n=-\infty}^{\infty} f(n)\,z^n.$$
Conversely given $z \in S^1$ we may define $\alpha_z \in \tilde{\mathcal{B}}$ so that
$$\alpha_z(f) = \sum_{n=-\infty}^{\infty} f(n)\,z^n \;\forall\; f \in \mathcal{B}$$
and so we have shown
$$\tilde{\mathcal{B}} = \operatorname{spec}(\mathcal{B}) = \left\{\alpha_z : z \in S^1\right\}.$$
Consequently for $f \in \mathcal{B}$ we have
$$\sigma(f) = \left\{\alpha_z(f) : z \in S^1\right\} = \operatorname{Ran}\left(S^1 \ni z \to \alpha_z(f)\right).$$
From this we may conclude that $f \in \mathcal{B}$ is invertible iff
$$S^1 \ni z \to \alpha_z(f) = \sum_{n=-\infty}^{\infty} f(n)\,z^n$$
is never zero. Since $f^{-1} f = \delta_0$ we find that
$$\alpha_z\left(f^{-1}\right) = \frac{1}{\alpha_z(f)} = \frac{1}{\sum_{n=-\infty}^{\infty} f(n)\,z^n}$$
which implies the result if $f \in \ell^1(\mathbb{Z})$ such that $\alpha_z(f) \ne 0$ for all $z \in S^1$ then there exists a unique $g \in \ell^1(\mathbb{Z})$ such that
$$\alpha_z(g) = \frac{1}{\alpha_z(f)} \text{ for all } z \in S^1.$$
Further notice that $\alpha_z(f) = 0$ for all $z \in S^1$ implies (Fourier theory) that $f = 0$ and this shows $\operatorname{rad}(\mathcal{B}) = \{0\}$ so the Gelfand map
$$\ell^1(\mathbb{Z}) \ni f \to (z \to \alpha_z(f)) \in C\left(S^1\right)$$
is an injective $*-$ homomorphism with dense range.

8.3 Gelfand Theory Exercises

In each of the following two problems a commutative $*$ algebra $\mathcal{A}$ with identity is given. In each case

1. Find the spectrum of $\mathcal{A}$.
2. Determine whether $\mathcal{A}$ is semi–simple or symmetric or a C^*-algebra, or several of these.
3. Determine whether the Gelfand map is one to one, or onto or both or neither or has dense range.

Exercise 8.1. Let $\mathcal{A}$ be the $*$-algebra of 2×2 complex matrices of the form
$$A = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \text{ but with } A^* := \begin{pmatrix} \bar{a} & \bar{b} \\ 0 & \bar{a} \end{pmatrix}.$$
The norm on $\mathcal{A}$ is still taken to be the operator norm, $\|A\|$, associated to the usual inner product on $\mathbb{C}^2$ with associated norm
$$\left\|\begin{pmatrix} c \\ d \end{pmatrix}\right\| = (|c|^2 + |d|^2)^{1/2}.$$

Exercise 8.2. $\mathcal{A} = \ell^1(\mathbb{Z})$ where $\mathbb{Z}$ is the set of all integers. For f and g in $\mathcal{A}$ define
$$(fg)(x) = \sum_{n=-\infty}^{\infty} f(x-n)g(n)$$
and $f^*(x) = \overline{f(-x)}$. Show first that $\mathcal{A}$ is a commutative $*$-Banach algebra with identity which is **not** a C^*-algebra. You may cite any results from [41].

* More C^*-Algebras Properties

This is another **optional chapter** which contains some more interesting C^*-algebra properties along with some alternate proof of results already proved above. **Warning:** this chapter has not been properly edited!!

Lemma 9.1. *If $\rho : \mathcal{A} \to \mathcal{B}$ is a homomorphism between two unital C^*-algebras, then*

1. *$\sigma(\rho(a)) \subset \sigma(a)$ and $r(\rho(a)) \leq r(a)$ for all $a \in \mathcal{A}$.*
2. *ρ is contractive, i.e. $\|\rho\|_{op} \leq 1$.*
3. *If $a \in \mathcal{A}$ is normal, then*

$$f(\rho(a)) = \rho(f(a)) \text{ for all } f \in C(\sigma(a)).$$

Proof. 1. If a is invertible in $\mathcal{A}$ then $\rho(a^{-1}) = \rho(a)^{-1}$ so that $\rho(a)$ is invertible in $\mathcal{B}$. From this it follows $\sigma(a)^c \subset \sigma(\rho(a))^c$, i.e. $\sigma(\rho(a)) \subset \sigma(a)$ and this suffices to show $r(\rho(a)) \leq r(a)$.

2. Using the C^*-condition and its consequence, $r(b) = \|b\|$ when $b = b^*$ in Proposition 4.3, we find

$$\begin{aligned}\|\rho(a)\|^2 &= \|\rho(a)^* \rho(a)\| = \|\rho(a^*a)\| \\ &= r(\rho(a^*a)) \leq r(a^*a) = \|a^*a\| = \|a\|^2.\end{aligned}$$

3. Let $\mathcal{A}_0 = C^*(a, 1)$ and $\mathcal{B}_0 = C^*(\rho(a), 1)$ and recall that for $f \in C(\sigma(a))$ that $f(a)$ is the unique element of $\mathcal{A}_0$ such that $\alpha(f(a)) = f(\alpha(a))$ for all $\alpha \in \widetilde{\mathcal{A}}_0$, see Eq. (7.9). Since $\sigma(\rho(a)) \subset \sigma(a)$, $f(\rho(a))$ is well defined and of course also uniquely determined by $\beta(f(\rho(a))) = f(\beta(\rho(a)))$ for all $\beta \in \tilde{\mathcal{B}}_0$.

Since $\rho_0 := \rho|_{\mathcal{A}_0} : \mathcal{A}_0 \to \mathcal{B}_0$ is a C^*-homomorphism, if $\beta \in \tilde{\mathcal{B}}_0$, then $\alpha = \beta \circ \rho_0 \in \widetilde{\mathcal{A}}_0$ and hence

$$\beta(\rho_0(f(a))) = (\beta \circ \rho_0)(f(a)) = f((\beta \circ \rho_0)(a)) = f(\beta(\rho(a))).$$

As this holds for all $\beta \in \tilde{\mathcal{B}}_0$ if follows that

$$\rho(f(a)) = \rho_0(f(a)) = f(\rho_0(a)) = f(\rho(a)).$$

■

Lemma 9.2. *If $\rho : \mathcal{A} \to \mathcal{B}$ is an injective unital $*$-homomorphism and $a \in \mathcal{A}$ is a normal element, then $\sigma(\rho(a)) = \sigma(a)$.*

Proof. By Lemma 9.1 we know that $\operatorname{spec}(\rho(a)) \subset \operatorname{spec}(a)$. For sake of contradiction, suppose that $\operatorname{spec}(\rho(a)) \subsetneq \operatorname{spec}(a)$. Let $f(x) := \mathit{dist}(x, \operatorname{spec}(\rho(a)))$ in which case f is non-zero continuous function on $\operatorname{spec}(a)$ which vanishes on $\sigma(\rho(a))$. As we have already shown in Lemma 9.1, $\rho(f(a)) = f(\rho(a)) = 0$. As ρ is injective this would lead to the contradiction that $f(a) = 0$ which is a contradiction since f is not the zero function on $\sigma(a)$. ■

Theorem 9.3. *If $\rho : \mathcal{A} \to \mathcal{B}$ is an injective unital $*$-homomorphism, then ρ is isometric.*

Proof. If $a \in \mathcal{A}$ is self-adjoint, then by Lemma 9.2, $\|\rho(a)\| = r(\rho(a)) = r(a) = \|a\|$. For general $a \in \mathcal{A}$,

$$\|\rho(a)\|^2 = \|\rho(a)^* \rho(a)\| = \|\rho(a^*a)\| = \|a^*a\| = \|a\|^2.$$

■

Next we give another (more elementary but trickier) proof of a special case of Theorem 7.35.

Theorem 9.4. *Suppose that $\mathcal{B}$ is a unital C^*-algebra (not assumed to be commutative) and $\mathcal{A} \subset \mathcal{B}$ is a unital commutative C^*—sub-algebra. Then for every $a \in \mathcal{A}$, $\sigma_{\mathcal{A}}(a) = \sigma_{\mathcal{B}}(a)$.*

Proof. Since it is easier to find an inverse if we are allowed to look for this inverse in $\mathcal{B}$ rather than just in $\mathcal{A}$, it follows that if $a \in \mathcal{A}$ is invertible in $\mathcal{A}$ then it is invertible in $\mathcal{B}$. The contrapositive is that if a is not invertible in $\mathcal{B}$ then it is not invertible in $\mathcal{A}$ which directly shows that $\sigma_{\mathcal{B}}(a) \subset \sigma_{\mathcal{A}}(a)$ for all $a \in \mathcal{A}$. To prove the converse inclusion it suffices to show $\sigma_{\mathcal{B}}(a)^c \subset \sigma_{\mathcal{A}}(a)^c$ which amounts to showing if $a \in \mathcal{A}$ has in inverse in $\mathcal{B}$ then a has an inverse in $\mathcal{A}$.

So suppose that $a \in \mathcal{A}$ has an inverse in $\mathcal{B}$. Then $a^* \in \mathcal{A}$ also has an inverse in $\mathcal{B}$ and hence $b = a^*a \in \mathcal{A}$ has an inverse in $\mathcal{B}$. As we have now seen we also know

$$\sigma_{\mathcal{B}}(a^*a) \subset \sigma_{\mathcal{A}}(a^*a) = \left\{\alpha(a^*a) : \alpha \in \widetilde{\mathcal{A}}\right\} = \left\{|\alpha(a)|^2 : \alpha \in \widetilde{\mathcal{A}}\right\}.$$

So we may conclude that

$$\sigma_{\mathcal{B}}(a^*a) \subset \sigma_{\mathcal{A}}(a^*a) \subset [0, \|a^*a\|]$$

and a^*a is invertible in $\mathcal{B}$ we know that $0 \notin \sigma_{\mathcal{B}}(a^*a)$. Since $\sigma_{\mathcal{B}}(a^*a)$ is compact there exists $\delta > 0$ such that

$$\sigma_{\mathcal{B}}(a^*a) \subset \sigma_{\mathcal{A}}(a^*a) \subset [\delta, \|a^*a\|].$$

We now work as in Proposition 9.16. Let $\lambda := (\delta + \|a^*a\|)/2$ so that

$$\sigma_{\mathcal{A}}(a^*a - \lambda) \subset [\delta, \|a^*a\|] - \lambda = \left[-\frac{\ell}{2}, \frac{\ell}{2}\right],$$

$\lambda > \ell/2$ and

$$\sigma_{\mathcal{A}}\left(\frac{a^*a - \lambda}{\lambda}\right) \subset \left[-\frac{\ell}{2\lambda}, \frac{\ell}{2\lambda}\right].$$

From this it follows $\left\|\frac{a^*a-\lambda}{\lambda}\right\| \le \frac{\ell}{2\lambda} =: \gamma < 1$ and therefore

$$\left(1 + \frac{a^*a - \lambda}{\lambda}\right)^{-1} \text{ exists in } \mathcal{A}.$$

From this we conclude that

$$\lambda\left(1 + \frac{a^*a - \lambda}{\lambda}\right) = \lambda a^*a$$

is invertible in $\mathcal{A}$ and so is a^*a. Finally this implies that $a^{-1} = (a^*a)^{-1} a^*$ exists in $\mathcal{A}$ and the proof is complete. ■

Lemma 9.5. *Suppose a is a self-adoint element of a unital C^* — algebra $\mathcal{B}$ (not assumed to be commutative. Then the following two statements hold;*

1. *If $\sigma(a) \subset [0, \infty)$, then for all $t \ge \|a\|$, $\|t - a\| \le t$.*
2. *If there exists $t \ge 0$ such that $\|t - a\| \le t$, then $\sigma(a) \subset [0, \infty)$.*

Proof. Let $\mathcal{A} := C^*(I, a)$. Since $\sigma_{\mathcal{B}}(a) = \sigma_{\mathcal{A}}(a)$, by the continuous form of the spectral theorem we may assume that a is a compactly supported function on $\mathbb{R}$. For the first item we know that $0 \le a \le t$ and hence $0 \le t - a \le t$ which implies $\|t - a\| \le t$. For the second item, if $\|t - a\| \le t$ then $|t - a| \le t$ pointwise and therefore $t - a \le t$ which implies $a \ge 0$. ■

Corollary 9.6. *If $a, b \in \mathcal{B}$ (unital C^*-algebra) such that $\sigma(a)$ and $\sigma(b)$ are contained in $[0, \infty)$, then $\sigma(a + b) \subset [0, \infty)$.*

Proof. By Lemma 9.5, for $s \ge \|a\|$ and $t \ge \|b\|$ we know that

$$\|s - a\| \le s \text{ and } \|t - b\| \le t.$$

This inequalities along with the triangle inequality then implies,

$$\|(t + s) - (a + b)\| \le \|s - a\| + \|t - b\| \le s + t$$

which, by Lemma 9.5 again, shows $\sigma(a + b) \subset [0, \infty)$. ■

Theorem 9.7. *$\mathcal{A}$ is a unital algebra and $a, b \in \mathcal{A}$, then $\sigma(ab) \setminus \{0\} = \sigma(ba) \setminus \{0\}$. In fact, if $0 \ne \lambda \notin \sigma(ab)$, then*

$$x = \frac{1}{\lambda} + \frac{1}{\lambda} b(\lambda - ab)^{-1} a = \lambda^{-1}\left[1 + b(\lambda - ab)^{-1} a\right]$$

is the inverse to $(\lambda - ba)$.

Proof. To motivate the formula for x, let us note that for $|\lambda|$ large,

$$\begin{aligned}
\lambda(\lambda - ba)^{-1} &= \frac{\lambda}{\lambda - ba} = \frac{1}{1 - \lambda^{-1}ba} = \sum_{n=0}^{\infty} \lambda^{-n}(ba)^n \\
&= 1 + \sum_{n=0}^{\infty} \lambda^{-(n+1)}(ba)^{n+1} = 1 + \frac{1}{\lambda}\sum_{n=0}^{\infty} \lambda^{-n} b(ab)^n a \\
&= 1 + \frac{1}{\lambda} b\left[\sum_{n=0}^{\infty} \lambda^{-n}(ab)^n\right] a \\
&= 1 + \frac{1}{\lambda} b \frac{1}{1 - \lambda^{-1}ab} a = 1 + b\frac{1}{\lambda - ab} a
\end{aligned}$$

and so we expect that

$$(\lambda - ba)^{-1} = \frac{1}{\lambda}\left[1 + b\frac{1}{\lambda - ab} a\right].$$

For the formal proof we need only shows $x(\lambda - ba) = 1 = (\lambda - ba)x$. For example,

$$\begin{aligned}
x(\lambda - ba) &= \lambda^{-1}\left[(\lambda - ba) + b(\lambda - ab)^{-1} a(\lambda - ba)\right] \\
&= \lambda^{-1}\left[(\lambda - ba) + b(\lambda - ab)^{-1}(\lambda - ab)a\right] \\
&= \lambda^{-1}[(\lambda - ba) + ba] = 1
\end{aligned}$$

and

$$
\begin{aligned}
(\lambda - ba)\,x &= \lambda^{-1}(\lambda - ba)\left[1 + b(\lambda - ab)^{-1} a\right] \\
&= \lambda^{-1}\left[(\lambda - ba) + (\lambda - ba)\,b(\lambda - ab)^{-1} a\right] \\
&= \lambda^{-1}\left[(\lambda - ba) + b(\lambda - ab)(\lambda - ab)^{-1} a\right] \\
&= \lambda^{-1}\left[(\lambda - ba) + ba\right] = 1.
\end{aligned}
$$

■

Corollary 9.8. *If a Banach algebra has unit* 1*, then* 1 *cannot be a commutator; i.e.,* $1 \neq [x,y]$ *for any* $x, y \in \mathcal{B}$*.*

Proof. This is because xy and yx have the same spectrum except possibly 0 while if $xy = 1 + yx$ we would have $\sigma(xy) = 1 + \sigma(yx)$. ■

Theorem 9.9 (Positivity in a C^*-algebra). *Let* $a = a^*$ *in a unital* C^**-algebra* $\mathcal{A}$*. Then* a *is positive (i.e.* $a = b^* b$ *for some* $b \in \mathcal{A}$*) iff* $\sigma(a) \subset [0,\infty)$*.*

Proof. We have see that if $\sigma(a) \subset [0,\infty)$ then we $a = (\sqrt{a})^2$ which shows a is positive. So we now need to show that if $a = b^* b$ for some[1] $b \in \mathcal{A}$, then $\sigma(a) \subset [0,\infty)$. To this end let $g(x) = (-x) \vee 0$. We then hope to show $g(a) = 0$ which would then imply that $g|_{\sigma(a)} = 0$ and hence $\sigma(a) \subset [0,\infty)$.

To carry out the proof we consider $c = bg(a)$ so that

$$c^* c = g(a)\, b^* b g(a) = g(a)\, a g(a) = -g(a)^3.$$

Moreover if we write $c = x + iy$ with x and y being self-adjoint in $\mathcal{A}$, then

$$c^* c + c c^* = (x - iy)(x + iy) + (x + iy)(x - iy) = 2\left(x^2 + y^2\right)$$

or equivalently,

$$c c^* = 2\left(x^2 + y^2\right) - c^* c = 2\left(x^2 + y^2\right) + g(a)^3.$$

Since $t \to 2t^2$ and g^3 are positive functions it follows that $2x^2$, $2y^2$, and $g(a)^3$ are self-adjoint with spectrum in $[0,\infty)$ and therefore by Corollary 9.6, $\sigma(cc^*) \subset [0,\infty)$. On the other hand $\sigma(c^* c) = \sigma\left(-g(a)^3\right) \subset (-\infty, 0]$. But an application of Theorem 9.7 with $a = c$ and $b = c^*$, show $\sigma(cc^*) \setminus \{0\} = \sigma(c^* c) \setminus \{0\}$ which along with the previous inclusions shows $\sigma(c^* c) = \{0\}$ which implies $\|c\|^2 = \|c^* c\| = r(c^* c) = 0$. Therefore $g^3(a) = 0$ and as $g^3 > 0$ on $(-\infty, 0)$ it follows that $\sigma(a) \subset [0,\infty)$. ■

[1] If b is assumed to be normal, life is easier as we saw in Corollary 7.36.

Corollary 9.10 (Polar Decomposition). *Suppose that* $x \in \mathcal{B}$ *is an invertible element of a unital* C^**-algebra. Then there exists a unique* $u \in \mathcal{B}$ *and Hermitian* $a \in \mathcal{B}$ *with* $\sigma(a) \subset [0,\infty)$ *such that* $x = ua$*.*

Proof. If such a decomposition exists we must have

$$x^* x = a u^* u a = a^2$$

and so $a = \sqrt{x^* x}$. In order for this to make sense we are going to need to know that $\sigma(x^* x) \subset [0,\infty)$. ■

9.1 Alternate Proofs

The goal of this section is to give an elementary proof of Proposition **??**, i.e. without the aid of the spectral theorem. We need to do some preparation first which is of interest in its own right. The results in this section could be proved using the spectral theorem as shown in Proposition **??** below.

Proposition 9.11. *If* $A = A^*$*, then* $\sigma(A) \subset \left[-\|A\|_{op}, \|A\|_{op}\right]$ *and if* $A \geq 0$*, then* $\sigma(A) \subset \left[0, \|A\|_{op}\right]$ *and moreover*

$$\left\|(A + \lambda)^{-1}\right\| \leq \lambda^{-1} \text{ for all } \lambda > 0. \tag{9.1}$$

Proof. First proof. By the spectral theorem we may assume there exists a probability space, $(\Omega, \mathcal{F}, \mu)$ and a bounded measurable function, f, on Ω so that $A = M_f$ acting on $H = L^2(\mu)$. We then know $A = A^*$ iff f is real a.e. and

$$\sigma(A) = \mathrm{essran}_\mu(f) \subset \left[-\|f\|_\infty, \|f\|_\infty\right] = \left[-\|A\|_{op}, \|A\|_{op}\right].$$

Moreover, $A \geq 0$ iff $\int_\Omega f|g|^2 d\mu \geq 0$ for all $g \in L^2(\mu)$ and this then implies by taking $g = 1_A$, that $\int_\Omega f 1_A d\mu \geq 0$ for all $A \in \mathcal{F}$. This last assertion is equivalent to $f \geq 0$ a.e. and hence

$$\sigma(A) = \sigma(M_f) \subset \left[0, \|f\|_\infty\right] = \left[0, \|A\|_{op}\right].$$

Finally if $\lambda > 0$, then $f + \lambda \geq \lambda$ a.e. and therefore

$$0 \leq \frac{1}{f + \lambda} \leq \frac{1}{\lambda} \text{ a.e.}$$

which implies $\left\|(A + \lambda)^{-1}\right\| = \left\|\frac{1}{f+\lambda}\right\|_\infty \leq \frac{1}{\lambda}$.

Second proof. If $A = A^*$ and $\lambda = a + ib \in \mathbb{C}$, then

$$\begin{aligned}\|(A-\lambda) f\|^2 &= \|(A-a) f - ibf\|^2 \\ &= \|(A-a) f\|^2 - 2\operatorname{Re}\langle (A-a) f, ibf\rangle + b^2\|f\|^2 \\ &= \|(A-a) f\|^2 + b^2\|f\|^2 \geq b^2\|f\|^2.\end{aligned}$$

Hence if $b \neq 0$, then $\operatorname{Ran}(A-\lambda)$ is closed and hence

$$\operatorname{Ran}(A-\lambda) = \operatorname{Nul}\left(A - \bar{\lambda}\right)^{\perp} = \{0\}$$

where the latter assertion follows from the inequality we have proved with λ replaced by $\bar{\lambda}$. Thus we see that $A - \lambda$ is a bijection with

$$\left\|(A-\lambda)^{-1}\right\|_{op} \leq (\operatorname{Im}\lambda)^{-1} < \infty.$$

So we have shown $\sigma(A) \subset \mathbb{R}$ and this completes the proof that $\sigma(A) \subset \left[-\|A\|_{op}, \|A\|_{op}\right]$, since (as always) $\sigma(A) \subset \overline{D\left(0, \|A\|_{op}\right)}$.

If we further assume that $A \geq 0$ and $\lambda > 0$, then we have

$$\begin{aligned}\|(A+\lambda) f\|^2 &= \|Af + \lambda f\|^2 \\ &= \|Af\|^2 + 2\operatorname{Re}\langle Af, \lambda f\rangle + \lambda^2\|f\|^2 \\ &\geq \|Af\|^2 + \lambda^2\|f\|^2 \geq \lambda^2\|f\|^2.\end{aligned}$$

The same argument as above now shows that $(A+\lambda)^{-1}$ exists and Eq. (9.1) holds. ■

Lemma 9.12. *Suppose $A \in B(H)$ with $A \geq 0$ (this means $A = A^*$ and $\langle Ax, x\rangle \geq 0$ for all $x \in H$), then*

1. $\operatorname{Nul}(A) = \{x \in H : \langle x, Ax\rangle = 0\}$.
2. $\operatorname{Nul}(A) = \operatorname{Nul}(A^2)$.
3. *If $A, B \in B(H)$ are two positive operators then* $\operatorname{Nul}(A+B) = \operatorname{Nul}(A) \cap \operatorname{Nul}(B)$.

Proof. Items 2. and 3. are fairly easy and will be left to the reader.

To prove Item 1., it suffices to show $\{x \in H : \langle x, Ax\rangle = 0\} \subset \operatorname{Nul}(A)$ since the reverse inclusion is trivial. For sake of contradiction suppose there exists $x \neq 0$ such that $y = Ax \neq 0$ and $\langle x, Ax\rangle = 0$. Using $x \perp y$, we have for $\lambda \in \mathbb{R}$ that

$$\begin{aligned}\langle (x+\lambda y), A(x+\lambda y)\rangle &= \langle x + \lambda y, y + \lambda Ay\rangle \\ &= \lambda\langle x, Ay\rangle + \lambda\|y\|^2 + \lambda^2\langle Ay, y\rangle \\ &= 2\lambda\|y\|^2 + \lambda^2\langle Ay, y\rangle.\end{aligned}$$

From this expression we easily deduce that

$$0 \leq \langle (x+\lambda y), A(x+\lambda y)\rangle < 0$$

for all $\lambda < 0$ sufficiently close to zero which is a contradiction. But this contradicts the positivity of A. ■

We can make item 1. of Lemma 9.12 more quantitative as follows.

Theorem 9.13. *If $A \geq 0$ and A^{-1} exists then there exists $\delta > 0$ so that $A \geq \delta I$. Conversely if $A \geq \delta I$ for some $\delta > 0$ then A^{-1} exists and $\left\|A^{-1}\right\|_{op} \leq \delta^{-1}$.*

Proof. First proof. By Proposition 9.11 we know that $\sigma_{B(H)}(A) \subset [0,\infty)$ and by Corollary 7.34 it follows that $\sigma_{C^*(A,I)}(A) = \sigma_{B(H)}(A) \subset [0,\infty)$. Hence we may apply Corollary 7.36 (or Corollary **??** below) with $\mathcal{B} = C^*(A, I)$ to find Hence we may apply to find $B \in C^*(A, I)$ so that $B = B^*$, $\sigma_{B(H)}(B) = \sigma_{C^*(A,I)}(B) \subset [0,\infty)$ and $A = B^2$. Thus if $A \geq \delta I$, then $\|Bx\|^2 = \langle Ax, x\rangle \geq \delta\|x\|^2$ implies B^{-1} exists in $B(H)$ and $\left\|B^{-1}\right\| \leq \frac{1}{\sqrt{\delta}}$ and therefore A^{-1} exists and

$$\left\|A^{-1}\right\| = \left\|\left(B^{-1}\right)^2\right\| \leq \frac{1}{\delta}.$$

Conversely if A^{-1} exists then B^{-1} exists and there exists $\delta > 0$ so that $\|Bx\|^2 \geq \delta\|x\|^2$ for all $x \in H$. As in the above argument, it now follows that $A \geq \delta I$.

Second proof. [This proof uses no C^*-algebra technology.] Suppose that A^{-1} exists and let $\varepsilon := \left\|A^{-1}\right\|_{op}^{-1} > 0$. We then have

$$\begin{aligned}&\left\|A^{-1}f\right\| \leq \left\|A^{-1}\right\|_{op}\|f\| \implies \|f\| \leq \left\|A^{-1}\right\|_{op}\|Af\|, \text{ i.e.} \\ &\|Af\| \geq \varepsilon\|f\| \text{ for all } f \in H.\end{aligned}$$

If $\lambda \geq 0$ and $f \in H$ with $\|f\| = 1$, then

$$\begin{aligned}0 &\leq \langle A(f - \lambda Af), (f - \lambda Af)\rangle = \left\langle Af - \lambda A^2 f, f - \lambda Af\right\rangle \\ &= \langle Af, f\rangle - 2\lambda\|Af\|^2 + \lambda^2\left\langle A^3 f, f\right\rangle \\ &\leq \langle Af, f\rangle - 2\lambda\varepsilon^2 + \lambda^2\|A\|_{op}^3.\end{aligned}$$

Minimizing the right side of this inequality by taking $\lambda = \varepsilon^2/\|A\|_{op}^3$ shows

$$\langle Af, f\rangle \geq \frac{\varepsilon^4}{\|A\|_{op}^3} = \frac{1}{\|A\|_{op}^3\left\|A^{-1}\right\|_{op}^4} =: \delta > 0.$$

Conversely if $A \geq \delta I$, then $A - \delta I \geq 0$ and (by Proposition 9.11) it follows that $\sigma(A - \delta I) \subset [0,\infty)$, i.e. $\sigma(A) \subset [\delta,\infty)$. Hence $0 \notin \sigma(A)$, i.e. A^{-1} exists. Moreover,

$$\delta\|f\|^2 \leq \langle Af, f\rangle \leq \|Af\|\|f\| \implies \|Af\| \geq \delta\|f\|$$

from which it follows that $\left\|A^{-1}\right\|_{op} \leq \delta^{-1}$. ■

Lemma 9.14. *If $\mathcal{B}$ is a C^*-sub-algebra of $B(H)$ with unit and $B \in \mathcal{B}$ with $B \geq 0$, then $(I+B)^{-1} \in \mathcal{B}$. Actually what is proved here is; if $B \in \mathcal{B}(H)$ with $B \geq 0$, then $(I+B)^{-1} \in C^*(I, B)$.*

Proof. Let $A_\lambda := I + \lambda B$ for all $\lambda \geq 0$. As $A_\lambda \geq I$ we know that A_λ^{-1} exists in $B(H)$ by Theorem 9.13. Moreover if $0 \leq \lambda < \|B\|_{op}^{-1}$, then

$$A_\lambda^{-1} = \sum_{n=0}^{\infty} (-1)^n \lambda^n B^n \in \mathcal{B}.$$

Suppose that we have shown that $A_\lambda \in \mathcal{B}$ for some $\lambda > 0$. Then for $\varepsilon > 0$ we have

$$A_{\lambda+\varepsilon} = A_\lambda + \varepsilon B = A_\lambda \left[I + \varepsilon A_\lambda^{-1} B\right]$$

where $A_\lambda^{-1} B \in \mathcal{B}$. Thus if $\varepsilon \left\|A_\lambda^{-1} B\right\| < 1$, then

$$\left[I + \varepsilon A_\lambda^{-1} B\right]^{-1} = \sum_{n=0}^{\infty} (-1)^n \lambda^n \left(B A_\lambda^{-1}\right)^n \in \mathcal{B}$$

and so $A_{\lambda+\varepsilon}^{-1} = \left[I + \varepsilon A_\lambda^{-1} B\right]^{-1} A_\lambda^{-1} \in \mathcal{B}$. On the other hand, from Proposition 9.11,

$$\begin{aligned} \left\|A_\lambda^{-1}\right\| &= \left\|(I+\lambda B)^{-1}\right\| = \lambda^{-1} \left\|\left(B + \frac{1}{\lambda}\right)^{-1}\right\| \\ &\leq \lambda^{-1} \left(\lambda^{-1}\right)^{-1} = 1 \text{ for all } \lambda \geq 0 \end{aligned}$$

and so

$$\varepsilon \left\|A_\lambda^{-1} B\right\| \leq \varepsilon \|B\|$$

and the previous construction works provided $\varepsilon < \|B\|_{op}^{-1}$ where the bound on ε is independent of λ! Putting this all together if we fix $n \in \mathbb{N}$ so that $1/n < \varepsilon$, then we can show inductively that $A_{k/n} \in \mathcal{B}$ for $k = 1, 2, \dots$ and hence $A_1 \in \mathcal{B}$. ■

Corollary 9.15. *Let $\mathcal{B}$ be a C^*-sub-algebra of $B(H)$ with identity and $A \in \mathcal{B}$ with $A \geq 0$. If A^{-1} exists in $B(H)$ then $A^{-1} \in \mathcal{B}$. [In other words, if $A \in \mathcal{B}$ and $A \geq 0$, then invertibility of A in $B(H)$ and in $\mathcal{B}$ are the same notions.]*

Proof. By Theorem 9.13, if A^{-1} exists there exists $\delta > 0$ so that $A \geq \delta I$. We may replace A by $\delta^{-1} A$ and henceforth assume that $A \geq I$. Then $B := A - I \in \mathcal{B}$ and $A = I + B$ with $B \geq 0$. It now follows from Lemma 9.14 that $A^{-1} = (I+B)^{-1} \in \mathcal{B}$. ■

Proposition 9.16. *Suppose that A is a self-adjoint operator on a Hilbert space, H, such that $A \geq \varepsilon I$ for some $\varepsilon > 0$. Then A^{-1} exists and if M is chosen so that $\varepsilon I \leq A \leq MI$ and $\lambda := (M+\varepsilon)/2$, then*

$$A^{-1} = \sum_{n=-0}^{\infty} \frac{1}{\lambda^{n+1}} (\lambda I - A)^n.$$

Proof. The key idea of the proof is to shift A by some $\lambda \in \mathbb{R}$ in such a way that we make $\|A - \lambda I\|_{op}$ as small as possible. To this end let $\alpha := (M-\varepsilon)/2$ and set $\lambda = (M+\varepsilon)/2 = \varepsilon + \alpha = M - \alpha$ and note that

$$A = \lambda I + (A - \lambda) = \lambda \left[I + \frac{1}{\lambda}(A - \lambda)\right]$$

where

$$\frac{1}{\lambda}(\varepsilon - \lambda) I \leq \frac{1}{\lambda}(A - \lambda) \leq \frac{1}{\lambda}(M - \lambda) I$$

and

$$\frac{1}{\lambda}(\varepsilon - \lambda) = -\frac{\alpha}{\varepsilon + \alpha} \text{ and } \frac{1}{\lambda}(M - \lambda) = \frac{\alpha}{\varepsilon + \alpha}$$

and hence

$$-\frac{\alpha}{\varepsilon+\alpha} I \leq \frac{1}{\lambda}(A - \lambda) \leq \frac{\alpha}{\varepsilon+\alpha} I.$$

From this we conclude that

$$\left\|\frac{1}{\lambda}(A - \lambda)\right\|_{op} \leq \frac{\alpha}{\varepsilon + \alpha} < 1$$

and hence $I + \frac{1}{\lambda}(A - \lambda)$ is invertible and moreover,

$$A^{-1} = \frac{1}{\lambda} \sum_{n=-0}^{\infty} \frac{1}{\lambda^n} (\lambda I - A)^n = \sum_{n=-0}^{\infty} \frac{1}{\lambda^{n+1}} (\lambda I - A)^n.$$

■

Corollary 9.17. *If $T \in B(H)$ is an invertible operator, then there exists $\lambda \in (0, \infty)$ such that*

$$T^{-1} = \sum_{n=-0}^{\infty} \frac{1}{\lambda^{n+1}} (\lambda I - T^* T)^n T^* \tag{9.2}$$

and in particular $T^{-1} \in C^*(I, T)$ where $C^*(I, T)$ is the C^*–algebra generated by $\{I, T\}$, i.e. $C^*(I, T)$ is the smallest Banach sub-algebra of $B(H)$ containing $\{I, T, T^*\}$.

Proof. Let $\varepsilon > 0$ be defined so that $\left\|T^{-1}\right\|_{op} = \varepsilon^{-1/2}$, i.e.

$$\left\|T^{-1}v\right\|^2 \le \varepsilon^{-1} \|v\|^2 \text{ for all } v \in H.$$

Replacing v by Tv in this inequality and then multiplying by ε shows,

$$\varepsilon \|v\|^2 \le \|Tv\|^2 = \langle T^*Tv, v\rangle .$$

Hence $A = T^*T$ is a self-adjoint operator such such $A \ge \varepsilon I$ for some $\varepsilon > 0$, hence by Proposition 9.16, there exists $\lambda > \varepsilon > 0$ such that

$$A^{-1} = \sum_{n=-0}^{\infty} \frac{1}{\lambda^{n+1}} (\lambda I - A)^n = \sum_{n=-0}^{\infty} \frac{1}{\lambda^{n+1}} (\lambda I - T^*T)^n .$$

Equation (9.2) follows from this equation and the observation that $T^{-1} = A^{-1}T^*$. ∎

As a consequence of Corollary 9.17, if $T \in B(H)$, then

$$\sigma_{B(H)}(T) = \sigma_{C^*(I,T)}(T) .$$

9.2 The Spectral Theorem Again

In this section, let H be a complex Hilbert space. Our goal here is to give another C^*-algebra style proof of the spectral theorem.

9.2.1 First variant of the spectral theorem proof.

Theorem 9.18. *Let H be a Hilbert space, $\mathcal{B}$, be a commutative C^*-subalgebra of $B(H)$ with identity, $x \in H \setminus \{0\}$, and $H_x := \overline{\mathcal{B}x}^H$. Then there exists a compact Hausdorff space, X, a Radon measure μ on X, and a unitary map, $U : L^2(X,\mu) \to H_x$, and $f_A \in L^\infty(X,\mu)$ such that $U^*AU = M_{f_A}$ for all $A \in \mathcal{B}$.*

Proof. Let $X := \tilde{\mathcal{B}} = \operatorname{spec}(\mathcal{B})$ and $f_A := \hat{A} \in C\left(\tilde{\mathcal{B}}\right)$ for all $A \in \mathcal{B}$. To construct the measure, μ, let Λ be the linear functional on $C\left(\tilde{\mathcal{B}}\right)$ defined by,

$$\Lambda\left(\hat{A}\right) := \langle Ax, x\rangle \text{ for all } A \in \mathcal{B}.$$

If $\hat{A} \ge 0$ then $\sqrt{\hat{A}} \in C\left(\tilde{\mathcal{B}}\right)$ and $\sqrt{\hat{A}} = \hat{B}$ for some $B \in \mathcal{B}$ and moreover, $B = B^*$ and $B^2 = A$ since $\widehat{B^2} = \hat{B}^2 = \hat{A}$. Therefore,

$$\Lambda\left(\hat{A}\right) = \left\langle B^2x, x\right\rangle = \langle Bx, Bx\rangle = \|Bx\|^2$$

which shows Λ is a positive linear functional on $C\left(\tilde{\mathcal{B}}\right)$. An application of the Riesz-Markov theorem (see [41, Theorem 3.14, p. 69]) there exists a unique Radon measure such that $\Lambda(f) = \int_X f d\mu$ for all $f \in C(X)$, i.e.

$$\langle Ax, x\rangle = \Lambda\left(\hat{A}\right) = \int_{\tilde{\mathcal{B}}} \hat{A} d\mu \ \forall\ A \in \mathcal{B}.$$

Let us further observe that

$$\int_{\tilde{\mathcal{B}}} \left|\hat{A}\right|^2 d\mu = \int_{\tilde{\mathcal{B}}} \widehat{A^*A} d\mu = \langle A^*Ax, x\rangle = \|Ax\|^2 . \tag{9.3}$$

From this identity it follows that $\hat{A} = 0$ μ – a.e. iff $Ax = 0$. Thus we may define a linear operator, $U_0 : C\left(\tilde{\mathcal{B}}\right) \to H_x \subset H$ by defining

$$U_0\hat{A} := Ax \text{ for all } A \in \mathcal{B}.$$

This map is an isometry on the dense subspace,[2] $C(X)$, of $L^2(X,\mu)$ and therefore extends uniquely to an isometry $U : L^2(X,\mu) \to H_x$. As the $\mathcal{B}x = \operatorname{Ran}(U_0) \subset \operatorname{Ran}(U)$, $\mathcal{B}x$ is dense in H_x, and $\operatorname{Ran}(U)$ is complete and hence closed, we conclude that $\operatorname{Ran}(U) = H_x$. This shows $U : L^2(X,\mu) \to H_x$ is unitary. Finally if $A, B \in \mathcal{B}$, then

$$U^*AU\hat{B} = U^*ABx = \widehat{AB} = \hat{A}\hat{B} = M_{\hat{A}}\hat{B}.$$

Since $C\left(\tilde{\mathcal{B}}\right)$ is dense in $L^2(\mu)$ we may conclude that $U^*AU = M_{\hat{A}}$ and the proof is complete. ∎

Corollary 9.19. *Let H be a separable Hilbert space and $\mathcal{B}$ be a commutative C^*-subalgebra of $B(H)$. Then there exists a finite measure space $(X, \mathcal{F}, \mu)$, $f_A \in L^\infty(\mu)$ for all $A \in \mathcal{B}$, and a unitary map, $U : L^2(\mu) \to H$, such that $U^*AU = M_{f_A}$ for all $A \in \mathcal{B}$.*

Proof. Follow the same proof strategy as in the proof of Corollary 7.40 with Theorem 9.18 playing the role of Theorem **??**. ∎

[2] Recall that $C(X)$ is a dense subspace of $L^p(X,\mu)$ for an $1 \le p < \infty$, see [41, Theorem 3.14, p. 69].

9.2.2 Second variant of the spectral theorem proof.

In this section we are going to give another variant of the proof of Corollary 9.19. The key idea is to take $\mathcal{B}$ and embedded it in a larger (maximal) commutative C^*-subalgebra $\mathcal{A}$ of $B(H)$. We then show that $\mathcal{A}$ has cyclic vector and therefore we will get the result of Corollary 9.19 by directly applying Theorem 9.18 with $\mathcal{B}$ replaced by $\mathcal{A}$ as long as we choose x to be cyclic vector for $\mathcal{A}$.

Notation 9.20 *If $S \subset B(H)$ then*

$$S' = \{A \in B(H) : AB = BA \ \forall B \in S\}.$$

S' is clearly a subalgebra of $B(H)$ for any set S. S' is called the ***commutator algebra*** *of S.*

Remark 9.21. Recall from Proposition **??** that S' is w.o.t. closed and hence also s.o.t. and operator norm closed. It is of course easy to directly verify that S' is closed under operator norm convergence. Indeed, if $A_n \in S'$ and $A \in B(H)$ such $\|A - A_n\|_{op} \to 0$, then for any $B \in S$,

$$[B, A] = \lim_{n\to\infty} [B, A_n] = 0 \text{ for all } B \in S$$

which shows $A \in S'$.

Also observe that if S is $*-$ closed then so is S'. Indeed, if $A \in S'$, then

$$[A^*, B] = -[A, B^*]^* = 0 \text{ for all } B \in S.$$

Definition 9.22. *A* ***maximal abelian algebra*** *on H is a commutative subalgebra, $\mathcal{A} \subset B(H)$, which is not contained in any larger commutative subalgebra of $B(H)$.*

Proposition 9.23. *Let H be a Hilbert space and $\mathcal{A}$ be a sub-algebra of $B(H)$. Then;*

1. *$\mathcal{A} \subset B(H)$ is a maximal abelian subalgebra iff $\mathcal{A}' = \mathcal{A}$.*
2. *If $\mathcal{A}$ is maximal abelian then $\mathcal{A}$ is operator norm closed. [More generally $\mathcal{A}$ is w.o.t. and s.o.t. closed.]*

Proof. We consider each item separately.

1. Suppose $\mathcal{A}$ is a maximal abelian algebra and $B \in \mathcal{A}'$. Then the algebra generated by $\mathcal{A} \cup \{B\}$ consisting of operators of the form
$$A_0 + A_1 B + A_2 B^2 + \cdots + A_n B^n \text{ with } A_j \in \mathcal{A}$$
is a commutative algebra containing $\mathcal{A}$ and therefore it is $\mathcal{A}$. Thus we have shown $\mathcal{A}' \subset \mathcal{A}$ and as $\mathcal{A}$ is commutative we also have $\mathcal{A} \subset \mathcal{A}'$, i.e. $\mathcal{A}' = \mathcal{A}$. Conversely, if $\mathcal{A}$ is not maximal abelian then there exist a commutative algebra, $\mathcal{B} \subset B(H)$, such that $\mathcal{A} \subsetneqq \mathcal{B}$. As $\mathcal{B} \subset \mathcal{A}'$ it follows that $\mathcal{A} \subsetneqq \mathcal{A}'$.
2. This follows from item 1. and Remark 9.21.

■

Definition 9.24. *A* ***maximal abelian self–adjoint (m.a.s.a.) algebra*** *on H is a commutative $*$-subalgebra, $\mathcal{A} \subset B(H)$, which is not contained in any larger commutative $*$-subalgebra of $B(H)$.*

Proposition 9.25. *Let H be a Hilbert space and $\mathcal{A}$ be a $*$-subalgebra of $B(H)$.*

1. *$\mathcal{A}$ is a maximal abelian self-adjoint algebra iff $\mathcal{A}' = \mathcal{A}$.*
2. *A m.a.s.a. algebra $\mathcal{A}$ is a C^*-algebra.*[3]

Proof. 1. If $\mathcal{A}' = \mathcal{A}$, then $\mathcal{A}$ is maximal abelian by Proposition 9.23 and in particular $\mathcal{A}$ must be a m.a.s.a. Conversely if $\mathcal{A}$ is m.a.s.a. and $B \in \mathcal{A}'$, then $B^* \in \mathcal{A}'$ by Remark 9.21 and hence

$$X := \operatorname{Re} B = \frac{1}{2}(B + B^*) \in \mathcal{A}' \text{ and}$$
$$Y := \operatorname{Im} B = \frac{1}{2i}(B - B^*) \in \mathcal{A}'.$$

Since X and Y are self-adjoint, the algebras generated by $\mathcal{A} \cup \{X\}$ and $\mathcal{A} \cup \{Y\}$ are both commutative self-adjoint algebras containing $\mathcal{A}$ and hence must be $\mathcal{A}$. This shows that both $X, Y \in \mathcal{A}$ and hence $B = X + iY \in \mathcal{A}$. Thus we have shown $\mathcal{A}' \subset \mathcal{A}$ and therefore $\mathcal{A}' = \mathcal{A}$ as we always have $\mathcal{A} \subset \mathcal{A}'$ when $\mathcal{A}$ is commutative.

2. If $\mathcal{A}$ is m.a.s.a., then $\mathcal{A} = \mathcal{A}'$ is operator norm closed by Remark 9.21 and is $*$-closed by definition, i.e. $\mathcal{A}$ is a C^*-subalgebra of $B(H)$. ■

Definition 9.26. *Let (X, μ) be a measure space. The* ***multiplication algebra*** *(denoted by $\mathcal{M}(X, \mu)$) of (X, μ) is the algebra of operators on $L^2(X, \mu)$ consisting of all M_f, $f \in L^\infty$.*

The next proposition is essentially a repeat of item 1. of Proposition **??** below.

Proposition 9.27. *If (X, μ) is a σ–finite measure space, then $\mathcal{M}(X, \mu)$ is a m.a.s.a. algebra.*

Proof. Assume first $\mu(X) < \infty$. Write $\mathcal{M} = \mathcal{M}(X, \mu)$ and assume $T \in \mathcal{M}'$. Let $g = T(1)$. If $f \in L^\infty$ then $TM_f 1 = M_f T1$. Therefore $T(f) = fg$. Thus $Tf = M_g f$ for f in L^∞. The proof in the preceding example shows $\|g\|_\infty \le \|T\|$. Since M_g is bounded the equation $T \mid L^\infty = M_g \mid L^\infty$, already established,

[3] It also a von Neumann algebra.

extends by continuity to L^2. Hence $T \in \mathcal{M}$ and $\mathcal{M}$ is maximal abelian. Since $M_g^* = M_{\bar{g}}$, $\mathcal{M}$ is self–adjoint.

In the general case, write $X = \cup_{j=1}^\infty X_j$, where the X_j are disjoint subsets of finite measure. If T is in $\mathcal{M}'$ it commutes with $M_{\chi_{X_j}}$ and therefore leaves invariant the subspace $L^2(X_j)$ which we identify with $\{f \in L^2(X) : f = 0 \text{ off } X_j\}$. Apply the finite measure case and piece together the result to get the general case. ∎

Definition 9.28. *If $\mathcal{A}$ is a subalgebra of $B(H)$ a vector x in H is called a* ***cyclic****vector for $\mathcal{A}$ if $\mathcal{A}x \equiv \{Ax : A \in \mathcal{A}\}$ is dense in H.*

Lemma 9.29. *Let $\mathcal{A}$ be any $*$ subalgebra of $B(H)$. Suppose K is a closed subspace of H and P is the projection on K. Then K is invariant under $\mathcal{A}$ iff $P \in \mathcal{A}'$.*

Proof. (This result may be deduced rather quickly from Exercise 5.2. Nevertheless, we will give the proof here for completeness.)

($\Leftarrow$) If $P \in \mathcal{A}'$ and $x \in K$, then

$$Ax = APx = PAx \in K \;\forall\; A \in \mathcal{A}.$$

($\Rightarrow$) If $\mathcal{A}K \subset K$, then for any $x \in H$, $Px \in K$ and so $\mathcal{A}Px \in K$. Thus it follows that $APx = PAPx$ for all $A \in \mathcal{A}$ and $x \in H$, i.e. $AP = PAP$ for all $A \in \mathcal{A}$. Since $A^* \in \mathcal{A}$ for $A \in \mathcal{A}$ we also have $A^*P = PA^*P$ for all $A \in \mathcal{A}$. Using these observations we find,

$$PA = P^*A = (A^*P)^* = (PA^*P)^* = PAP = AP \;\forall\; A \in \mathcal{A},$$

i.e. $P \in \mathcal{A}'$.

Alternative proof. Let $x, y \in H$ and $A \in \mathcal{A}$, then using $APx \in K$ and $A^*Py \in K$ we find,

$$\begin{aligned}\langle APx, y\rangle &= \langle APx, Py\rangle = \langle Px, A^*Py\rangle \\ &= \langle x, A^*Py\rangle = \langle PAx, y\rangle .\end{aligned}$$

As $x, y \in H$ were arbitrary we have shown $[A, P] = 0$ so that $P \in \mathcal{A}'$. ∎

Lemma 9.30. *If H is separable and $\mathcal{A}$ is a m.a.s.a. on H then $\mathcal{A}$ has a cyclic vector.*

Proof. For any $x \in H$, let $\overline{\mathcal{A}x}$ be the closed subspace containing $\mathcal{A}x$. Since $I \in \mathcal{A}$, $x \in \mathcal{A}x$. Since $\mathcal{A}x$ is invariant under $\mathcal{A}$, so is $\overline{\mathcal{A}x}$. If $y \perp \mathcal{A}x$ then $\mathcal{A}y \perp \mathcal{A}x$ since $\langle Ay, Bx\rangle = \langle y, A^*Bx\rangle = 0$. Let $E = \{x_\alpha\}$ be an orthonormal set such that $\mathcal{A}x_\alpha \perp \mathcal{A}x_\beta$ if $\alpha \neq \beta$. Such sets exist (e.g. singletons). Zorn's lemma gives us a maximal such set. For this E, $H = \text{closed span}_\alpha\{\mathcal{A}x_\alpha\}$ for otherwise we could adjoin to E any unit vector in $(\text{span}\{\mathcal{A}x_\alpha\})^\perp$. Now, since H is separable, E is countable; $E = \{x_1, x_2, \ldots\}$ put $z = \sum_{n=1}^\infty 2^{-n}x_n$.

Claim: z is a cyclic vector for $\mathcal{A}$. To prove this recall from Lemma 9.29 that the orthogonal projection operator, P_n, from H onto $\overline{\mathcal{A}x_n}$ is in $\mathcal{A}' = \mathcal{A}$. Therefore,

$$\mathcal{A}z \supset \mathcal{A}P_n z = \mathcal{A}2^{-n}x_n = \mathcal{A}x_n \;\forall n \in \mathbb{N}$$

and hence

$$H = \text{closed span}_n\{\mathcal{A}x_n\} \subset \overline{\mathcal{A}z}.$$

∎

Theorem 9.31. *Let $\mathcal{A}$ be a m.a.s.a. on separable Hilbert space H. Then there exists finite measure space (X, μ) and a unitary operator $U : H \twoheadrightarrow L^2(X, \mu)$ such that $U\mathcal{A}U^{-1} = \mathcal{M}(X, \mu)$.*

Proof. Let z be a unit cyclic vector for $\mathcal{A}$ and then apply Theorem 9.18 with $\mathcal{B}$ replaced by $\mathcal{A}$ and $x = z$ in order to find a Radon measure, μ , on $x := \widetilde{\mathcal{A}}$ such that

$$UAU^{-1} = M_{\hat{A}} \text{ for all } A \in \mathcal{A}. \tag{9.4}$$

Let $\mathcal{N} = U\mathcal{A}U^{-1}$ and let $\mathcal{M} = \{M_f : f \in L^\infty(X, \mu)\}$ be the multiplication algebra of (X, μ). Then

$$\mathcal{N}' = \left[U\mathcal{A}U^{-1}\right]' = \mathcal{N} = U\mathcal{A}'U^{-1} = \mathcal{N} = U\mathcal{A}U^{-1} = \mathcal{N}$$

and $\mathcal{N} \subset \mathcal{M}$ by Eq. (9.4). Therefore, (using Proposition 9.27) we find

$$\mathcal{M} = \mathcal{M}' \subset \mathcal{N}' = \mathcal{N} \subset \mathcal{M},$$

i.e. $\mathcal{M} = \mathcal{N}$.

∎

Remark 9.32. The compact Hausdorff space, $X = \text{spec}(\mathcal{A})$ in the above proof is rather pathological. It has the bizarre property that every element, $f \in L^\infty(X, \mu)$, has a continuous representative!

Theorem 9.33 (Spectral Theorem). *Let $\{A_\alpha\}_{\alpha \in I}$ be a family of bounded normal operators on a complex separable Hilbert space. Assume that the family is a commuting set in the sense that:*

$$A_\alpha A_\beta = A_\beta A_\alpha \quad \forall \alpha, \beta$$

and

$$A_\alpha A_\beta^* = A_\beta^* A_\alpha \quad \forall \alpha, \beta$$

Then there exists a finite measure space (X, μ) and a unitary operator $U : H \to L^2(X, \mu)$ and for each α there exists a function $f_\alpha \in L^\infty$ such that

$$UA_\alpha U^{-1} = M_{f_\alpha}.$$

Proof. Let $\mathcal{A}_0$ be the algebra generated by the $\{A_\alpha, A_\alpha^*\}_{\alpha \in I}$. Then $\mathcal{A}_0$ is a commutative $*$ algebra. Order the set of all commutative self-adjoint algebras containing $\mathcal{A}_0$ by inclusion. By Zorn's lemma there exists a largest such algebra, $\mathcal{A}$. [Bruce does not see the need for the following argument as it seems to me that clearly $\mathcal{A}$ is m.a.s.a. by construction. We assert that $\mathcal{A} = \mathcal{A}'$. Indeed if $B \in \mathcal{A}'$ then $B^* \in \mathcal{A}'$ also because $\mathcal{A}$ is self-adjoint. Hence $C := B + B^* \in \mathcal{A}'$. But the algebra generated by $\mathcal{A}$ and C is commutative and self-adjoint. Therefore $C \in \mathcal{A}$. Similarly $i(B - B^*) \in \mathcal{A}$. Hence $B \in \mathcal{A}$. So $\mathcal{A}' = \mathcal{A}$. Therefore $\mathcal{A}$ is maximal abelian and self-adjoint.]

Now by the Theorem 9.31, there exists a measure space (X, μ) with $\mu(X) = 1$ and a unitary operator, $U : H \to L^2(X)$, such that $U\mathcal{A}U^{-1} = \mathcal{M}(X, \mu)$. Therefore $UA_\alpha U^{-1} = M_{f_\alpha}$ for some $f_\alpha \in L^\infty$. ■

*Projection Valued Measure Spectral Theorem

It is natural to put this material here but in fact I will likely be able to avoid using projection valued measures in the remainder of these notes. So for the time being you may consider this chapter optional as well. [This chapter needs editing!]

Recall that if A is a $n \times n$ self-adjoint matrix, then one may express the spectral theorem as,

$$A = \sum_{\lambda \in \sigma(A)} \lambda E_\lambda$$

where E_λ is orthogonal projection onto $\mathrm{Nul}\,(A - \lambda I)$ for all $\lambda \in \sigma(A)$. The goal of this chapter is to rewrite the general spectral theorem in this same form. The statement we are aiming for is if $A \in B(H)$ is self-adjoint, then there exists a unique "projection valued measure," $E(\cdot)$, on $\sigma(A)$ so that

$$A = \int_{\sigma(A)} \lambda dE(\lambda).$$

Clearly to make sense of this assertion we have to develop the notion of projection valued measures.

10.1 Projection valued measures

Definition 10.1. *A sequence A_n of bounded operators on a Banach space B **converges strongly** to a bounded operator A if $A_n x \to Ax$ for each $x \in B$. A_n **converges weakly** to A if $\langle A_n x, y\rangle_{B\times B^*} \to \langle Ax, y\rangle_{B\times B^*}$ for all $x \in B$, $y \in B^*$. If B is a Hilbert space weak convergence is equivalently defined as $\langle A_n x, y\rangle_H \to \langle Ax, y\rangle_H$ for all $x, y \in H$.*

Definition 10.2. *If P and Q are two projections in H, then P is called **orthogonal** to Q if $\mathrm{Ran}(P) \perp \mathrm{Ran}(Q)$.*

Proposition 10.3. *A bounded operator P with range M is the orthogonal projection onto M iff $P^2 = P$ and $P^* = P$.*

Proof. We already know that the orthogonal projection onto a closed subspace M has these properties. Suppose then that $P^2 = P$ and $P^* = P$ and $M = \mathrm{Ran}\,(P)$. If $x \in M$ then $x = Py$ for some y. Hence: $Px = P^2 y = Py = x$. So $P|_M = I_M$. The subspace, M, is closed since, if $x_n \in M$ and $x_n \to x$ then $Px = \lim Px_n = \lim x_n = x$. Hence $x \in M$. It remains to show that $\mathrm{Nul}(P) = M^\perp$.

If $x \in M$ and $Py = 0$ then

$$\langle x, y\rangle = \langle Px, y\rangle = \langle x, Py\rangle = 0$$

and therefore $\mathrm{Nul}\,(P) \subset M^\perp$. If $y \in M^\perp$ then

$$\langle x, Py\rangle = \langle Px, y\rangle = 0 \;\forall\; x \in H$$

which implies $Py = 0$, i.e. $y \in \mathrm{Nul}\,(P)$. ■

Note: Henceforth **projection** means "**orthogonal projection**".

Corollary 10.4. *If P_1, P_2 are two projections with ranges M_1, M_2, respectively, then*

1. *$M_1 \perp M_2$ implies $P_1 P_2 = P_2 P_1 = 0$.*
2. *$P_1 P_2 = 0$ implies $M_1 \perp M_2$.*
3. *In case of 1. or 2., $P_1 + P_2$ is the projection onto span $\{M_1, M_2\}$.*

[If either of the equivalent conditions in items 1. or 2. hold we say P_1 is orthogonal to P_2 and write $P_1 \perp P_2$.]

Proof. We will take each item in turn.

1. If $M_1 \perp M_2$, then for any $x \in H$, $P_1 x \in M_1 \subset M_2^\perp = \mathrm{Nul}(P_2)$ and hence $P_2 P_1 x = 0$. Similarly one shows $P_1 P_2 = 0$.
2. If $P_1 P_2 = 0$ and $x \in M_1$, $y \in M_2$, then
$$\langle x, y\rangle = \langle P_1 x, P_2 y\rangle = \langle x, P_1 P_2 y\rangle = 0,$$
which shows $M_1 \perp M_2$.
3. If $P_1 P_2 = 0$, then
$$(P_1 + P_2)^2 = P_1^2 + P_1 P_2 + P_2 P_1 + P_2^2 = P_1 + P_2$$

and moreover $(P_1+P_2)^* = P_1+P_2$. Therefore by Proposition 10.3, $P = P_1+P_2$ is the projection onto some closed subspace M. If $x \in M_1$, $y \in M_2$ then

$$P(x+y) = P_1x + P_2x + P_1y + P_2y = P_1x + P_2y = x+y$$

and therefore $M \supseteq M_1+M_2$. If $z \in M$, then $z = Pz = P_1z+P_2z \in M_1+M_2$. ■

Proposition 10.5. *If P_n is a sequence of mutually orthogonal projections, then strong $\lim_{n\to\infty}\sum_{k=1}^n P_k$ exists and is the projection onto the closure of span $\{\operatorname{Ran}(P_n)\}_{n=1}^\infty$.*

Proof. Let $Q_n = \sum_{k=1}^n P_k$. Then Q_n is the projection on $M_1+\cdots+M_n$ where $M_j = \operatorname{Ran}(P_j)$ by Corollary 10.4 and induction. The Proposition is now a consequence of the Martingale Convergence Theorem of Exercise **??**. For those who did not do that exercise, I will complete the proof here.

As Q_n is orthogonal projection we know, for all $x \in H$, that $\|Q_nx\|^2 \leqq \|x\|^2$. This inequality is, by Pythagorean's theorem, equivalent to

$$\|x\|^2 \geqq \left\|\sum_{k=1}^n P_kx\right\|^2 = \left\langle \sum_{k=1}^n P_kx, \sum_{j=1}^n P_jx\right\rangle = \sum_{k=1}^n \|P_kx\|^2$$

from which it follows that $\sum_{k=1}^\infty \|P_kx\|^2 \le \|x\|^2$. But if $n > m$,

$$\|(Q_n - Q_m)x\|^2 = \sum_{k=m+1}^n \|P_kx\|^2 \to 0 \text{ as } m,n\to\infty.$$

Hence $Qx := \lim_{n\to\infty} Q_nx$ exists for all $x \in H$. The operator Q is clearly a bounded linear operator and $\|Q\| \leqq 1$. Since

$$\langle Qx,y\rangle = \lim_{n\to\infty}\langle Q_nx,y\rangle = \lim_{n\to\infty}\langle x,Q_ny\rangle = \langle x,Qy\rangle \ \forall\, x,y\in H,$$

$Q^* = Q$. Using $Q_mQ_n = Q_m$ if $n \geqq m$, it follows that for any $x \in H$,

$$Q^2x = \lim_m Q_mQx = \lim_m\lim_n Q_mQ_nx = \lim_m Q_mx = Qx,$$

i.e. $Q^2 = Q$. Thus Q is the projection on some closed subspace M.

If $x \in M_k$, then $Q_nx = x$ for $n \geqq k$ and therefore $Qx = x$. This shows $M_k \subset M$ and as k was arbitrary we may conclude $N := \overline{\operatorname{span}\{M_n\}} \subset M$. Finally, if $x \in N^\perp$, then $x \perp M_n$ for all $n \in \mathbb{N}$, i.e. $Q_nx = 0$ for all n. Therefore $Qx = 0$ and we have shown $N^\perp \subset M^\perp$ which implies $M \subset N$. ■

Definition 10.6 (Projection valued measures). *Let Ω be a set and let $\mathcal{S}$ be a sub σ–field of 2^Ω. A* ***projection valued measure*** *on $\mathcal{S}$ is a function $E(\cdot)$ from $\mathcal{S}$ to projections on a Hilbert space H such that*

1. *$E(\emptyset) = 0$,*
2. *$E(\Omega) = I$,*
3. *$E(A\cap B) = E(A)E(B)$ where $A,B \in \mathcal{S}$, and*
4. *if $A_1, A_2, \ldots$ is a disjoint sequence in $\mathcal{S}$, then*

$$E(\cup_{n=1}^\infty A_n) = \sum_{n=1}^\infty E(A_n) \text{ (strong sum).} \tag{10.1}$$

Remark 10.7. Items 1. and 3. of Definition 10.6 imply; if $A\cap B = \emptyset$ then $E(A)$ and $E(B)$ are mutually orthogonal. Hence the strong sum in Eq. (10.1) converges to a projection by Proposition 10.5.

Remark 10.8. If $E(\cdot)$ a projection valued measure on a measurable space, $(\Omega,\mathcal{S})$, then for every $x,y\in H$, $\mathcal{S} \ni B \to \langle E(B)x,y\rangle$ is a complex measure on $\mathcal{S}$.

Notation 10.9 *If $E(\cdot)$ a projection valued measure on a measurable space, $(\Omega,\mathcal{S})$, and $v \in H$ we let μ_v denote the positive measure on $\mathcal{S}$ defined by*

$$\mu_v(B) := \langle E(B)v,v\rangle \text{ for all } B \in \mathcal{S}.$$

Example 10.10. Suppose $(\Omega,\mathcal{S})$ is a measure space, $(Z,\mathcal{M},\mu)$ is a $\sigma-$ finite measure space, and $G : Z \to \Omega$ is a measurable function. Then one easily checks $E(A) := M_{1_A\circ G} = M_{1_{G^{-1}(A)}} \in \mathcal{B}\left(L^2(X,\mu)\right)$ for all $A \in \mathcal{S}$ defines a projection valued measure.

Example 10.11. Suppose that $\mathbf{T} = (T_1,T_2,\ldots,T_n)$ are $n-$ commuting normal operators, $\Omega = \sigma(\mathbf{T})$, $\mathcal{S} = \mathcal{B}(\sigma_{ap}(\mathbf{T}))$, and for $A \in \mathcal{S}$, let $E(A) := 1_A(\mathbf{T}) := \varphi_{\mathbf{T}}(1_A)$. Then $\{E(A) : A \in \mathcal{S}\}$ is a projection valued measure. We will see shortly that in this case, if $v \in H$ and μ_v is the measure in Notation 10.9, then μ_v agrees with the measure used in Eq. (**??**) of Theorem **??**.

Definition 10.12. *Let $E(\cdot)$ be a projection valued measure $(\Omega,\mathcal{S})$. As usual if $f : \Omega \to \mathbb{C}$ is an $\mathcal{S}$ -simple function, then we define*

$$\int_\Omega f dE := \sum_{\lambda\in\mathbb{C}} \lambda E(f=\lambda)$$

where

$$E(f=\lambda) := E(\{\omega\in\Omega : f(\omega)=\lambda\}) = E\left(f^{-1}(\{\lambda\})\right).$$

[To simplify notation we will often write $\int f dE$ for $\int_\Omega f dE$.]

As for any vector valued measure, the above integral is a $\mathcal{B}(H)$ – valued linear transformation on the $\mathcal{S}$ – simple functions, see the proof of Proposition 1.28 which goes through without any significant change. Let us summarize some more properties of this integral.

Proposition 10.13 (Properties of $f \to \int f dE$). *The map*

$$\{\mathcal{S}\text{-simple functions}\} \ni f \to \int_\Omega f dE \in \mathcal{B}(H)$$

is a linear transformation. Moreover this integral satisfies the following identities;

$$\int \bar{f} dE = (\int f dE)^*, \tag{10.2}$$

$$\int fg dE = \left(\int f dE\right)\left(\int g dE\right), \tag{10.3}$$

and for $x, y \in H$,

$$\left\langle \left(\int_\Omega f(\omega)\, dE(\omega)\right) x, y\right\rangle = \int_\Omega f(\omega)\, d\langle E(\omega) x, y\rangle \text{ and} \tag{10.4}$$

$$\left\langle \left(\int f dE\right) x, \left(\int g dE\right) y\right\rangle = \int f\bar{g} d\langle E(\cdot) x, y\rangle. \tag{10.5}$$

In particular,

$$\left\|\left(\int f dE\right) x\right\|^2 = \int_\Omega |f(\omega)|^2\, d\mu_x(\omega) \text{ and} \tag{10.6}$$

$$\left\|\int f dE\right\|_{op} \le \sup_{\omega\in\Omega} |f(\omega)|. \tag{10.7}$$

Proof. We take the identities in turn. Equation (10.2) is proved by;

$$\begin{aligned}\left(\int f dE\right)^* &= \left(\sum_{\lambda\in\mathbb{C}} \lambda E(f=\lambda)\right)^* = \sum_{\lambda\in\mathbb{C}} \bar{\lambda} E(f=\lambda) \\ &= \sum_{\lambda\in\mathbb{C}} \lambda E\left(f=\bar{\lambda}\right) = \sum_{\lambda\in\mathbb{C}} \lambda E\left(\bar{f}=\lambda\right) = \int \bar{f} dE.\end{aligned}$$

For Eq. (10.3) we first observe that

$$fg = \sum_{\alpha,\beta\in\mathbb{C}} \alpha\beta 1_{f=\alpha} 1_{g=\beta} = \sum_{\alpha,\beta\in\mathbb{C}} \alpha\beta 1_{\{f=\alpha, g=\beta\}}$$

and hence,

$$\begin{aligned}\int fg dE &= \sum_{\alpha,\beta\in\mathbb{C}} \alpha\beta E(f=\alpha, g=\beta) = \sum_{\alpha,\beta\in\mathbb{C}} \alpha\beta E(f=\alpha) E(g=\beta) \\ &= \sum_{\alpha\in\mathbb{C}} \alpha E(f=\alpha) \sum_{\beta\in\mathbb{C}} \beta E(g=\beta) = \left(\int f dE\right)\left(\int g dE\right).\end{aligned}$$

For Eq. (10.4) we have

$$\begin{aligned}\left\langle \left(\int_\Omega f(\omega)\, dE(\omega)\right) x, y\right\rangle &= \left\langle \left(\sum_{\lambda\in\mathbb{C}} \lambda E(f=\lambda)\right) x, y\right\rangle \\ &= \sum_{\lambda\in\mathbb{C}} \lambda \langle (E(f=\lambda)) x, y\rangle \\ &= \int_\Omega f d\langle E(\cdot) x, y\rangle.\end{aligned}$$

Equation (10.5) now follows easily from what we have already proved, namely;

$$\begin{aligned}\left\langle \left(\int f dE\right) x, \left(\int g dE\right) y\right\rangle &= \left\langle \left(\int g dE\right)^* \left(\int f dE\right) x, y\right\rangle \\ &= \left\langle \left(\int \bar{g} dE\right)\left(\int f dE\right) x, y\right\rangle \\ &= \left\langle \left(\int f\bar{g} dE\right) x, y\right\rangle \\ &= \int f\bar{g} d\langle E(\cdot) x, y\rangle.\end{aligned}$$

Taking $g = f$ and $x = y$ in the above identity implies Eq. (10.6) and Eq. (10.7) easily follows since $\mu_x(\Omega) = \|x\|^2$.

■

Definition 10.14. *If f is a bounded measurable function, let f_n be a sequence of simple measurable functions converging to f uniformly. Then by Eq. (10.6) of Proposition 10.13,*

$$\left\|\int f_n dE - \int f_m dE\right\|_{op} \to 0 \text{ as } m, n \to \infty.$$

Hence we may define $\int f dE$ by, $\int f dE = \lim_{n\to\infty} \int f_n dE$ (in operator norm topology).

The majority of the proof of the following corollary is straightforward and will be left to the reader.

Corollary 10.15. *All of the properties of the integral in Proposition 10.13 extend to the integral on bounded measurable functions in Definition 10.14. Moreover, if $f_n \to f$ boundedly then*

$$\int f_n dE \overset{s}{\to} \int f dE \text{ and } n \to \infty.$$

Proof. We only verify the last assertion here. The key point is that for any $x \in H$ we still have have

$$\left\| \left(\int f dE \right) x \right\|^2 = \int |f|^2 \, d\mu_x$$

as it holds for simple functions by Proposition 10.13 and then for bounded measurable functions by taking uniform limits. Thus if $f_n \to f$ boundedly we have

$$\left\| \left(\int f dE \right) x - \left(\int f_n dE \right) x \right\|^2 = \int |f - f_n|^2 \, d\mu_x \to 0 \text{ as } n \to \infty$$

by DCT. ■

Remark 10.16 (Truncation). If $B \in \mathcal{S}$ and f is a bounded measurable function on Ω, then

$$\left(\int f dE \right) E(B) = \left(\int f dE \right) \left(\int 1_B dE \right) = \int 1_B f dE.$$

As usual we let

$$\int_B f dE := \int_\Omega 1_B f dE = \left(\int_\Omega f dE \right) E(B).$$

Example 10.17 (Continuation of Example 10.10). Let us continue the setup in Example 10.10, i.e. $(\Omega, \mathcal{S})$ is a measure space, $(Z, \mathcal{M}, \mu)$ is a σ – finite measure space, $G : Z \to \Omega$ is a measurable function, and

$$E(A) := M_{1_A \circ G} = M_{1_{G^{-1}(A)}} \in \mathcal{B}\left(L^2(Z, \mu)\right) \text{ for all } A \in \mathcal{S}.$$

In this case if $f : \Omega \to \mathbb{C}$ is a bounded measurable function then

$$\int_\Omega f dE = M_{f \circ G}.$$

Indeed, if $f = 1_A$ for some $A \in \mathcal{S}$ then

$$\int_\Omega 1_A dE = E(A) = M_{1_A \circ G}$$

and hence if f is an $\mathcal{S}$ – simple function,

$$\begin{aligned} \int f dE &= \sum \lambda E(f = \lambda) = \sum \lambda M_{1_{\{f=\lambda\}} \circ G} \\ &= \sum \lambda M_{1_{\{f \circ G = \lambda\}}} = M_{f \circ G}. \end{aligned}$$

For general bounded $\mathcal{S}$ – measurable functions f, we may choose $\mathcal{S}$ – simple functions, f_n, so that $f_n \to f$ uniformly and therefore,

$$\int f dE = \lim_{n \to \infty} \int f_n dE = \lim_{n \to \infty} M_{f_n \circ G} = M_{f \circ G}$$

wherein we have used

$$\|M_{f_n \circ G} - M_{f \circ G}\|_{op} \le \sup_{z \in Z} |f_n(G(z)) - f(G(z))| \le \|f_n - f\|_\infty \to 0 \text{ as } n \to \infty.$$

10.2 Spectral Resolutions

Definition 10.18 (Support). *The support of a projection measure E on $(\mathbb{C}^n, \mathcal{B}(\mathbb{C}^n))$ is the set*

$$\operatorname{supp}(E) := \{\lambda \in \mathbb{C}^n : E(B(\lambda, \varepsilon)) \neq 0 \text{ for all } \varepsilon > 0\}.$$

Remark 10.19. Here are a few simple remarks about $\operatorname{supp}(E)$.

1. $\mathbb{C}^n \setminus \operatorname{supp}(E)$ is an open set and hence $\operatorname{supp}(E)$ is a closed set.
2. If $K = \operatorname{supp}(E)$ and $f : \mathbb{C}^n \to \mathbb{C}$ is a bounded measurable function, then
$$\int_{\mathbb{C}^n} f(z) \, dE(z) = \int_{\mathbb{C}^n} 1_K(z) f(z) \, dE(z).$$
This follows directly from Remark 10.16 and the observation that $E(K) = I$ so that and hence
$$\int_{\mathbb{C}^n} f dE = \left(\int_{\mathbb{C}^n} f dE \right) E(K) = \int_{\mathbb{C}^n} 1_K f dE.$$
3. If $A \subset \mathbb{C}^n$ is a Borel set and $v \in \operatorname{Ran} E(A)$, then $\operatorname{supp}(\mu_v) \subset \bar{A}$. Indeed we have
$$\begin{aligned} \mu_v(B) &= \langle E(B) v, v \rangle = \langle E(B) E(A) v, v \rangle \\ &= \langle E(B \cap A) v, v \rangle = \mu_v(B \cap A). \end{aligned}$$
Thus if $\lambda \in \mathbb{C}^n$ and $\varepsilon > 0$ so that $B(\lambda, \varepsilon) \cap A = \emptyset$ we must have
$$\mu_v(B(\lambda, \varepsilon)) = \mu_v(B(\lambda, \varepsilon) \cap A) = \mu_v(\emptyset) = 0.$$
This shows that if $\lambda \in \bar{A}^c$ then $\lambda \in \operatorname{supp}(\mu_v)^c$, i.e. $\operatorname{supp}(\mu_v) \subset \bar{A}$.

Lemma 10.20 (Optional). *Suppose that E is a compactly supported projection valued measure on $(\mathbb{C}^n, \mathcal{B}(\mathbb{C}^n))$ and for $1 \le j \le n$,*

$$T_j := \int_K z_j dE(\mathbf{z})$$

where $K = \operatorname{supp}(E)$. Then $\operatorname{supp}(E) = \sigma_{ap}(\mathbf{T})$ where $\mathbf{T} = (T_1, \dots, T_n)$.

Proof. First observe if $\lambda \in \mathbb{C}^n$ and $v \in H$, then

$$\begin{aligned}\sum_{j=1}^n \|(T_j - \lambda_j) v\|^2 &= \sum_{j=1}^n \int_K |z_j - \lambda_j|^2 d\mu_v(\mathbf{z}) \\ &= \int_K |\mathbf{z} - \lambda|^2 d\mu_v(\mathbf{z}). \qquad (10.8)\end{aligned}$$

If we now assume $\lambda \in \operatorname{supp}(E)$ and $\varepsilon > 0$, we can find $v_\varepsilon \in \operatorname{Ran} E(B(\lambda, \varepsilon))$ such that $\|v_\varepsilon\| = 1$. Taking $v = v_\varepsilon$ in Eq. (10.8) and making use of Item 3. in Remark 10.19 we find,

$$\begin{aligned}\sum_{j=1}^n \|(T_j - \lambda_j) v_\varepsilon\|^2 &= \int_{\overline{B(\lambda,\varepsilon)}} |\mathbf{z} - \lambda|^2 d\mu_{v_\varepsilon}(\mathbf{z}) \\ &\le \varepsilon^2 \|v_\varepsilon\|^2 = \varepsilon^2\end{aligned}$$

and as $\varepsilon > 0$ was arbitrary we have shown $\lambda \in \sigma_{ap}(\mathbf{T})$. Conversely if $\lambda \notin \operatorname{supp}(E)$ there exists $\varepsilon > 0$ so that $E(B(\lambda, \varepsilon)) = 0$ and hence for any $v \in H$ we also have $\mu_v(B(\lambda, \varepsilon)) = 0$. Using this remark back in Eq. (10.8) shows,

$$\sum_{j=1}^n \|(T_j - \lambda_j) v\|^2 = \int_{K \setminus B(\lambda,\varepsilon)} |\mathbf{z} - \lambda|^2 d\mu_v(\mathbf{z}) \ge \varepsilon^2 \|v\|^2$$

and hence $\lambda \notin \sigma_{ap}(\mathbf{T})$. ■

Definition 10.21 (Spectral resolution). *Suppose that $\mathbf{T} = (T_1, \dots, T_n)$ is a list of commuting normal operators in $B(H)$. A spectral valued measure, E, on $(\mathbb{C}^n, \mathcal{B}(\mathbb{C}^n))$ is a* ***spectral resolution*** *for $\mathbf{T}$ provided E is compactly supported and*

$$T_j = \int_{\sigma_{ap}(\mathbf{T})} z_j dE(z) \text{ for } 1 \le j \le n. \qquad (10.9)$$

Theorem 10.22 (Spectral Resolution Theorem). *If $\mathbf{T} = (T_1, \dots, T_n)$ is a list of commuting normal operators in $B(H)$, then $\mathbf{T}$ has a unique spectral resolution, namely $E(B) = 1_B(\mathbf{T})$ for all $B \in \mathcal{B}(\mathbb{C}^n)$. Moreover,* $\operatorname{supp}(E) = \sigma_{ap}(\mathbf{T})$.

Proof. Uniqueness. Let $E(\cdot)$ be a spectral resolution of $\mathbf{T}$ and $K := \operatorname{supp}(E) \cup \sigma_{ap}(\mathbf{T})$. For any bounded measurable function, f, on K let

$$\psi(f) := \int_E f dE.$$

It is now easily verified that ψ is a $*-$ homomorphism satisfying the hypothesis of satisfying the same properties as φ in Theorem 7.42 and hence we in fact must have $\psi(f) = \varphi\left(f 1_{\sigma_{ap}(\mathbf{T})}\right)$ by the same uniqueness proof used there.

Existence. Let $E(B) = 1_{B \cap \sigma_{ap}(\mathbf{T})}(\mathbf{T})$ for all $B \in \mathcal{B}(\mathbb{C}^n)$ in which case $\operatorname{supp}(E(\cdot)) \subset \sigma_{ap}(\mathbf{T})$. Moreover, by the spectral Corollary 7.40 we may find a finite measure space $(X, \mathcal{F}, \mu)$ and a measurable functions $\pi : X \to \sigma_{ap}(\mathbf{T}) \subset \mathbb{C}^n$ such that $T_j = U M_{\pi_j} U^*$ for some unitary map, $U : L^2(\mu) \to H$. With this notation we have $E(B) = U M_{1_A \circ \pi} U^*$. If $f : \sigma_{ap}(\mathbf{T}) \to \mathbb{C}$ is a bounded measurable function and $\tilde{E}(A) := M_{1_A \circ \pi}$, it is easy to verify that

$$\int_{\sigma_{ap}(\mathbf{T})} f dE = U \left[\int_{\sigma_{ap}(\mathbf{T})} f d\tilde{E}\right] U^*.$$

From Example 10.17, we know that

$$\int_{\sigma_{ap}(\mathbf{T})} f d\tilde{E} = M_{f \circ \pi}$$

and therefore,

$$\int_{\sigma_{ap}(\mathbf{T})} f dE = U M_{f \circ \pi} U^* = f(\mathbf{T}).$$

Taking $f(z) = z_j$ then shows that E is a spectral resolution of $\mathbf{T}$. ■

Corollary 10.23. *Let $\mathbf{T} = (T_1, \dots, T_n)$be a list of commuting normal operators in $B(H)$ and $E(\cdot)$ be the corresponding spectral resolution. If $D \in B(H)$ satisfies $[T_j, D] = 0$ for $1 \le j \le n$, then $[E(B), D] = 0$ for all $B \in \mathcal{B}(\sigma_{ap}(\mathbf{T}))$. In other words, $E(B) \in C(I, \mathbf{T})''$ – the double commutant of $C(I, \mathbf{T})$.*

Proof. This follows directly from item 7. of Theorem 7.42 and the fact that $E(B) = \varphi_{\mathbf{T}}(1_B)$ for all $B \in \mathcal{B}(\sigma_{ap}(\mathbf{T}))$. ■

Corollary 10.24. *If A is a bounded normal operator on H with spectral resolution $E(\cdot)$ then support $E = \sigma(A)$.*

Proof. From the construction of E, we see that the support of E is the essential range of f. But essential range of $f = \sigma(M_f) = \sigma(A)$ since unitary equivalences preserve spectrum. ■

Lemma 10.25. *For any bounded operator A,*

$$\sigma(A^*) = \operatorname{conj}(\sigma(A)) := \{\bar{z} : z \in \sigma(A)\}.$$

Furthermore, if A is invertible, then $\sigma(A^{-1}) = \sigma(A)^{-1}$.

The proof follows easily from the definitions of $\sigma(A)$, A^* and A^{-1}.

Proposition 10.26. *If $A = A^*$ then $\sigma(A) \subset \mathbb{R}$ and if U is unitary then*

$$\sigma(U) \subset S^1 := \{z \in \mathbb{C} : |z| = 1\}.$$

Proof. The first assertion easily follows from Lemma 10.25. For the second, since $\|U\| = 1$ we know $\sigma(U) \subseteq \{z : |z| \le 1\}$. If $0 < |z| < 1$ and $z \in \sigma(U)$ then

$$z^{-1} \in \sigma(U^{-1}) = \sigma(U^*) \subset \{z : |z| \le 1\},$$

a contradiction. Finally, it is clear that $0 \notin \sigma(U)$. Alternatively just notice that U is unitarily equivalent to M_f for some function f which is necessarily taking values if S^1. Indeed if f took values outside of S^1 with positive measure it would be easy to show M_f is not an isometry. ∎

Corollary 10.27 (Spectral theorem for a bounded Hermitian operator). *If A is a bounded Hermitian operator on a separable Hilbert space H, then there exists a unique projection–valued Borel measure $E(\cdot)$ on the line with compact support such that*

$$A = \int_{-\infty}^{\infty} \lambda dE(\lambda).$$

For all real Borel sets B, $E(B) \subset \{A\}''$.

Proof. $\sigma(A) \subset (-\infty, \infty)$ by the proposition. Apply Corollary 10.24. ∎

Corollary 10.28 (Spectral theorem for a unitary operator). *If U is a unitary operator on a separable Hilbert space, then there exists a unique projection–valued Borel measure $E(\cdot)$ on $[0, 2\pi)$ such that*

$$U = \int_0^{2\pi} e^{i\theta} dE(\theta),$$

and $E(B) \subset \{U\}''$ for all Borel sets B.

Proof. The same as for Corollary 10.27 if we map $[0, 2\pi)$ onto $\{z : |z| = 1\}$ with $\theta \to e^{i\theta}$. ∎

10.3 Spectral Types

Definition 10.29 (Atoms). *Suppose that E is a projection valued measure on a measurable space, $(\Omega, \mathcal{S})$. We say $A \in \mathcal{S}$ is an **atom of** E provided $E(A) \neq 0$ and either $E(A \cap B) = E(A)$ or 0 for every $B \in \mathcal{S}$.*

Exercise 10.1. Suppose that E is a projection valued measure on a measurable space, $(\mathbb{R}^n, \mathcal{B} = \mathcal{B}_{\mathbb{R}^n})$ for some $n \in \mathbb{N}$. If $B \in \mathcal{B}$ is an atom of E, then there exists a unique point $\lambda \in B$ such that $E(B) = E(\{\lambda\})$.

Definition 10.30 (Point spectrum). *If $\mathbf{T} = (T_1, \ldots, T_n)$ be a list of bounded operators and $\lambda \in \mathbb{C}^n$, let*

$$\operatorname{Nul}(\mathbf{T} - \lambda) := \cap_{j=1}^n \operatorname{Nul}(T_j - \lambda_j)$$

and let

$$\sigma_p(\mathbf{T}) := \{\lambda \in \mathbb{C}^n : \operatorname{Nul}(\mathbf{T} - \lambda) \neq \{0\}\}.$$

In other words, $\lambda \in \sigma_p(\mathbf{T})$ iff there exists $v \neq 0$ so that $T_j v = \lambda_j v$ for all $1 \le j \le n$.

Theorem 10.31 (Joint eigen-vectors). *Let $\mathbf{T} = (T_1, \ldots, T_n)$ be a list of commuting normal operators in $B(H)$ and $E(\cdot)$ be the corresponding spectral resolution. Then for all $\lambda \in \mathbb{C}^n$ we have*

$$\operatorname{Nul}(\mathbf{T} - \lambda) = \operatorname{Ran} E(\{\lambda\}) \tag{10.10}$$

and in particular $\lambda \in \sigma_p(\mathbf{T})$ iff $E(\{\lambda\}) \neq 0$.

Proof. For $\lambda \in \mathbb{C}^n$ and $v \in H$ we have

$$\begin{aligned}
\sum_{j=1}^n \|(T_j - \lambda_j) v\|^2 &= \sum_{j=1}^n \int_{\sigma_{ap}(\mathbf{T})} |z_j - \lambda_j|^2 \, d\mu_v(z) \\
&= \int_{\sigma_{ap}(\mathbf{T})} \sum_{j=1}^n |z_j - \lambda_j|^2 \, d\mu_v(z) \\
&= \int_{\sigma_{ap}(\mathbf{T})} |z - \lambda|^2 \, d\mu_v(z). \qquad (10.11)
\end{aligned}$$

Hence if $v \in \operatorname{Nul}(\mathbf{T} - \lambda)$, then

$$0 = \int_{\sigma_{ap}(\mathbf{T})} |z - \lambda|^2 \, d\mu_v(z)$$

from which it follows that $\mu_v(\{\lambda\}^c) = 0$. This then implies that

$$\|v\|^2 = \mu_v(\{\lambda\}) = \|E(\{\lambda\})v\|^2 \implies v \in \operatorname{Ran}(E(\{\lambda\})).$$

[This is true even if $v = 0$.] Thus we have shown

$$\operatorname{Nul}(\mathbf{T} - \lambda) \subset \operatorname{Ran} E(\{\lambda\}).$$

For any $A, B \in \mathcal{B}(\sigma_{ap}(\mathbf{T}))$ we have

$$\mu_{E(B)v}(A) = \langle E(A)E(B)v, E(B)v\rangle = \langle E(A \cap B)v, v\rangle = \mu_v(B \cap A).$$

Therefore if follows from Eq. (10.11) with v replaced by $E(B)v$ that

$$\sum_{j=1}^{n} \|(T_j - \lambda_j)E(B)v\|^2 = \int_B |z - \lambda|^2 d\mu_v(z).$$

In particular if $B = \{\lambda\}$, this shows

$$\sum_{j=1}^{n} \|(T_j - \lambda_j)E(\{\lambda\})v\|^2 = \int_{\{\lambda\}} |z - \lambda|^2 d\mu_v(z) = 0.$$

Hence if $v \in \operatorname{Ran} E(\{\lambda\})$, then $v \in \operatorname{Nul}(\mathbf{T}-\lambda)$ and the proof of Eq. (10.10) is complete. ■

Definition 10.32. *Let A be any bounded operator. The set $\sigma_p(A)$ of all eigenvalues of A is called the* ***point spectrum*** *of A. Let H_p be the closed subspace of H spanned by the eigenvectors of A. If $H_p = H$ then A is said to have* ***pure point spectrum****.*

Example 10.33. $H = \ell^2$. If $x = \{a_n\}_{n=1}^{\infty} \in \ell^2$, put

$$Ax = \left\{\frac{1}{n}a_n\right\}_{n=1}^{\infty}.$$

Then A is a bounded multiplication operator by a real function, and is hence Hermitian.

$$\sigma(A) = \{1, 1/2, 1/3, \ldots, 0\}.$$

Each point is an eigenvalue, except 0. The eigenvector corresponding to $1/n$ is

$$x_n = (0, 0, \ldots, 1, 0, \ldots).$$

In this example, $H_p = H$ but $\sigma_p(A) \neq \sigma(A)$ since $0 \notin \sigma_p(A)$.

Remark 10.34. Every element $x \in H_p$ is of the form, $x = \sum_{j=1}^{N} x_j$ where $\{x_j\}_{j=1}^{N}$ is an orthogonal set of eigenvectors of A $/\!\!($ $N = \infty$ allowed). Assuming $Ax_j = \lambda_j x_j$ it then follows that

$$E(B)x = \sum_{j=1}^{N} E(B)x_j = \sum_{j=1}^{N} 1_B(\lambda_j)x_j$$

and hence

$$\mu_x(B) = \langle E(B)x, x\rangle = \left\langle \sum_{j=1}^{N} 1_B(\lambda_j)x_j, x \right\rangle = \sum_{j=1}^{N} 1_B(\lambda_j)\|x_j\|^2$$

which is to say

$$\mu_x = \sum_{j=1}^{N} \|x_j\|^2 \delta_{\lambda_j}.$$

Definition 10.35. *If $H_p = \{0\}$ then A is said to have* ***purely continuous spectrum.***

Example 10.36. $H = L^2(0,1)$, $A = M_{x+2}$. Then A has no eigenvalues, as we have seen before. Hence $\sigma_p(A) = \emptyset$. Thus A has purely continuous spectrum. Note that $\sigma(A) = [2, 3]$.

Example 10.37. Let $Q =$ rationals in $[0,1]$ with the counting measure. Let $A = M_{x+2}$. Then

$$\sigma(A) = \operatorname{essran}(x + 2) = [2, 3].$$

But every rational number in $[2, 3]$ is an eigenvalue of A because the function

$$f(x) = \begin{cases} 1 \text{ if } & x = r \\ 0 \text{ if } x \neq r, & x \in [0,1] \end{cases}$$

is an eigenfunction associated to the eigenvalue $2 + r$ if r is a rational in $[0, 1]$. Since these functions form an Orthonormal basis of H we have $H_p = H$. Thus A has pure point spectrum in spite of the fact that $\sigma(A) = [2, 3]$, which is the same spectrum as in Example 10.36.

Exercise 10.2. Suppose that μ is a measure on $(\mathbb{R}, \mathcal{B})$ such that $B \in \mathcal{B}$ is a finite atom, see Definition 2.57. Show there exists a unique point $\lambda \in B$ such that $\mu(\{\lambda\}) = \mu(B)$. [This exercise easily extends to the case of measure on $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ as well.]

Exercise 10.3 (Decomposition by spectral type). Let A be a bounded Hermitian operator on a complex Hilbert space H. Suppose that $A = \int_{-\infty}^{\infty} \lambda dE(\lambda)$ is its spectral resolution. Denote by H_{ac} the set of all vectors x in H such that the measure $B \to \mu_x(B) := \|E(B)x\|^2$ is absolutely continuous with respect to Lebesgue measure.

1. Show that H_{ac} is a closed subspace of H.

2. Show that $H_p \perp H_{ac}$.
3. Define $H_{sc} = (H_p + H_{ac})^{\perp}$. (So we have the decomposition $H = H_p \oplus H_{ac} \oplus H_{sc}$.) Show that if $x \in H_{sc}$ and $x \neq 0$ then the measure
$$B \to \mu_x(B) := \langle E(B)x, x\rangle = \|E(B)x\|^2$$
has no atoms and yet there exists a Borel set B of Lebesgue measure zero such that $E(B)x \neq 0$.
4. Show that the decomposition of part c) reduces A. That is, $AH_i \subset H_i$, for $i = p$, ac, or sc.

10.4 More Exercise

Exercise 10.4 (Behavior of the resolvent near an isolated eigenvalue). We saw in the proof of Corollary 3.44 in Chapter 2 that if A is a bounded operator on a complex Banach space and λ_0 is not in $\sigma(A)$ then $(A-\lambda)^{-1}$ has a power series expansion: $(A-\lambda)^{-1} = \sum_{n=0}^{\infty}(\lambda-\lambda_0)^n B_n$ valid in some disk $|\lambda - \lambda_0| < \varepsilon$, where each B_n is a bounded operator.

1. Suppose that A is the operator on the two dimensional Hilbert space $\mathbb{C}^2$ given by the two by two matrix
$$A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}.$$
As you (had better) know, $\sigma(A) = \{3\}$. Show that the resolvent $(A-\lambda)^{-1}$ has a Laurent expansion near $\lambda = 3$ with a pole of order two. That is
$$(A-\lambda)^{-1} = (\lambda-3)^{-2}B_{-2} + (\lambda-3)^{-1}B_{-1} + \sum_{n=0}^{\infty}(\lambda-\lambda_0)^n B_n$$
which is valid in some punctured disk $0 < |\lambda - 3| < a$. Find B_{-2} and B_{-1} and show that neither operator is zero.
2. Suppose now that A is a bounded self-adjoint operator on a complex, separable, Hilbert space H. Suppose that λ_0 is an isolated eigenvalue of A, by which we mean that, for some $\varepsilon > 0$
$$\sigma(A) \cap \{\lambda \in \mathbb{C} : |\lambda - \lambda_0| < \varepsilon\} = \{\lambda_0\}.$$
Prove that $(A-\lambda)^{-1}$ has a pole of order one around λ_0, in the sense that, for some $\delta > 0$,
$$(A-\lambda)^{-1} = (\lambda-\lambda_0)^{-1}B_{-1} + \sum_{n=0}^{\infty}(\lambda-\lambda_0)^n B_n, \; 0 < |\lambda-\lambda_0| < \delta,$$
where the operators B_j, $j = -1, 0, 1, \ldots$ are bounded operators on H. Express B_{-1} in terms of the spectral resolution of A.

Definition 10.38. *A* ***one parameter unitary group*** *is a function $U : \mathbb{R} \to$ unitary operators on a Hilbert space H such that*
$$U(t+s) = U(t)\,U(s) \;\; \forall\, s, t \in \mathbb{R}.$$

Exercise 10.5. Let A be a bounded Hermitian operator on a separable Hilbert space H. Denote by $E(\cdot)$ its spectral resolution. Assume that $A \geq 0$ and write $P = E(\{0\})$ (which may or may not be the zero projection). Prove that for any vector u in H
$$\lim_{t\to+\infty} e^{-tA}u = Pu.$$

Exercise 10.6. Let V be a unitary operator on a separable complex Hilbert space H. Prove that there exists a one parameter group $U(t)$ on H such that

(a) $U(1) = V$,
(b) $U(\cdot)$ is continuous in the operator norm.

Part IV

Unbounded operators

Unbounded Operator Introduction

Definition 11.1. *If X and Y are Banach spaces and $\mathcal{D}$ is a subspace of X, then a linear transformation T from $\mathcal{D}$ into Y is called a linear transformation (or operator) from X to Y with domain $\mathcal{D}(T) = \mathcal{D}$. If $\mathcal{D}(T)$ is dense in X, we say T is said to be* ***densely defined.***

Notation 11.2 *If S, T are two operators from X to Y we say T is an* ***extension*** *of S if $\mathcal{D}(S) \subset \mathcal{D}(T)$ and $S = T$ on $\mathcal{D}(S)$. We abbreviate this by writing $S \subset T$, see Remark 11.8.*

Definition 11.3. *If $S, T : X \to Y$ are linear operators we define $S + T : X \to Y$ by setting $\mathcal{D}(S+T) := \mathcal{D}(S) \cap \mathcal{D}(T)$ and for $x \in \mathcal{D}(S+T)$ we let $(S+T)x = Sx + Tx$. If $T : X \to Y$ and $S : Y \to Z$ are linear operators we define $ST : X \to Z$ by setting*

$$\mathcal{D}(ST) := \{x \in \mathcal{D}(T) : Tx \in \mathcal{D}(S)\}$$

and for $x \in \mathcal{D}(ST)$ we let $(ST)x = S(Tx)$.

Proposition 11.4 (Properties of sums and products). *Let A, B and C be operators from H to H, then*

1. *$A(BC) = (AB)C$*
2. *$(A+B)C = AC + BC$*
3. *$AB + AC \subset A(B+C)$ with equality if A is everywhere defined.*

Proof. The only real issue to check in each of this assertions is that the domains of the operators on both sides of the equations are the same because it is easily checked that equality holds on the intersection of the domains of the operators on each side of the equation.

1. We have
$$\begin{aligned}\mathcal{D}(A(BC)) &= \{h \in \mathcal{D}(BC) : BCh \in \mathcal{D}(A)\} \\ &= \{h \in \mathcal{D}(C) : Ch \in \mathcal{D}(B) \text{ and } BCh \in \mathcal{D}(A)\}\end{aligned}$$
while
$$\begin{aligned}\mathcal{D}((AB)C) &= \{h \in \mathcal{D}(C) : Ch \in \mathcal{D}(AB)\} \\ &= \{h \in \mathcal{D}(C) : Ch \in \mathcal{D}(B) \text{ and } BCh \in \mathcal{D}(A)\}.\end{aligned}$$
2. For the second item;
$$\begin{aligned}\mathcal{D}((A+B)C) &= \{h \in \mathcal{D}(C) : Ch \in \mathcal{D}(A) \cap \mathcal{D}(B)\} \\ &= C^{-1}(\mathcal{D}(A) \cap \mathcal{D}(B))\end{aligned}$$
while
$$\begin{aligned}\mathcal{D}(AC + BC) &= \mathcal{D}(AC) \cap \mathcal{D}(BC) \\ &= \{h \in \mathcal{D}(C) : Ch \in \mathcal{D}(A) \cap \mathcal{D}(B)\} \\ &= C^{-1}(\mathcal{D}(A) \cap \mathcal{D}(B)).\end{aligned}$$
3. Lastly, we have $h \in \mathcal{D}(AB + AC) = \mathcal{D}(AB) \cap \mathcal{D}(AC)$ iff $h \in \mathcal{D}(B) \cap \mathcal{D}(C)$ and $Bh, Ch \in \mathcal{D}(A)$ which implies $h \in \mathcal{D}(B) \cap \mathcal{D}(C)$ and $(B+C)h \in \mathcal{D}(A)$, i.e. $h \in \mathcal{D}(A(B+C))$. If we further assume that A is everywhere defined then
$$\begin{aligned}&\mathcal{D}(AB + AC) = \mathcal{D}(AB) \cap \mathcal{D}(AC) = \mathcal{D}(B) \cap \mathcal{D}(C) \text{ and} \\ &\mathcal{D}(A(B+C)) = \mathcal{D}(B+C) = \mathcal{D}(B) \cap \mathcal{D}(C).\end{aligned}$$

■

Remark 11.5. The inclusion in item 3. may be strict. For example, suppose $A = B = -C = \frac{d}{dx}$ with common domains being $C_c^1(\mathbb{R}) \subset L^2(\mathbb{R}) = H$. Then

$$\begin{aligned}\mathcal{D}(AB + AC) &= \left\{h \in C_c^1(\mathbb{R}) : h' \in C_c^1(\mathbb{R})\right\} \\ &= C_c^2(\mathbb{R}) \subsetneqq C_c^1(\mathbb{R}) = \mathcal{D}(A(B+C)),\end{aligned}$$

wherein we have used $B + C = \mathbf{0}|_{C_c^1(\mathbb{R})}$.

Exercise 11.1. Suppose that $A, B : H \to K$ are (unbounded) operators such that; 1) $A \subset B$, 2) A is surjective, and 3) B is injective. Show $A = B$. [**Hint:** this result would hold for arbitray functions A, B between two abstarct sets H and K. This has nothing to do with linearity! In general if $A \subset B$ and B is injective, then A is injective and $A^{-1} \subset B^{-1}$.]

We note that $X \times Y$ is a Banach space in either of the equivalent norms;

$$\| (x,y) \| = \|x\| + \|y\| \text{ or} \tag{11.1}$$

$$\| (x,y) \| = \sqrt{\|x\|^2 + \|y\|^2}. \tag{11.2}$$

When H and K are Hilbert spaces, then $H \times K$ and $K \times H$ become Hilbert spaces by defining

$$\langle (x,y), (x',y') \rangle_{H\times K} := \langle x, x' \rangle_H + \langle y, y' \rangle_K \text{ and} \tag{11.3}$$

$$\langle (y,x), (y',x') \rangle_{K\times H} := \langle x, x' \rangle_H + \langle y, y' \rangle_K \tag{11.4}$$

respectively. The Hilbert norm associated to the inner-product in Eq. (11.1) is the norm in Eq. (11.2).

Definition 11.6 (Graph of an operator). *If T is an operator from X to Y with domain $\mathcal{D}$, the* graph*of T is*

$$\Gamma(T) := \{(x, Tx) : x \in \mathcal{D}(T)\} \subset H \times K.$$

Note that $\Gamma(T)$ is a subspace of $X \times Y$.

The linearity of T assures that $Z := \Gamma(T)$ is a subspace of $X \times Y$. Moreover it is easy to check that $\pi_X(Z) = \mathcal{D}(T)$ and $\#\{y \in Y : (x,y) \in Z\} = 1$. The next lemma shows that operators $T : X \to Y$ are in one to one correspondence with subspaces $Z \subset X \times Y$ such that Z passes the *vertical line test*, i.e.

$$\#\{y \in Y : (x,y) \in Z\} = 1 \text{ for all } x \in \pi_X(Z). \tag{11.5}$$

Lemma 11.7 (Vertical line test at $x = 0$ suffices). *If Z is a subspace of $X \times Y$ such that $(0,y) \in Z$ happens iff $y = 0$ (i.e. $[\{0\} \times Y] \cap Z = \{(0,0)\}$), then Z is the graph of an operator $T : X \to Y$. Explicitly, $\mathcal{D}(T) := \pi_X(Z)$ and for $x \in \mathcal{D}(T)$ we let $Tx = y$ where $y \in Y$ is uniquely determined by requiring $(x,y) \in Z$. Alternatively stated, $T = \pi_Y \circ \left(\pi_X|_Z^{-1}\right)$.*

Proof. Suppose that Z is a subspace of $X \times Y$ satisfying the assumption of the lemma. If $z = (x,y) \in Z$ is such that $\pi_X(z) = x = 0$,then (by assumption) $y = 0$ and hence $z = 0$. This shows that $\pi_X|_Z : Z \to \pi_X(Z) = \mathcal{D}(T)$ is a linear isomorphism and hence we may define $T = \pi_Y \circ \left(\pi_X|_Z^{-1}\right) : \mathcal{D}(T) \to Y$. Clearly T is linear and moreover for $x \in \mathcal{D}(T)$ and $Tx = y$ we have $(x,y) \in Z$. Similarly, if $(x,y) \in Z$, then $\pi_X(x,y) = x \in \mathcal{D}(T)$ and therefore $\pi_X|_Z^{-1}(x) = (x,y)$ and so $Tx = y$. Thus we have shown $\Gamma(T) = Z$. ∎

Remark 11.8. The reader should verify for herself that if S and T are two operators from X to Y, then $S \subset T$ iff $\Gamma(S) \subset \Gamma(T)$.

Definition 11.9 (Closed operators). *We say an operator $T : X \to Y$ is **closed** if $\Gamma(T)$ is a closed subspace of $X \times Y$. Recall the closed graph theorem.*

Remark 11.10. An operator $T : X \rightsquigarrow Y$ is closed iff for all sequences $\{x_n\} \subset \mathcal{D}(T)$ such that $x_n \to x$ in X and $Tx_n \to y$ in Y implies $x \in \mathcal{D}(T)$ and $y = Tx$. In other words, $\lim_{n\to\infty} Tx_n = T \lim_{n\to\infty} x_n$ provided $\lim_{n\to\infty} x_n$ and $\lim_{n\to\infty} Tx_n$ exist. So T is **closed** if $\lim_{n\to\infty} Tx_n = T \lim_{n\to\infty} x_n$ for all $\{x_n\}_{n=1}^\infty \subset \mathcal{D}(T)$ where both $\lim_{n\to\infty} x_n$ and $\lim_{n\to\infty} Tx_n$ exists. [If $T : X \to Y$ everywhere defined, then T is continuous iff $\lim_{n\to\infty} Tx_n = T \lim_{n\to\infty} x_n$ whenever $\lim_{n\to\infty} x_n$ exists in X.]

Example 11.11. If $T : X \to Y$ is an everywhere defined bounded operator then T is closed. Indeed, if $(x_n, y_n) = (x_n, Tx_n) \in \Gamma(T)$ and $(x_n, y_n) \to (x,y)$, then

$$y = \lim_{n\to\infty} y_n = \lim_{n\to\infty} Tx_n = T \lim_{n\to\infty} x_n = Tx$$

so that $(x,y) \in \Gamma(T)$. It turns out the converse is true as well provided X and Y are Banach spaces.

Theorem 11.12 (Closed Graph Theorem). *If X and Y are Banach spaces and $T : X \to Y$ is closed and everywhere defined and linear operator, then T is bounded. **Moral:** Unbounded closed operators from one Banach space to another cannot be everywhere defined.*

Exercise 11.2. Suppose that (X,μ) is a measure space and that $\mu(X) < \infty$. Let $T : L^2(\mu) \to L^2(\mu)$ be a bounded operator. Suppose that range T is contained in $L^5(\mu)$. Show that T is bounded as an operator from $L^2(\mu)$ into $L^5(\mu)$. **Hint:** Use the closed graph theorem. (See [41, Chapter 5, poblem 16.]. The solution to this problem depends on Theorem 5.10 of the same reference.)

Exercise 11.3. Suppose that $X = Y = BC(\mathbb{R}, \mathbb{C})$, the bounded continuous functions on $\mathbb{R}$ equipped with the supremum norm. Let $\mathcal{D}(T) = BC^1(\mathbb{R}, \mathbb{C})$ be those $f \in X$ which are differentiable with $f' \in X$ and for $f \in \mathcal{D}(T)$, let $Tf = f'$. Show T is a closed operator. [Hint: this is a standard undergraduate theorem in diguise.]

Lemma 11.13. *If X and Y are Banach spaces, $\mathcal{D}(A)$ is a dense subspace of X, and $A : \mathcal{D}(A) \to Y$ is a closed operator which is a bijection, then $A^{-1} : Y \to X$ is bounded.*

Proof. By the closed graph theorem we need only show A^{-1} has a closed graph. To this end suppose that $\{y_n\} \subset Y$ is a sequence such that both

$$y = \lim_{n\to\infty} y_n \text{ and } x = \lim_{n\to\infty} A^{-1} y_n$$

exist. Letting $x_n := A^{-1} y_n$ we have $x_n \to x$ and $Ax_n = y_n \to y$. As A is a closed operator, it follows that $x \in \mathcal{D}(A)$ and $Ax = y$ and so $x = A^{-1}y$ and hence A^{-1} is a closed operator. ∎

Theorem 11.14. *Let X and Y be Banach spaces and suppose that $T : X \to Y$ is an unbounded densely defined operator. For $x \in \mathcal{D}(T)$ let*

$$\|x\|_T := \|x\|_X + \|Tx\|_Y .$$

Then T is closed iff $(\mathcal{D}(T), \|\cdot\|_T)$ is a Banach space.

Proof. Short proof. The linear transformation, $\mathcal{D}(T) \ni x \to (x, Tx) \in \Gamma(T)$ is as surjective isometry of normed spaces. Thus $(\mathcal{D}(T), \|\cdot\|_T)$ will be a Banach space iff $\Gamma(T)$ is a Banach space which happens iff $\Gamma(T)$ is closed in $X \times Y$. ∎

Definition 11.15 (Closable operators). *An operator $T : X \to Y$ is **closable** if $\overline{\Gamma(T)}$ is the graph of an operator,see Lemma 11.7. If T is closable, we let $\bar{T} : X \to Y$ be the unique operator such that $\overline{\Gamma(T)} = \Gamma(\bar{T})$.*

Lemma 11.16. *An operator $T : X \to Y$ is closable iff for every $\{x_n\} \subset \mathcal{D}(T)$ such that $x_n \to 0$ and $y := \lim_{n\to\infty} Tx_n$ exists Y we must have $y = 0$.*

Proof. Let us first observe that $(0, y) \in \overline{\Gamma(T)}$ iff there exists $\{x_n\} \subset \mathcal{D}(T)$ such that $(x_n, Tx_n) \to (0, y)$, i.e. iff $x_n \in \mathcal{D}(T)$, $x_n \to 0$ in X, and $Tx_n \to y$. Thus according to Lemma 11.7, T is closable iff $(0, y) \in \overline{\Gamma(T)}$ implies $y = 0$ iff for every $\{x_n\} \subset \mathcal{D}(T)$ such that $x_n \to 0$ and $y := \lim_{n\to\infty} Tx_n$ exists Y we must have $y = 0$. ∎

Example 11.17. Let $1 \le p < \infty$ and $T : L^p(\mathbb{R}) \to \mathbb{C}$ defined by $\mathcal{D}(T) = C_c(\mathbb{R})$ and $Tf = f(0)$ is not closable. In fact, $\overline{\Gamma(T)} = L^p(\mathbb{R}) \times \mathbb{C}$. To see this is the case let $\varphi_k(x)$ be the tent function which is 0 on $\mathbb{R} \setminus \left[-\frac{1}{k}, \frac{1}{k}\right]$, $\varphi_k(0) = 1$, and φ_k is linear on $\left[-\frac{1}{k}, 0\right]$ and $\left[0, \frac{1}{k}\right]$. Given $f \in L^p(\mathbb{R})$, choose $g_n \in C_c(\mathbb{R})$ so that $g_n \to f$ in $L^p(\mathbb{R})$. Given an $a \in \mathbb{C}$, let

$$f_n(x) = g_n(x) + (a - g_n(0))\,\varphi_{k_n}$$

where $k_n \uparrow \infty$ sufficiently rapidly so that

$$|(a - g_n(0))|\,\|\varphi_{k_n}\|_p \to 0 \text{ as } n \to \infty.$$

We then have $f_n \to f$ in $L^p(\mathbb{R})$ while $Tf_n = f_n(0) = a \to a$ which shows $(f, a) \in \overline{\Gamma(T)}$.

Proposition 11.18 (IBP $\Longrightarrow$ Closable). *Suppose that $1 < p < \infty$ and $X = Y = L^p(\mathbb{R}^n, m)$. Let $\mathcal{D}(T) = C_c^\infty(\mathbb{R}^n)$ and for $f \in \mathcal{D}(T)$, let*

$$Tf = \sum_{|\alpha| \le m} a_\alpha D^\alpha f \text{ where } a_\alpha \in C^\infty(\mathbb{R}^n) \text{ and } D^\alpha = \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n}.$$

Here $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ and $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$. Then $T : X \to X$ is closable.

Proof. Let $q = \frac{p}{p-1}$ be the conjugate exponent to p. For $f \in L^p$ and $g \in L^q$, let $\langle f, \varphi \rangle = \int_{\mathbb{R}^n} f(x)\,\varphi(x)\,dx$. Then by integration by parts for $f, \varphi \in \mathcal{D}(T)$,

$$\langle Tf, \varphi \rangle = \left\langle f, T^\dagger \varphi \right\rangle \text{ where } T^\dagger = \sum_{|\alpha| \le m} (-1)^{|\alpha|} D^\alpha M_{a_\alpha}.$$

Thus if $f_n \in \mathcal{D}(T)$ is such that $f_n \to 0$ and $Tf_n \to g$ in L^p, then for all $\varphi \in \mathcal{D}(T)$,

$$\langle g, \varphi \rangle = \lim_{n\to\infty} \langle Tf_n, \varphi \rangle = \lim_{n\to\infty} \left\langle f_n, T^\dagger \varphi \right\rangle = \langle 0, \varphi \rangle = 0.$$

As $\mathcal{D}(T) = C_c^\infty(\mathbb{R}^n)$ is dense in $L^q(m)$, it follows that $g = 0$ a.e and hence T is closable. ∎

Remark 11.19. If $S : X \to Y$ is a linear operator, then S is closable iff S has at least one closed extension T. Moreover, if S is closable and T is a close extension of S, then $\bar{S} \subset T$. Thus $\bar{S}$ is the smallest closed extension of S. Indeed, if S is closable, then $S \subset \bar{S}$ and hence S has a closed extension. Conversely if T is a closed extension of S then $\Gamma(S) \subset \Gamma(T)$ with $\Gamma(T)$ being closed and hence $\overline{\Gamma(S)} \subset \Gamma(T)$ which implies $\overline{\Gamma(S)}$ is necessarily the graph of a linear operator, i.e. S is closable. Moreover we see that $\Gamma(\bar{S}) = \overline{\Gamma(S)} \subset \Gamma(T)$ and therefore $\bar{S} \subset T$.

Remark 11.20. If $S, T : X \to Y$ are linear operators such that $S \subset T$ and T is closable, then S is closable and $\bar{S} \subset \bar{T}$. Indeed, $S \subset \bar{T}$ and so the result follows directly from Remark 11.19.

Alternatively: if $x_n \in \mathcal{D}(S) \subset \mathcal{D}(T)$ such that $x_n \to 0$ and $Sx_n \to y$, then $Tx_n = Sx_n \to y$ and since T is closable we must have $y = 0$ showing S is closable. Moreover, if $x \in \mathcal{D}(\bar{S})$ and $\bar{S}x = y$, there exists $x_n \in \mathcal{D}(S) \subset \mathcal{D}(T)$ such that $x_n \to x$ and $Sx_n \to y$. As $Tx_n = Sx_n \to y$ it follows that $x \in \mathcal{D}(\bar{T})$ and $\bar{T}x = y$ as well showing $\bar{S} \subset \bar{T}$.

Definition 11.21 (Cores). *A **core** for a closed operator T is a subspace $\mathcal{D}_0 \subset \mathcal{D}(T)$ such that $T = \overline{T|_{\mathcal{D}_0}}$, i.e. T is the closure of its restriction to $\mathcal{D}_0$.*

Example 11.22. If $A : X \to Y$ is a bounded everywhere defined operator and $\mathcal{D}_0$ is a dense subspace of X, then $A_0 := A|_{\mathcal{D}_0}$ is not closed but is closable and $\bar{A}_0 = A$. So any dense subspace of X is a core for all bounded operators on X.

Notation 11.23 (Multiplication Operators) *Given a measure space $(X, \mathcal{M}, \mu)$ and a measurable function $q : X \to \mathbb{C}$, let $M_q : L^2(\mu) \to L^2(\mu)$ denote the operation of multiplication by q. More precisely, $M_q : \mathcal{D}(M_q) \to L^2(\mu)$ is defined by $M_q f = qf$ where*

$$\mathcal{D}(M_q) := \{f \in L^2(\mu) : qf \in L^2(\mu)\} \subset L^2(\mu).$$

Lemma 11.24. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $q : \Omega \to \mathbb{C}$ be a measurable function (not assumed to be bounded!), and M_q be the corresponding multiplication operator on $H = L^2(\mu)$. Then M_q is a closed operator and*

$$\mathcal{D}_0 := \cup_{n=1}^{\infty} \left\{ f \in L^2(\mu) : 1_{|q| \le n} f = f \right\}$$

is a core for M_q.

Proof. **M_q is closed.** Suppose that $\{f_n\} \subset \mathcal{D}(M_q)$ is s such that $f_n \to f$ and $M_q f_n = q f_n \to g$ in $L^2(\mu)$. By passing to a subsequence if necessary we may assume $f_n \to f$ and $q f_n \to g$ a.e. and from this we learn that $qf = g$ a.e. This shows $f \in \mathcal{D}(M_q)$ and that $M_q f = g$, i.e. M_q is closed.

$\mathcal{D}_0$ is a core. For $f \in \mathcal{D}(M_q)$, let $f_n := 1_{|q| \le n} f \in \mathcal{D}_0$. Then by DCT, $f_n \to f$ and $q f_n \to q f$ in $L^2(\mu)$ which shows $\overline{M_q|_{\mathcal{D}_0}} = M_q$, i.e. $\mathcal{D}_0$ is a core for M_q. ■

Definition 11.25. *Let $(\Omega, \mathcal{S})$ be a measurable space, H be a Hilbert space, $E(\cdot)$ be a projection valued measure $(\Omega, \mathcal{S})$, $\mu_x(B) := \langle E(B)x, x \rangle$ for all $x \in H$, and $f : \Omega \to \mathbb{C}$ be a measurable function. Let*

$$\mathcal{D}(T_f) := \left\{ x \in H : \int_\Omega |f|^2 \, d\mu_x < \infty \right\}$$

and for $x \in \mathcal{D}(T_f)$ let

$$T_f x := \lim_{n \to \infty} \left(\int_\Omega f 1_{\{|f| \le n\}} dE \right) x. \tag{11.6}$$

Notice that for $m \le n$,

$$\begin{aligned} &\left\| \left(\int_\Omega f 1_{\{|f| \le n\}} dE \right) x - \left(\int_\Omega f 1_{\{|f| \le m\}} dE \right) x \right\|^2 \\ &\quad = \left\| \left(\int_\Omega f 1_{\{m < |f| \le n\}} dE \right) x \right\|^2 = \int_\Omega 1_{\{m < |f| \le n\}} |f|^2 \, d\mu_x \to 0 \text{ as } m, n \to \infty \end{aligned}$$

and so the limit in Eq. (11.6) exists.

Theorem 11.26 (Projection valued measures II). *Continuing the notation in Definition 11.25, $\mathcal{D}(T_f)$ is a dense subspace of H and $T_f : H \to H$ is a closed linear operator on H which satisfies, for all $x \in \mathcal{D}(T_f)$,*

$$\|T_f x\|^2 = \int_\Omega |f|^2 \, d\mu_x \text{ and} \tag{11.7}$$

$$E(\{|f| \le n\}) T_f x = T_f E(\{|f| \le n\}) x = \left(\int_\Omega f 1_{\{|f| \le n\}} dE \right) x. \tag{11.8}$$

Moreover $L := \cup_{n=1}^{\infty} \operatorname{Ran}(E(\{|f| \le n\}))$ is a core for T_f.

Proof. If $x, y \in H$, then

$$\sqrt{\mu_{x+y}(B)} = \|E(B)(x+y)\| \le \|E(B)x\| + \|E(B)y\| = \sqrt{\mu_x(B)} + \sqrt{\mu_y(B)}$$

and therefore,

$$\mu_{x+y} \le \mu_x + \mu_y + 2\sqrt{\mu_x}\sqrt{\mu_y} \le 2(\mu_x + \mu_y).$$

From this one easily shows if $x, y \in \mathcal{D}(T_f)$ then $x + y \in \mathcal{D}(T_f)$ and since $\mu_{cx} = |c|^2 \mu_x$ we also have $cx \in \mathcal{D}(T_f)$ whenever $x \in \mathcal{D}(T_f)$. This show $\mathcal{D}(T_f)$ is a vector space. It is now a simple matter to use the definition of T_f in Eq. (11.6) to verify that T_f is linear and the equality in Eq. (11.7) holds. To see that $\mathcal{D}(T_f)$ is dense in H let $P_n := E(\{|f| \le n\})$ and observe that $P_n x \in \mathcal{D}(T_f)$ for all $x \in H$. Indeed,

$$\mu_{P_n x}(B) = \mu_x(B \cap \{|f| \le n\})$$

and hence

$$\int_\Omega |f|^2 \, d\mu_{P_n x} = \int_\Omega 1_{|f| \le n} |f|^2 \, d\mu_x \le n^2 \|x\|^2 < \infty.$$

So it only remains to show T_f is closed.

Let $x_n \in \mathcal{D}(T_f)$ and $x, y \in H$ be chosen so that $x_n \to x$ and $T_f x_n \to y$. Notice that, for any $B \in \mathcal{S}$,

$$\mu_{x_n}(B) = \langle E(B)x_n, x_n \rangle \to \langle E(B)x, x \rangle = \mu_x(B).$$

Hence if g is any simple function on Ω such that $0 \le g \le |f|^2$, then

$$\begin{aligned} \int_\Omega g d\mu_x &= \lim_{n \to \infty} \int_\Omega g d\mu_{x_n} \\ &\le \liminf_{n \to \infty} \int_\Omega |f|^2 \, d\mu_{x_n} = \lim_{n \to \infty} \|T_f x_n\|^2 = \|y\|^2 < \infty. \end{aligned}$$

From this inequality if follows that $\int_\Omega |f|^2 \, d\mu_x \le \|y\|^2 < \infty$ so that $x \in \mathcal{D}(T_f)$. So it only remains to show $T_f x = y$.

However, for any $m \in \mathbb{N}$ and $z \in \mathcal{D}(T_f)$,

$$\begin{aligned} P_m T_f z &= \lim_{n \to \infty} \left(P_m \int_\Omega f 1_{\{|f| \le n\}} dE \right) z = \lim_{n \to \infty} \left(\int_\Omega f 1_{\{|f| \le m\}} 1_{\{|f| \le n\}} dE \right) z \\ &= \left(\int_\Omega f 1_{\{|f| \le m\}} dE \right) z. \end{aligned}$$

Therefore it follows that

$$\begin{aligned}P_m T_f x &= \left(\int_\Omega f 1_{\{|f|\le m\}} dE\right) x = \lim_{n\to\infty} \left(\int_\Omega f 1_{\{|f|\le m\}} dE\right) x_n \\ &= \lim_{n\to\infty} P_m T_f x_n = P_m y.\end{aligned}$$

Letting $m \to \infty$ in this last equation shows $T_f x = y$. Since $P_m T_f x = T_f P_m x$, this also shows L is a core for T_f. ∎

Definition 11.27. *If $T : X \to Y$ is a linear operator we let*

$$\begin{aligned}\mathrm{Nul}(T) &:= \{x \in \mathcal{D}(T) : Tx = 0\} \subset X \text{ and} \\ \mathrm{Ran}(T) &:= \{Tx \in Y : x \in \mathcal{D}(T)\} \subset Y.\end{aligned}$$

Lemma 11.28. *If $T : X \to Y$ is a closed operator, then $\mathrm{Nul}(T)$ is closed.*

Proof. If $x_n \in \mathrm{Nul}(T)$ such that $x_n \to x$, then $Tx_n = 0 \to 0$ as $n \to \infty$ which shows $x \in \mathcal{D}(T)$ and $Tx = 0$, i.e. $x \in \mathrm{Nul}(T)$. ∎

Lemma 11.29. *If $T : X \to Y$ is a closable operator and there exists $\varepsilon > 0$ such that*

$$\|Tx\| \ge \varepsilon \|x\| \ \ \forall\, x \in \mathcal{D}(T), \tag{11.9}$$

then $\mathrm{Nul}(\bar{T}) = \{0\}$ and $\mathrm{Ran}(\bar{T}) = \overline{\mathrm{Ran}(T)}$ and Eq. (11.9) extends to the inequality,

$$\|\bar{T}x\| \ge \varepsilon \|x\| \ \ \forall\, x \in \mathcal{D}(\bar{T}). \tag{11.10}$$

Proof. If $x \in \mathcal{D}(\bar{T})$, there exists $x_n \in \mathcal{D}(T)$ such that $x_n \to x$ and $Tx_n \to \bar{T}x$ and hence

$$\|\bar{T}x\| = \lim_{n\to\infty} \|T_n x_n\| \ge \varepsilon \lim_{n\to\infty} \|x_n\| = \varepsilon \|x\|$$

which shows Eq. (11.10) also holds. In is now trivial to verify $\mathrm{Nul}(\bar{T}) = \{0\}$.

If $y_n = Tx_n \in \mathrm{Ran}(T)$ is convergent to y, then by Eq. (11.9) it follows that $\{x_n\}_{n=1}^\infty$ is a Cauchy sequence. Hence it follows that $x := \lim_{n\to\infty} x_n \in \mathcal{D}(\bar{T})$ and $y = \bar{T}x$, i.e. $y \in \mathrm{Ran}(\bar{T})$. Thus we have shown $\overline{\mathrm{Ran}(T)} \subset \mathrm{Ran}(\bar{T})$. Conversely if $y = \bar{T}x \in \mathrm{Ran}(\bar{T})$, there exist $x_n \in \mathcal{D}(T)$ such that $x_n \to x$ and $Tx_n \to y$ which shows $y \in \overline{\mathrm{Ran}(T)}$. ∎

Notation 11.30 *If $T : X \to Y$ is a linear operator we say T is **weakly invertible** if $\mathrm{Nul}(T) := \{0\}$. Under this assumption we define $T^{-1} : Y \to X$ to be the linear operator with $\mathcal{D}(T^{-1}) = \mathrm{Ran}(T)$ such that $T^{-1}y = x \in \mathcal{D}(T)$ iff $Tx = y$. We say T is **invertible** if $\mathrm{Nul}(T) = \{0\}$ and $\mathcal{D}(T^{-1}) = \mathrm{Ran}(T) = Y$.*

Proposition 11.31. *If $T : X \to Y$ is a closed weakly invertible linear operator, then T^{-1} is closed. Moreover if T is closed and invertible, then T^{-1} is bounded. [Note well that a weakly invertible bounded everywhere defined bounded operator, $T : X \to Y$, need not have bounded inverse but it is always a closed operator.]*

Proof. If $y_n \in \mathcal{D}(T^{-1}) = \mathrm{Ran}(T)$ is such that $y_n \to y$ in Y and $T^{-1}y_n \to x$ in X. Let $x_n := T^{-1}y_n \to x$ and $Tx_n = y_n \to y$ as $n \to \infty$. Since T is closed it follows that $x \in \mathcal{D}(T)$ and $Tx = y$, i.e. $y \in \mathcal{D}(T^{-1})$ and $T^{-1}y = x$. This shows T^{-1} is closed. If T is closed and invertible, then T^{-1} is an everywhere defined closed operator and hence bounded by the closed graph theorem. ∎

Example 11.32. Suppose that $X = Y = L^1([0,1], dm)$,

$$\mathcal{D}(A) := \left\{f \in AC([0,1]) : f' \in L^1(m) \text{ and } f(0) = 0\right\}$$

and for $f \in \mathcal{D}(A)$ let $Af = f'$. Then by the fundamental theorem of calculus, $f(x) = \int_0^x f'(y)\,dy$ from which we learn that A is invertible with

$$\left(A^{-1}g\right)(x) = \int_0^x g(y)\,dy.$$

Let us note that

$$\left\|A^{-1}g\right\|_1 \le \int_{0\le y\le x\le 1} |g(y)|\,dy = \int_0^1 (1-y)\,|g(y)|\,dy \le \|g\|_1$$

so that A^{-1} is bounded and hence A is a closed operator by Proposition 11.31.

Lemma 11.33. *Suppose that $B : Y \to X$ is a bounded everywhere defined weakly invertible operator, $A := B^{-1}$, Y_0 is a dense subspace of Y, and $A_0 := B^{-1}|_{\mathcal{D}(A_0)}$ where $\mathcal{D}(A_0) := BY_0$. Then $\bar{A}_0 = A$.*

Proof. Suppose that $x \in \mathcal{D}(A)$ and $Ax = y$, i.e. $By = x$. Then choose $y_n \in Y_0$ so that $y_n \to y$ as $n \to \infty$ and let $x_n := By_n \in \mathcal{D}(A_0)$. We then have $x_n = By_n \to By =: x$ and $A_0 x_n = y_n \to y$ and therefore $x \in \mathcal{D}(\bar{A}_0)$ and $\bar{A}_0 x = y$. This shows that $\bar{A}_0 = A$. ∎

Lemma 11.34. *For $g \in L^1([0,1], dm)$, let*

$$\bar{g} := \int_0^1 g(x)\,dx \text{ and } \tilde{L}^1 := \left\{g \in L^1([0,1], dm) : \bar{g} := 0\right\}.$$

Then $\tilde{C}_c := \{g \in C_c(0,1) : \bar{g} = 0\}$ is a dense subspace of $\tilde{L}^1$.

Proof. Let $\psi \in C_c((0,1))$ be chosen so that $\int_0^1 \psi(x)\,dx = 1$. Then given $g \in \tilde{L}^1$ we may find $\gamma_n \in C_c(0,1)$ so that $\gamma_n \to g$ in L^1. Since $\bar{\gamma}_n \to \bar{g} = 0$ it follows that $g_n := \gamma_n - \bar{\gamma}_n \psi \in \tilde{C}_c$ also converges to g in $L^1(m)$. ∎

Example 11.35 (Continuation of Example 11.32). If $X = Y = L^1([0,1], dm)$, and $A : X \to X$ be the operator in Example 11.32. Further let $\mathcal{D}(A_0) :=$

$C_c^1((0,1))$, and for $f \in \mathcal{D}(A_0)$ let $A_0 f = f'$ so that $A_0 \subset A$. Then $\bar{A}_0 \subsetneq A$. More precisely,

$$\mathcal{D}(\bar{A}_0) = \{f :\in \mathcal{D}(A) : f(0) = 0 = f(1)\}.$$

Indeed, applying Lemma 11.33 with $X = L^1([0,1], dm)$, $Y = \tilde{L}^1$, $Y_0 = \tilde{C}_c$ which is dense in Y by Lemma 11.34, and $B = A^{-1}|_Y$ we have $B\tilde{C}_c = C_c^1(0,1) = \mathcal{D}(A_0)$ and

$$\mathcal{D}(\bar{A}_0) = B\tilde{L}^1 = \{f :\in \mathcal{D}(A) : f(0) = 0 = f(1)\}.$$

In particular this shows $C_c^1((0,1))$ is **not** a core for A.

Definition 11.36. *If $T : X \to X$ is a densely defined linear operator and $D \in B(X)$, we say T and D* ***commute*** *if $DT \subset TD$. [The condition $DT \subset TD$ is equivalent to; $D(\mathcal{D}(T)) \subset \mathcal{D}(T)$ and $DT = TD$ on $\mathcal{D}(T)$.]*

Exercise 11.4. Let $T : X \to X$ be a densely defined closable linear operator and $D \in B(X)$. If $D(\mathcal{D}(T)) \subset \mathcal{D}(T)$ and $DT = TD$ on $\mathcal{D}(T)$, then $D(\mathcal{D}(\bar{T})) \subset \mathcal{D}(\bar{T})$ and D commutes with $\bar{T}$, i.e. $D\bar{T} = \bar{T}D$ on $\mathcal{D}(\bar{T})$.

Proposition 11.37. *Suppose that T is a densely defined invertible (i.e. $\mathrm{Nul}(T) = \{0\}$ and $\mathrm{Ran}(T) = X$) and $D \in B(X)$, then $DT \subset TD$ iff $[T^{-1}, D] = 0$.*

Proof. If $DT \subset TD$, then $D = DTT^{-1} \subset TDT^{-1}$ and as $\mathcal{D}(D) = X$ we find $D = TDT^{-1}$. Therefore $T^{-1}D = T^{-1}TDT^{-1} = DT^{-1}$.

Conversely if $T^{-1}D = DT^{-1}$, then $D = TT^{-1}D = TDT^{-1}$ and hence

$$DT = TDT^{-1}T = TD1_{\mathcal{D}(T)} \subset TD.$$

■

Definition 11.38 (Spectrum and Resolvents). *For any (possibly unbounded) linear operator $T : X \to X$, a complex number $\lambda \in \mathbb{C}$ is said to be in the* ***resolvent set*** *$(\rho(T))$ of T if $T - \lambda I$ is one to one and onto and $(T - \lambda I)^{-1}$ is bounded. Otherwise λ is said to be in the* ***spectrum*** *$(\sigma(T))$ of T. For $\lambda \in \rho(T)$ we let $R_\lambda := (T - \lambda)^{-1}$.*

Proposition 11.39. *If $T : X \to X$ is an unbounded operator and $\rho(T) \neq \emptyset$, then T is closed. [So if T is not closed, then $\sigma(T) = \mathbb{C}$!]*

Proof. Let $\lambda \in \rho(T)$, so that $R_\lambda = (\lambda - T)^{-1} \in B(X)$. Suppose that $v_n \in \mathcal{D}(T)$ is such that $v_n \to v$ and $Tv_n \to w \in X$. It then follows that $(\lambda - T)v_n \to \lambda v - w$ and therefore,

$$v = \lim_{n\to\infty} v_n = \lim_{n\to\infty} (\lambda - T)^{-1}(\lambda - T)v_n = (\lambda - T)^{-1}(\lambda v - w) \in \mathcal{D}(T).$$

Applying $\lambda - T$ to this idenity shows,

$$(\lambda - T)v = (\lambda - T)(\lambda - T)^{-1}(\lambda v - w) = \lambda v - w,$$

i.e. $Tv = w$ and we have shown T is closed. ■

Lemma 11.40 (Spectrum of M_q). *Let $(X, \mathcal{M}, \mu)$ be a σ – finite measure space and $q : X \to \mathbb{C}$ be a measurable function, then*

1. *M_q is always a closed operator, i.e. $\Gamma(M_q) := \{(f, qf) : f \in \mathcal{D}(M_q)\}$ is a closed subspace of $L^2(\mu) \times L^2(\mu)$.*
2. *M_q is bounded iff $q \in L^\infty$ in which case $\|M_q\| = \|q\|_{L^\infty(\mu)}$.*
3. *The following are equivalent:*
   a) *$M_q^{-1} : L^2(\mu) \to L^2(\mu)$ exists in the algebraic sense, i.e. $M_q : \mathcal{D}(M_q) \subset L^2(\mu) \to L^2(\mu)$ is a bijection.*
   b) *$M_q^{-1} : L^2(\mu) \to L^2(\mu)$ exists as a bounded linear operator.*
   c) *$q \neq 0$ a.e. and $q^{-1} \in L^\infty(\mu)$.*
4. *$\sigma(M_q) = \mathrm{essran}_\mu(q)$. [Recall that $\lambda \in \mathrm{essran}_\mu(q)$ iff $\mu(|q - \lambda| < \varepsilon) > 0$ for all $\varepsilon > 0$, see Definition 1.32.]*

Proof. We take each item in turn.

1. This was proved in Lemma 11.24.

2. Let $K := \|q\|_{L^\infty(\mu)}$ which we assume to be positive for otherwise $M_q \equiv 0$. If $K < \infty$, then $|q| \le K$ a.e. from which it easily follows that $\|M_q\|_{op} \le K$. We wish to prove the reverse inequality for $0 < K \le \infty$. By assumption of $0 < k < K$, then $\mu(|q| \ge k) > 0$. Because μ is σ – finite we can find a set $A_k \in \mathcal{M}$ such that $|q| \ge k$ on A_k and $0 < \mu(A_k) < \infty$. We then take $f := \frac{\bar{q}}{|q|} 1_{A_k} \in L^2(\mu)$ and for this f we have

$$M_q f = |q| 1_{A_k} \ge k 1_{A_k}$$

and therefore,

$$\frac{\|M_q f\|_2^2}{\|f\|_2^2} \ge \frac{k^2 \mu(A_k)}{\mu(A_k)} = k^2 \implies \|M_q\|_{op} \ge k.$$

Since $0 < k < K$ was arbitrary we may conclude that $\|M_q\|_{op} \ge K$.

3. a. $\Longrightarrow$ b. is a special case of Lemma 11.13.
   b. $\Longrightarrow$ c. Observe if $\mu(q=0)>0$ there exists $A\subset\{q=0\}$ such that $0<\mu(A)<\infty$. Then $1_A\in \mathrm{Nul}(M_q)$ and M_q is not invertible. So if M_q^{-1} exists we must have $q\neq 0$ a.e. and if $f\in L^2$, $g:=M_q^{-1}f\in L^2$, then $f=M_q g=qg$ a.e. Therefore $g=q^{-1}f$ a.e. so that M_q^{-1} is necessarily given by $M_{q^{-1}}$. By item (1) this operator is bounded iff $q^{-1}\in L^\infty(\mu)$.
   c. $\Longrightarrow$ a. Indeed, $M_{q^{-1}}M_q=I_{D(M_q)}$ and $M_qM_{q^{-1}}=I_{L^2(\mu)}$.
4. By item (1), $\lambda\in\sigma(M_q)$ iff $\left\|(q-\lambda)^{-1}\right\|_{L^\infty}=\infty$ iff $\lambda\in\mathrm{essran}_\mu(q)$. To prove the last assertion, suppose first that $\lambda\in\mathrm{essran}_\mu(q)$. Then for all $\varepsilon>0$, $\mu(|q-\lambda|<\varepsilon)>0$ and hence $\left\|(q-\lambda)^{-1}\right\|_{L^\infty}\geq\varepsilon^{-1}$ which implies $\left\|(q-\lambda)^{-1}\right\|_{L^\infty}=\infty$ since $\varepsilon>0$ was arbitrary. Conversely if $\left\|(q-\lambda)^{-1}\right\|_{L^\infty}=\infty$ then for all $0<M<\infty$ we have $\mu\left(\left|(q-\lambda)^{-1}\right|>M\right)>0$, i.e. $\mu\left(|(q-\lambda)|<M^{-1}\right)>0$ for all $M>0$ which implies $\lambda\in\mathrm{essran}_\mu(q)$.

■

Proposition 11.41. *The set $\rho(T)$ is open and if $\rho(T)\neq\emptyset$, then $\rho(T)\ni\lambda\to R_\lambda\in B(X)$ is a continuous and in fact analytic map.*

Proof. We may assume $\rho(T)\neq\emptyset$ for otherwise there is nothing to prove. If $\lambda\in\rho(T)$ and $h\in\mathbb{C}$ we have,

$$(T-(\lambda+h))=\left[I-h(T-\lambda)^{-1}\right](T-\lambda)=(I-hR_\lambda)(T-\lambda).$$

Thus if $|h|\,\|R_\lambda\|<1$, we have $(I-hR_\lambda)^{-1}$ is invertible and we may conclude that $(T-(\lambda+h)):\mathcal{D}(T)\to H$ is one to one and onto, i.e. is invertible and moreover,

$$R_{\lambda+h}=(T-(\lambda+h))^{-1}=(T-\lambda)^{-1}(I-hR_\lambda)^{-1}.$$

This show $\rho(T)$ is open as well as showing $\lambda\to R_\lambda$ is in fact analytic in λ since,

$$(I-hR_\lambda)^{-1}=\sum_{n=0}^\infty h^nR_\lambda^n \text{ for } |h|<\|R_\lambda\|^{-1}.$$

■

If $\mu,\lambda\in\rho(T)$ with $\mu\neq\lambda$ then working informally we find,

$$\begin{aligned}R_\lambda-R_\mu&=\frac{1}{T-\lambda}-\frac{1}{T-\mu}=\frac{T-\mu-(T-\lambda)}{(T-\lambda)(T-\mu)}\\&=\frac{\lambda-\mu}{(T-\lambda)(T-\mu)}=(\lambda-\mu)R_\lambda R_\mu.\end{aligned}$$

This heuristic computation serves as motivation for the following important identity.

Theorem 11.42 (Resolvent Identity). *Let $T:X\to X$ be a densely defined closed operator and suppose that $\mu,\lambda\in\rho(T)$ with $\mu\neq\lambda$. Then*

$$R_\lambda-R_\mu=(\lambda-\mu)R_\lambda R_\mu \tag{11.11}$$

and in particular, $[R_\lambda,R_\mu]=0$. [Just interchange the roles of μ and λ in Eq. (11.11).]

Proof. We can basically give the same proof we gave in the bounded case, namely,

$$\begin{aligned}(T-\lambda)^{-1}-(T-\mu)^{-1}&=(T-\lambda)^{-1}[(T-\mu)-(T-\lambda)](T-\mu)^{-1}\\&=(\lambda-\mu)(T-\lambda)^{-1}(T-\mu)^{-1}.\end{aligned}$$

wherein the second line we have used $\mathrm{Ran}\left((T-\mu)^{-1}\right)=\mathcal{D}(T)$ and so $[(T-\mu)-(T-\lambda)](T-\mu)^{-1}$ is everywhere defined.

■

Lemma 11.43. *If $T:X\to X$ is a densely defined closed operator and $\lambda\in\rho(T)$, then $R_\lambda T\subset TR_\lambda$ and $\overline{R_\lambda T}=TR_\lambda$.*

Proof. As $R_\lambda=(T+\lambda)^{-1}$, we have

$$R_\lambda(T+\lambda)=I_{\mathcal{D}(T)}\subset I=(T+\lambda)R_\lambda$$

which easily implies $R_\lambda T\subset TR_\lambda$. Since

$$TR_\lambda=(T+\lambda-\lambda)R_\lambda=I-\lambda R_\lambda, \tag{11.12}$$

TR_λ is a bounded operator. Hence if $x\in X$ and $\{x_n\}$ is any sequence in $\mathcal{D}(T)$ such that $x_n\to x$, then

$$R_\lambda Tx_n=TR_\lambda x_n\to TR_\lambda x \text{ as } n\to\infty.$$

This shows $x\in\mathcal{D}\left(\overline{R_\lambda T}\right)$ and $\overline{R_\lambda T}x=TR_\lambda x$. ■

Remark 11.44. Here is another proof that $[R_\lambda,R_\mu]=0$ based on Proposition 11.37. Since

$$\begin{aligned}(T-\lambda)^{-1}(T-\mu)&=(T-\lambda)^{-1}(T-\lambda+(\lambda-\mu))\\&=I+(\lambda-\mu)(T-\lambda)^{-1} \text{ on } \mathcal{D}(T)\end{aligned}$$

and

$$(T-\mu)(T-\lambda)^{-1} = (T-\lambda+(\lambda-\mu))(T-\lambda)^{-1} = I+(\lambda-\mu)(T-\lambda)^{-1},$$

it follows that

$$(T-\lambda)^{-1}(T-\mu) \subset (T-\mu)(T-\lambda)^{-1}. \tag{11.13}$$

An application of Proposition 11.37 now completes the proof.

Corollary 11.45. *Let $T : X \to X$ be a densely defined closed operator. If $\lambda,\mu \in \rho(T)$, then*

$$TR_\lambda R_\mu = R_\lambda T R_\mu = TR_\mu R_\lambda = R_\mu T R_\lambda \tag{11.14}$$

Proof. From Lemma 11.43 or from Eq. (11.13) with $\mu = 0$ we know that $R_\lambda T = TR_\lambda$ on $\mathcal{D}(T)$ for all $\lambda \in \rho(T)$. From this assertion and the fact that R_μ and R_λ commute, Eq. (11.14) holds on $\mathcal{D}(T)$. Since TR_λ is a bounded operator for all $\lambda \in \rho(T)$ (see Eq. (11.12)), Eq. (11.14) holds on all of X by continuity. ■

Exercise 11.5. Let $T : X \to X$ be a densely defined closed operator. If $\lambda,\mu \in \rho(T)$, show $[TR_\lambda, TR_\mu] = 0$.

11.1 Exercises

Exercise 11.6. Let U be an open subset of $\mathbb{R}^n$, $g \in C(U,(0,\infty))$, and Y be a Banach space. Given $v \in \mathbb{R}^n$ define ∂_v to be the unbounded derivative operator on $B_gC(U,Y)$ given as follows. First of we let $\partial_v f(x) := \frac{d}{dt}|_0 f(x+tv)$ provided the limit exists in Y. We then set $\mathcal{D}(\partial_v)$ to be those functions $f \in B_gC(U,Y)$ such that $\partial_v f(x)$ exists for all $x \in U$ and for which that the resulting function $\partial_v f$ is back in $B_gC(U,Y)$. Show ∂_v is a closed unbounded operator on $B_gC(U,Y)$.

Exercise 11.7. If $A_i : X \to Y_i$ are closed operators for $1 \le i \le p$, the operator $\mathbf{A} : X \to Y := \prod_{i=1}^p Y_i$ defined by $\mathcal{D}(\mathbf{A}) = \cap_{i=1}^p \mathcal{D}(A_i)$ and

$$\mathbf{A}x = (A_1x,\dots,A_px) \in Y$$

is again a closed operator.

Corollary 11.46. *Let U be an open subset of $\mathbb{R}^n$, $g \in C(U,(0,\infty))$, and Y be a Banach space and let $\partial_i := \partial_{e_i}$ be the operators on $B_gC(U,Y)$ as described in Exercise 11.6. Then the operator $\nabla : X \to Y^n$ defined by $\nabla f := (\partial_1 f,\dots,\partial_n f)$ with $\mathcal{D}(\nabla) = \cap_{i=1}^n \mathcal{D}(\partial_i)$ is a closed operator. Moreover, $\mathcal{D}(\nabla) = B_gC^1(U,Y)$ where $f \in B_gC^1(U,Y)$ iff f is continuously differentiable and $Df \in B_gC(U,B(\mathbb{R}^n,Y))$.*

Corollary 11.47. *Let U be an open subset of $\mathbb{R}^n$, $g \in C(U,(0,\infty))$, Y be a Banach space, $k \in \mathbb{N}_0$, and $B_gC^k(U,Y)$ denote those $f \in B_gC(U,Y)$ such that $\partial_{v_1}\dots\partial_{v_l} f$ exists for all $1 \le \ell \le k$ and $v_i \in \mathbb{R}^n$. For $f \in B_gC^k(U,Y)$ let $J_k f := \{\partial^\alpha f : |\alpha| \le k\} \in Y^N$ be the k – jet of f, where $N := \#\{\alpha \in \mathbb{N}_0^n : |\alpha| \le k\}$. Then J_k is a closed operator.*

Proof. I will only prove the case $k = 2$ here leaving the general induction argument to the reader. Suppose that $f_n \in B_gC^2(U,Y)$ such that $f_n \to f$ and $\partial^\alpha f_n \to u_\alpha$ for $1 \le |\alpha| \le 2$ in $B_gC(U,Y)$. From Corollary 11.46 it follows that $f \in B_gC^1(U,Y)$ and that $u_\alpha = \partial^\alpha f$ for $|\alpha| = 1$. Moreover, given $1 \le i,j \le n$ then we know $\partial^{e_j} f_n = \partial_j f_n \to u_{e_j}$ and $\partial^{e_i+e_j} f_n = \partial_i\partial_j f_n \to u_{e_i+e_j}$ and hence by Exercise 11.6 it follows hat $u_{e_j} = \partial^{e_j} f \in \mathcal{D}(\partial_i)$ for $1 \le i \le n$ and $\partial^{e_i+e_j} f = \partial^{e_i} u_{e_j} = u_{e_i+e_j}$. As this holds for all i and j it now follows that $f \in B_gC^2(U,Y)$ and $u_\alpha = \partial^\alpha f$ for $|\alpha| \le 2$. ■

Exercise 11.8. Let U be an open interval in $\mathbb{R}$, $BC(U)$ be the bounded continuous functions on U and $BC^1(U)$ denote those $f \in C^1(U) \cap BC(U)$ such that $f' \in BC(U)$. Define $D(\partial) := BC^1(U) \subset C^1(U)$ and for $f \in D(\partial)$ let $\partial f := f'$. Show $\partial : BC(U) \to BC(U)$ is a closed unbounded operator.

Exercise 11.9. Generalize Exercise 11.8 to the following set up. Let U be an open subset of $\mathbb{R}^n$, let $BC^k(U)$ denote those $f \in BC(U)$ such that $\partial^\alpha f$ exits for all $|\alpha| \le k$ and $\partial^\alpha f \in BC(U)$. Let $J : BC(U) \to \prod_{|\alpha|\le k} BC(U)$ be the linear operator such that $D(J) = BC^k(U)$ and for $f \in D(J)$ let

$$Jf := \{\partial^\alpha f : |\alpha| \le k\}.$$

Show J is a closed operator. We refer of Jf as the k-jet of f. Perhaps we should let g be any positive continuos function on U and generalize the above considerations to allow for $BC(U)$ to be replace by $B_gC(U)$ where $f \in B_gC(U)$ iff $f \in C(U)$ and

12

Contraction Semigroups

For this section, let $(X, \|\cdot\|)$ be a Banach space with norm $\|\cdot\|$. Also let $T := \{T(t)\}_{t>0}$ be a collection of bounded operators on X.

Definition 12.1. *Let X and T be as above.*

1. *T is a **semi-group** if $T(t+s) = T(t)\,T(s)$ for all $s, t > 0$.*
2. *A semi-group T is **strongly continuous** if $\lim_{t\downarrow 0} T(t)\,v = v$ for all $v \in X$. By convention if T is strongly continuous, set $T(0) := I$–the identity operator on X.*
3. *A semi-group T is a **contraction semi-group** if $\|T(t)\| \le 1$ for $t > 0$. The following examples will be covered in more detail in the exercises.*

Example 12.2 (Translation Semi-group). Let $X := L^2(\mathbb{R}^d, d\lambda)$, $w \in \mathbb{R}^d$ and

$$(T_w(t)\,f)(x) := f(x+t).$$

Then $T_w(t)$ is a strongly continuous contraction semi-group. In fact $T_w(t)$ is unitary for all $t \in \mathbb{R}$.

Example 12.3 (Rotation Semi-group). Suppose that $X := L^2(\mathbb{R}^d, d\lambda)$ and $O : \mathbb{R} \to O(d)$ is a one parameter semi-group of orthogonal operators. Set $(T_O(t)\,f)(x) := f(O(t)\,x)$ for all $f \in X$ and $x \in \mathbb{R}^d$. Then T_O is also a strongly continuous unitary semi-group.

12.1 Infinitesimal Generators

Definition 12.4. *The **infinitesimal generator** (L) of a strongly continuous contraction semi-group, $\{T(t)\}_{t\ge 0}$ is the (unbounded) operator on X defined by*

$$Lv := \frac{d}{dt}|_{0+} T(t)\,v \text{ for all } v \in \mathcal{D}(L) \tag{12.1}$$

where $\mathcal{D}(L)$ consists precisely of those $v \in X$ for which the derivative in Eq. (12.1) exists in X.

Proposition 12.5. *Let $T(t)$ be a strongly continuous contraction semi-group, then;*

1. *$[0,\infty) \ni t \to T(t)\,v \in X$ is continuous for all $v \in X$,*
2. *$\mathcal{D}(L)$ is a dense linear subspace of X.*
3. *If $v : [0,\infty) \to X$ is a continuous, then $w(t) := T(t)\,v(t)$ is also continuous on $[0,\infty)$.*

Proof. We take each item in turn.

1. By assumption $v(t) := T(t)\,v$ is continuous at $t = 0$. For $t > 0$ and $h > 0$,

$$\|v(t+h) - v(t)\| = \|T(t)\,(T(h) - I)v\| \le \|v(h) - v\| \to 0 \text{ as } h \downarrow 0.$$

Similarly if $h \in (0,t)$,

$$\|v(t-h) - v(t)\| = \|T(t-h)\,(I - T(h))v\| \le \|v - v(h)\| \to 0 \text{ as } h \downarrow 0.$$

Thus we have proved item 1.

2. Since $\sigma \to T(\sigma)\,v$ is continuous for any $v \in X$, we may define

$$v_s := \int_0^s T(\sigma)\,v d\sigma \text{ for all } 0 < s < \infty,$$

where the integral may be interpreted as X–valued Riemann integral or as a Bochner integral. Note

$$\left\|\frac{1}{s}v_s - v\right\| = \left\|\frac{1}{s}\int_0^s (T(\sigma)\,v - v)d\sigma\right\| \le \frac{1}{|s|}\left|\int_0^s \|T(\sigma)\,v - v\|d\sigma\right| \to 0$$

as $s \downarrow 0$, so that

$$\mathcal{D} := \{v_s \in X : s > 0 \text{ and } v \in X\}$$

is dense in X. Moreover for any $s > 0$ and $v \in X$ we have

$$\begin{aligned}\left.\frac{d}{dt}\right|_{0+} T(t)\,v_s &= \left.\frac{d}{dt}\right|_{0+} \int_0^s T(t+\sigma)v d\sigma \\ &= \left.\frac{d}{dt}\right|_{0+} \int_t^{t+s} T(\tau)\,v d\tau = T(s)\,v - v\end{aligned}$$

which shows $v_s \in \mathcal{D}(L)$ and $Lv_s = T(s)\,v - v$. In particular, $\mathcal{D} \subset \mathcal{D}(L)$ and hence $\mathcal{D}(L)$ is dense in X. It is easily checked that $\mathcal{D}(L)$ is a linear subspace of X.

3. This statement is really just a consequence of the fact that the bilinear map,
$$X \times B(X) \ni (v,T) \to Tv \in X$$
is continuous when X is equipped with the norm topology and $B(X)$ with the strong operator topology and $X \times B(X)$ is then given the corresponding product topology. For completeness we will give a direct proof in the setting at hand.
If $v:[0,\infty) \to X$ is a continuous function and $w(t) := T(t)v(t)$, then for $t>0$ and $h \in (-t,\infty)$,
$$w(t+h) - w(t) = (T(t+h) - T(t))v(t) + T(t+h)(v(t+h) - v(t))$$
The first term goes to zero as $h \to 0$ by item 1 and the second term goes to zero since v is continuous and $\|T(t+h)\| \le 1$. The above argument also works with $t=0$ and $h \ge 0$.

■

Theorem 12.6 (Solution Operator). *Let $T(t)$ be a contraction semi-group with infinitesimal generator L and $v \in \mathcal{D}(L)$. Then $T(t)v \in \mathcal{D}(L)$ for all $t>0$, the function $t \to T(t)v$ is differentiable for $t>0$, and*
$$\frac{d}{dt}T(t)v = LT(t)v = T(t)Lv.$$

Proof. Let $v \in \mathcal{D}(L)$ and $t>0$. Then for $h>0$ we have
$$\frac{T(t+h)-T(t)}{h}v = \frac{T(h)-I}{h}T(t)v = \frac{T(t)(T(h)-I)}{h}v.$$
Letting $h \downarrow 0$ in the last set of equalities show that $T(t)v \in \mathcal{D}(L)$ and
$$\frac{d}{dh}|_{0+}T(t+h)v = LT(t)v = T(t)Lv. \tag{12.2}$$
Similarly for $h \in (0,t)$,
$$\frac{T(t-h)-T(t)}{-h}v = \frac{T(t)-T(t-h)}{h}v = T(t-h)\frac{T(h)-I}{h}v. \tag{12.3}$$
In order to pass to the limit in this equation, let $u:[0,\infty) \to X$ be the continuous function defined by
$$u(h) = \begin{cases} h^{-1}(T(h)-I)v & \text{if } h>0 \\ Lv & \text{if } h=0 \end{cases}.$$
Hence by same argument as in the proof of item 3 of Proposition 12.5), $h \to T(t-h)u(h)$ is continuous at $h=0$ and therefore,
$$\frac{T(t-h)-T(t)}{-h}v = T(t-h)u(h) \to T(t-0)u(0) = T(t)Lv \text{ as } h \downarrow 0.$$
So we have shown
$$\frac{d}{dh}|_{0-}T(t+h)v = T(t)Lv$$
which coupled with Eq. (12.2) shows
$$\frac{d}{dt}T(t)v = \frac{d}{dh}|_0 T(t+h)v = T(t)Lv = LT(t)v.$$

■

Proposition 12.7 (L is Closed). *Let L be the infinitesimal generator of a contraction semi-group, $T(t)$, then L is a densely defined closed operator on X.*

Proof. Suppose that $v_n \in \mathcal{D}(L)$, $v_n \to v$, and $Lv_n \to w$ in X as $n \to \infty$. By Theorem 12.6 and the fundamental theorem of calculus, we have
$$T(t)v_n - v_n = \int_0^t T(\tau)Lv_n d\tau.$$
Passing to the limit as $n \to \infty$ in this equation (using $T(\cdot)Lv_n \to T(\cdot)w$ uniformly) allows us to show,
$$T(t)v - v = \int_0^t T(\tau)w d\tau.$$
It then follows by the fundamental theorem of calculus (one sided version) that $v \in \mathcal{D}(L)$ and $Lv = \frac{d}{dt}|_{0+}T(t)v = w$, i.e. L is closed. ■

Definition 12.8 (Evolution Equation). *Let T be a strongly continuous contraction semi-group with infinitesimal generator L. A function $v:[0,\infty) \to X$ is said to solve the differential equation*
$$\dot{v}(t) = Lv(t) \tag{12.4}$$
if i) $v(t) \in \mathcal{D}(L)$ for all $t \ge 0$, ii) $v \in C([0,\infty) \to X) \cap C^1((0,\infty) \to X)$, and iii) Eq. (12.4) holds for all $t>0$.

Theorem 12.9 (Evolution Equation). *Let T be a strongly continuous contraction semi-group with infinitesimal generator L. The for all $v_0 \in \mathcal{D}(L)$, there is a unique solution to (12.4) such that $v(0) = v_0$.*

Proof. Existence. By Theorem 12.6 and Proposition 12.5, $v(t) := T(t)v_0$ solves (12.4.

Uniqueness. Let $v(t)$ be any solution of Eq. (12.4), $\tau > 0$ be given, and set $w(t) := T(\tau - t)v(t)$. By item 3 of Proposition 12.5, w is continuous for $t \in [0, \tau]$. We will now show that w is also differentiable on $(0, \tau)$ and that $\dot{w}(t) := 0$ for $t \in (0, \tau)$.

To simplify notation let $P(t) := T(\tau - t)$ and for fixed $t \in (0, \tau)$ and $h \neq 0$ but sufficiently close to 0, let

$$\varepsilon(h) := \frac{v(t+h) - v(t)}{h} - \dot{v}(t)$$

so that $\lim_{h \to 0} \varepsilon(h) = 0$ by definition of the derivative. With this notation we have

$$\begin{aligned}
\frac{w(t+h) - w(t)}{h} &= \frac{1}{h}(P(t+h)v(t+h) - P(t)v(t)) \\
&= \frac{(P(t+h) - P(t))}{h} v(t) + P(t+h)\frac{(v(t+h) - v(t))}{h} \\
&= \frac{(P(t+h) - P(t))}{h} v(t) + P(t+h)(\dot{v}(t) + \varepsilon(h)) \\
&\to -P(t)Lv(t) + P(t)\dot{v}(t) \text{ as } h \to 0,
\end{aligned}$$

wherein we have used

$$\|P(t+h)\varepsilon(h)\| \leq \|\varepsilon(h)\| \to 0 \text{ as } h \to 0.$$

Hence we have shown,

$$\dot{w}(t) = -P(t)Lv(t) + P(t)\dot{v}(t) = -P(t)Lv(t) + P(t)Lv(t) = 0,$$

and thus $w(t) = T(t-t)v(t)$ is constant or $(0, \tau)$. By the continuity of w on $[0, \tau]$ we may now conclude that $w(\tau) = w(0)$, i.e.

$$v(\tau) = w(\tau) = w(0) = T(\tau)v(0) = T(\tau)v_0.$$

This proves the only solutions to Eq. (12.4) with initial condition, v_0, is $v(t) = T(t)v_0$. ∎

Corollary 12.10. *Suppose that T and $\hat{T}$ are two strongly continuous contraction semi-groups on a Banach space X which have the same infinitesimal generators L. Then $T = \hat{T}$.*

Proof. Let $v_0 \in \mathcal{D}(L)$ then $v(t) = T(t)v$ and $\hat{v}(t) = \hat{T}(t)\hat{v}$ both solve Eq. (12.4 with initial condition v_0. By Theorem 12.9, $v = \hat{v}$ which implies that $T(t)v_0 = \hat{T}(t)v_0$, i.e., $T = \hat{T}$. ∎

Because of the last corollary the following notation is justified.

Notation 12.11 *If T is a strongly continuous contraction semi-group with infinitesimal generator L, we will write $T(t)$ as e^{tL}.*

If $T(t) = e^{tL}$ is a contraction semigroup we expect L to be "negative" (more precisely non-positive) in some sense and so working formally we expect to have, for all $\lambda > 0$, that

$$\int_0^\infty e^{-t\lambda}e^{tL}dt = \frac{1}{L - \lambda}e^{t(L-\lambda)}|_{t=0}^\infty = \frac{1}{\lambda - L} = (\lambda - L)^{-1}.$$

The next theorem justifies these hopes.

Theorem 12.12. *Suppose $T(t) = e^{tL}$ is a strongly continuous contraction semi-group with infinitesimal generator L. For any $\lambda > 0$ the integral*

$$\int_0^\infty e^{-t\lambda}e^{tL}dt =: R_\lambda \tag{12.5}$$

exists as a $B(X)$-valued Bochner-integral (or as an improper Riemann integral). Moreover, $(\lambda - L) : \mathcal{D}(L) \to X$ is an invertible operator, $(\lambda - L)^{-1} = R_\lambda$, and $\|R_\lambda\| \leq \lambda^{-1}$.

Proof. First notice that

$$\int_0^\infty e^{-t\lambda}\|e^{tL}\|dt \leq \int_0^\infty e^{-t\lambda}dt = 1/\lambda.$$

Therefore the integral in Eq. (12.5) exists and the result, R_λ, satisfies $\|R_\lambda\| \leq \lambda^{-1}$. So we now must show that $R_\lambda = (\lambda - L)^{-1}$.

Let $v \in X$ and $h > 0$, then

$$e^{hL}R_\lambda v = \int_0^\infty e^{-t\lambda}e^{(t+h)L}vdt = \int_h^\infty e^{-(t-h)\lambda}e^{tL}vdt = e^{h\lambda}\int_h^\infty e^{-t\lambda}e^{tL}vdt. \tag{12.6}$$

Therefore

$$\left.\frac{d}{dh}\right|_{0+} e^{hL}R_\lambda v = -v + \int_0^\infty \lambda e^{-t\lambda}e^{tL}vdt = -v + \lambda R_\lambda v,$$

which shows that $R_\lambda v \in \mathcal{D}(L)$ and that $LR_\lambda v = -v + \lambda R_\lambda v$, i.e. $(\lambda - L)R_\lambda = I$. Similarly,

$$R_\lambda e^{hL}v = e^{h\lambda}\int_h^\infty e^{-t\lambda}e^{tL}vdt \tag{12.7}$$

and hence if $v \in \mathcal{D}(L)$, then

$$R_\lambda Lv = \left.\frac{d}{dh}\right|_{0+} R_\lambda e^{hL}v = -v + \lambda R_\lambda v,$$

i.e. $R_\lambda(\lambda - L) = I_{\mathcal{D}(L)}$. ∎

12.2 The Hille-Yosida Theorem

Theorem 12.13 (Hille-Yosida). *A closed densely defined linear operator, L, on a Banach space, X, is the generates a contraction semi-group iff for all $\lambda \in (0, \infty)$;*

1. *$R_\lambda := (\lambda - L)^{-1}$ exists as a bounded operator and*
2. *$\|R_\lambda\| = \left\|(\lambda - L)^{-1}\right\| \le \frac{1}{\lambda} \quad \forall \lambda > 0$.*

Remark 12.14. In what follows we will freely use the fact that $LR_\lambda = R_\lambda L$ on $\mathcal{D}(L)$. Indeed, if $v \in \mathcal{D}(L)$, then

$$\begin{aligned} LR_\lambda v &= L(\lambda - L)^{-1} v = (L - \lambda + \lambda)(\lambda - L)^{-1} v \\ &= -v + \lambda(\lambda - L)^{-1} v = -v + \lambda R_\lambda v \end{aligned}$$

while

$$R_\lambda L v = R_\lambda (L - \lambda + \lambda) v = -v + \lambda R_\lambda v.$$

Proposition 12.15 (Approximators). *Let L be an operator on X satisfying properties 1. and 2. of the Hille-Yosida theorem and for each $\lambda > 0$, let*

$$L_\lambda := \lambda L R_\lambda = \text{“}\frac{\lambda L}{\lambda - L}.\text{”}$$

Then

$$L_\lambda = -\lambda + \lambda^2 R_\lambda, \tag{12.8}$$

$$\lambda R_\lambda \xrightarrow{s} I \quad \text{as} \quad \lambda \to \infty, \ \text{and} \tag{12.9}$$

$$\lim_{\lambda\to\infty} L_\lambda v = Lv \ \forall \ v \in \mathcal{D}(L). \tag{12.10}$$

Proof. The first identity is easy;

$$L_\lambda = \lambda L(\lambda - L)^{-1} = \lambda(L - \lambda + \lambda)(\lambda - L)^{-1} = -\lambda + \lambda^2 R_\lambda.$$

To prove Eq. (12.9), we will use, for $v \in \mathcal{D}(L)$, that

$$\begin{aligned} \lambda R_\lambda v &= (\lambda - L)^{-1} \lambda v \\ &= (\lambda - L)^{-1}(\lambda - L) v + (\lambda - L)^{-1} L v = v + R_\lambda L v. \end{aligned} \tag{12.11}$$

From this idenitty along with assumption 2., if follows that

$$\|\lambda R_\lambda v - v\| = \|R_\lambda L v\| \le \frac{1}{\lambda}\|Lv\| \to 0 \text{ as } \lambda \to \infty.$$

Equation (12.9), now follows from the previous equation, the fact that $\|\lambda R_\lambda\| \le 1$, along with the standard 3ε–argument. Using Eq. (12.9) along with Remark 12.14; if $v \in \mathcal{D}(L)$, then

$$L_\lambda v = \lambda L R_\lambda v = \lambda R_\lambda L v \to Lv \text{ as } \lambda \to \infty$$

and the proof is complete. ∎

Lemma 12.16. *If X is a Banach space and $A, B \in B(X)$, then for any $t \in \mathbb{R}$*

$$e^{tB} - e^{tA} = \int_0^t e^{(t-\tau)A}(B - A)e^{\tau B} d\tau.$$

If we further assume that A and B commute, then for each $v \in X$,

$$\left\|\left(e^{tB} - e^{tA}\right) v\right\| \le \left|\int_0^t \left\|e^{(t-\tau)A}\right\| \left\|e^{\tau B}\right\| d\tau\right| \cdot \|(A - B) v\| \tag{12.12}$$

$$\le |t| \cdot M_A(t) M_B(t) \|(A - B) v\| \tag{12.13}$$

where

$$M_A(t) := \sup\left\{\left\|e^{\tau A}\right\| : \tau \text{ between } 0 \text{ and } t\right\}$$

with an analogous definition for $M_B(t)$. In particular, if $\|e^{tA}\|$ and $\|e^{tB}\|$ are bounded by 1 for all $t > 0$, then

$$\left\|\left(e^{tA} - e^{tB}\right) v\right\| \le t \|(A - B) v\| \ \text{ for all } v \in X. \tag{12.14}$$

Proof. By the fundamental theorem of calculus and the product rule we have,

$$e^{-tA} e^{tB} - I = \int_0^t \frac{d}{d\tau} e^{-\tau A} e^{\tau B} d\tau = \int_0^t e^{-\tau A}(-A + B)e^{\tau B} d\tau.$$

Multiplying this equation on the left by e^{tA} shows,

$$e^{tB} - e^{tA} = \int_0^t e^{(t-\tau)A}(B - A)e^{\tau B} d\tau.$$

If we now further assume that $[A, B] = 0$, the previously displayed equation may be written as

$$e^{tB} - e^{tA} = \int_0^t e^{(t-\tau)A} e^{\tau B}(B - A) d\tau.$$

Applying this identity to $v \in X$ and then taking norms and using the triangle inequality for integrals gives Eq. (12.12) which also clearly implies Eqs. (12.13) and (12.14). ∎

With this preparation we are now ready for the proof of Hille–Yoshida Theorem 12.13.

Proof. Hille–Yoshida Theorem 12.13. For $\lambda > 0$, let

$$T_\lambda(t) := e^{tL_\lambda} = e^{tL_\lambda} := \sum_{n=0}^{\infty} \frac{t^n}{n!} L_\lambda^n.$$

The outline of the proof is: i) show that $T_\lambda(t)$ is a contraction for all $t > 0$, ii) show for $t > 0$ that $T_\lambda(t)$ converges strongly to an operator $T(t)$, iii) we show $T(t)$ is a strongly continuous contraction semi-group, and iv) the generator of $T(t)$ is L.

Step i) Using $L_\lambda = -\lambda + \lambda^2 R_\lambda$ (see Eq. (12.8) we find that $e^{tL_\lambda} = e^{-t\lambda} e^{t\lambda^2 R_\lambda}$ and hence

$$\|T_\lambda(t)\| = \|e^{tL_\lambda}\| \le e^{-t\lambda} e^{t\lambda^2 \|R_\lambda\|} \le e^{-t\lambda} e^{t\lambda^2 \lambda^{-1}} = 1$$

wherein we have used assumtion 2. to conclude $\lambda \|R_\lambda\| \le 1$ for all $\lambda > 0$.

Step ii) Let $\alpha, \mu > 0$ and $v \in \mathcal{D}(L)$, then by Lemma 12.16 and Proposition 12.15,

$$\|(T_\alpha(t) - T_\mu(t))v\| \le t \|L_\alpha v - L_\mu v\| \to 0 \text{ as } \alpha, \mu \to \infty.$$

This shows, for all $v \in \mathcal{D}(L)$, that $\lim_{\alpha\to\infty} T_\alpha(t) v$ exists uniformly for t in compact subsets of $[0,\infty)$. For general $v \in X$, $w \in \mathcal{D}(L)$, $\tau > 0$, and $0 \le t \le \tau$, we have

$$\begin{aligned}\|(T_\alpha(t) - T_\mu(t))v\| &\le \|(T_\alpha(t) - T_\mu(t))w\| + \|(T_\alpha(t) - T_\mu(t))(v-w)\| \\ &\le \|(T_\alpha(t) - T_\mu(t))w\| + 2\|v-w\|. \\ &\le \tau \|L_\alpha w - L_\mu w\| + 2\|v-w\|.\end{aligned}$$

Thus

$$\limsup_{\alpha,\mu\to\infty} \sup_{t\in[0,\tau]} \|(T_\alpha(t) - T_\mu(t))v\| \le 2\|v-w\| \to 0 \text{ as } w \to v.$$

Hence for each $v \in X$, $T(t) v := \lim_{\alpha\to\infty} T_\alpha(t) v$ exists uniformly for t in compact sets of $[0,\infty)$.

Step iii) It is now easily follows that $\|T(t)\| \le 1$ and that $t \to T(t)$ is strongly continuous. Let us now fix $s, t > 0$ and $v \in X$ and note that

$$\varepsilon(\alpha) := T_\alpha(s) v - T(s) v \to 0 \text{ as } \alpha \downarrow 0.$$

We then have

$$T_\alpha(t+s)v = T_\alpha(t) T_\alpha(s) v = T_\alpha(t) T(s) v + T_\alpha(t) \varepsilon(\alpha).$$

Passing to the limit as $\alpha \downarrow 0$ in this identity then shows $T(t+s)v = T(t) T(s) v$ ((i.e. is a semi-group) where we have used

$$\|T_\alpha(t) \varepsilon(\alpha)\| \le \|\varepsilon(\alpha)\| \to 0 \text{ as } \alpha \downarrow 0.$$

Step iv) Let $\tilde{L} = \frac{d}{dt}|_{0^+} T(t)$ denote the infinitesimal generator of T. We are going to finish the proof by showing $\tilde{L} = L$.

If $v \in \mathcal{D}(L)$ and $\lambda > 0$, then

$$T_\lambda(t) v = e^{tL_\lambda} v = v + \int_0^t e^{\tau L_\lambda} L_\lambda v d\tau = v + \int_0^t T_\lambda(\tau) L_\lambda v d\tau. \tag{12.15}$$

Let us note that

$$\begin{aligned}\|T_\lambda(\tau) L_\lambda v - T(\tau) Lv\| &\le \|T_\lambda(\tau)[L_\lambda v - Lv]\| + \|T_\lambda(\tau) Lv - T(\tau) Lv\| \\ &\le \|L_\lambda v - Lv\| + \|T_\lambda(\tau) Lv - T(\tau) Lv\|\end{aligned}$$

and hence

$$\begin{aligned}&\max_{0\le\tau\le t} \|T_\lambda(\tau) L_\lambda v - T(\tau) Lv\| \\ &\qquad \le \|L_\lambda v - Lv\| + \max_{0\le\tau\le t} \|T_\lambda(\tau) Lv - T(\tau) Lv\| \to 0 \text{ as } \lambda \downarrow 0,\end{aligned}$$

wherein we have used Eq. (12.10) and step ii to deduce the limit. With this result in hand we may let $\lambda \downarrow 0$ in Eq. (12.15) in order to conclude,

$$T(t) v = v + \int_0^t T(\tau) Lv \, d\tau \text{ for all } v \in \mathcal{D}(L).$$

It then follows by the fundamental theorem of calculus that $\tilde{L} v = \frac{d}{dt}|_{0^+} T(t) v = Lv$, i.e. we have shown $L \subset \tilde{L}$ and therefore $\lambda - L \subset \lambda - \tilde{L}$ for any $\lambda > 0$. However both $\lambda - L$ and $\lambda - \tilde{L}$ are invertible and hence by the simple Exercise 11.1 with $A = \lambda - \tilde{L}$ and $B = \lambda - L$, it follows that $A = B$, i.e. $L = \tilde{L}$.

For the skeptical reader: here is a direct proof of the last part of the argument. Suppose that $\tilde{v} \in \mathcal{D}(\tilde{L})$. Fix $\lambda > 0$ and let $v := (\lambda - L)^{-1} \left(\lambda - \tilde{L}\right) \tilde{v} \in \mathcal{D}(L)$ so that $(\lambda - L) v = \left(\lambda - \tilde{L}\right) \tilde{v}$. Since $L \subset \tilde{L}$, we may conclude

$$\left(\lambda - \tilde{L}\right) v = (\lambda - L) v = \left(\lambda - \tilde{L}\right) \tilde{v}$$

and because $\lambda - \tilde{L}$ is invertible it follows that $\tilde{v} = v \in \mathcal{D}(L)$. ■

Part V

References and Apppendicies

References

1. Ralph Abraham and Jerrold E. Marsden, *Foundations of mechanics*, Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, Reading, Mass., 1978, Second edition, revised and enlarged, With the assistance of Tudor Ratiu and Richard Cushman. MR 81e:58025
2. T. Ando and F. Hiai, *Inequality between powers of positive semidefinite matrices*, Linear Algebra Appl. **208/209** (1994), 65–71. MR 1287340 (95k:15023)
3. Frank F. Bonsall and John Duncan, *Complete normed algebras*, Springer-Verlag, New York-Heidelberg, 1973, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 80. MR 0423029
4. J. Brüning and M. Lesch, *Hilbert complexes*, J. Funct. Anal. **108** (1992), no. 1, 88–132. MR MR1174159 (93k:58208)
5. Edward L. Bueler, *The heat kernel weighted Hodge Laplacian on noncompact manifolds*, Trans. Amer. Math. Soc. **351** (1999), no. 2, 683–713. MR MR1443866 (99d:58164)
6. Paul R. Chernoff, *Essential self-adjointness of powers of generators of hyperbolic equations*, J. Functional Analysis **12** (1973), 401–414. MR 0369890 (51 #6119)
7. John B. Conway, *A course in functional analysis*, second ed., Graduate Texts in Mathematics, vol. 96, Springer-Verlag, New York, 1990. MR 1070713
8. Claude Dellacherie, *Capacités et processus stochastiques*, Springer-Verlag, Berlin, 1972, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 67. MR MR0448504 (56 #6810)
9. Robert Doran, *Characterizations of c* algebras: the gelfand naimark theorems*, CRC Press, 05 2019.
10. Bruce K. Driver and Leonard Gross, *Hilbert spaces of holomorphic functions on complex Lie groups*, New trends in stochastic analysis (Charingworth, 1994), World Sci. Publishing, River Edge, NJ, 1997, pp. 76–106. MR MR1654507 (2000h:46029)
11. Gerald B. Folland, *Harmonic analysis in phase space*, Annals of Mathematics Studies, vol. 122, Princeton University Press, Princeton, NJ, 1989. MR 983366 (92k:22017)
12. Masatoshi Fujii and Yuki Seo, *Reverse inequalities of Araki, Cordes and Löwner-Heinz inequalities*, Nihonkai Math. J. **16** (2005), no. 2, 145–154. MR 2217555 (2006k:47036)
13. Matthew P. Gaffney, *The harmonic operator for exterior differential forms*, Proc. Nat. Acad. Sci. U. S. A. **37** (1951), 48–50. MR MR0048138 (13,987b)
14. ———, *The conservation property of the heat equation on Riemannian manifolds.*, Comm. Pure Appl. Math. **12** (1959), 1–11. MR MR0102097 (21 #892)
15. Robert F. Geitz, *Pettis integration*, Proc. Amer. Math. Soc. **82** (1981), no. 1, 81–86. MR 603606
16. ———, *Geometry and the Pettis integral*, Trans. Amer. Math. Soc. **269** (1982), no. 2, 535–548. MR 637707
17. Robert F. Geitz and J. J. Uhl, Jr., *Vector-valued functions as families of scalar-valued functions*, Pacific J. Math. **95** (1981), no. 1, 75–83. MR 631660
18. Robert Frederick Geitz, *THE PETTIS INTEGRAL*, ProQuest LLC, Ann Arbor, MI, 1980, Thesis (Ph.D.)–University of Illinois at Urbana-Champaign. MR 2630764
19. Brian C. Hall, *Phase space bounds for quantum mechanics on a compact Lie group*, Comm. Math. Phys. **184** (1997), no. 1, 233–250.
20. ———, *Quantum mechanics in phase space*, Perspectives on quantization (South Hadley, MA, 1996), Amer. Math. Soc., Providence, RI, 1998, pp. 47–62. MR 99e:22015

21. ———, *Holomorphic methods in analysis and mathematical physics*, First Summer School in Analysis and Mathematical Physics (Cuernavaca Morelos, 1998), Contemp. Math., vol. 260, Amer. Math. Soc., Providence, RI, 2000, pp. 1–59. MR 1770752
22. Erhard Heinz, *Beiträge zur Störungstheorie der Spektralzerlegung*, Math. Ann. **123** (1951), 415–438. MR 0044747 (13,471f)
23. Svante Janson, *Gaussian Hilbert spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997. MR MR1474726 (99f:60082)
24. Tosio Kato, *Notes on some inequalities for linear operators*, Math. Ann. **125** (1952), 208–212. MR 0053390 (14,766e)
25. Karl Löwner, *Über monotone Matrixfunktionen*, Math. Z. **38** (1934), no. 1, 177–216. MR 1545446
26. Laurent W. Marcoux, *An introduction to operator algebras*, University of Waterloo preprint book, 2016.
27. Theodore W. Palmer, *Banach algebras and the general theory of* **-algebras. Vol. I*, Encyclopedia of Mathematics and its Applications, vol. 49, Cambridge University Press, Cambridge, 1994, Algebras and Banach algebras. MR 1270014
28. ———, *Banach algebras and the general theory of* $*$*-algebras. Vol. 2*, Encyclopedia of Mathematics and its Applications, vol. 79, Cambridge University Press, Cambridge, 2001, $*$-algebras. MR 1819503
29. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983. MR 710486
30. Josip Pečarić, Takayuki Furuta, Jadranka Mićić Hot, and Yuki Seo, *Mond-Pečarić method in operator inequalities*, Monographs in Inequalities, vol. 1, ELEMENT, Zagreb, 2005, Inequalities for bounded selfadjoint operators on a Hilbert space. MR 3026316
31. Gert K. Pedersen, *Some operator monotone functions*, Proc. Amer. Math. Soc. **36** (1972), 309–310. MR 0306957 (46 #6078)
32. A. Prasad, *An easy proof of the Stone-von Neumann-Mackey theorem*, ArXiv e-prints (2009), 1–10.
33. Ian F. Putnam, *Lecture notes on c-star-algebras*, http://www.math.uvic.ca/faculty/putnam/ln/C*-algebras.pdf, 2019.
34. Michael Reed and Barry Simon, *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975. MR 5812429b
35. ———, *Methods of modern mathematical physics. III*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1979, Scattering theory. MR 80m:81085
36. ———, *Methods of modern mathematical physics. I*, second ed., Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980, Functional analysis. MR 85e:46002
37. Franz Rellich, *Halbbeschränkte gewöhnliche Differentialoperatoren zweiter Ordnung*, Math. Ann. **122** (1951), 343–368. MR 0043316 (13,240g)
38. Walter Roelcke, *Über den Laplace-Operator auf Riemannschen Mannigfaltigkeiten mit diskontinuierlichen Gruppen*, Math. Nachr. **21** (1960), 131–149. MR MR0151927 (27 #1908)
39. Jonathan Rosenberg, *A selective history of the Stone-von Neumann theorem*, Operator algebras, quantization, and noncommutative geometry, Contemp. Math., vol. 365, Amer. Math. Soc., Providence, RI, 2004, pp. 331–353. MR 2106827 (2005h:46001)
40. Walter Rudin, *Principles of mathematical analysis*, third ed., McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976, International Series in Pure and Applied Mathematics. MR 0385023
41. ———, *Real and complex analysis*, third ed., McGraw-Hill Book Co., New York, 1987. MR 924157
42. Robert Schatten, *Norm ideals of completely continuous operators*, Second printing. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 27, Springer-Verlag, Berlin, 1970. MR 41 #2449
43. Konrad Schmüdgen, *Unbounded self-adjoint operators on Hilbert space*, Graduate Texts in Mathematics, vol. 265, Springer, Dordrecht, 2012. MR 2953553
44. B. Simon, *Convergence in trace ideals*, Proc. Amer. Math. Soc. **83** (1981), no. 1, 39–43. MR 82h:47042
45. Barry Simon, *Trace ideals and their applications*, London Mathematical Society Lecture Note Series, vol. 35, Cambridge University Press, Cambridge, 1979. MR 80k:47048
46. Elias M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, Princeton, NJ, 1993, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. MR 95c:42002
47. Robert S. Strichartz, *Analysis of the Laplacian on the complete Riemannian manifold*, J. Funct. Anal. **52** (1983), no. 1, 48–79. MR 84m:58138
48. Michel Talagrand, *Pettis integral and measure theory*, Mem. Amer. Math. Soc. **51** (1984), no. 307, ix+224. MR 756174
49. Michael E. Taylor, *Partial differential equations I. Basic theory*, second ed., Applied Mathematical Sciences, vol. 115, Springer, New York, 2011. MR 2744150
50. J. v. Neumann, *Die Eindeutigkeit der Schrödingerschen Operatoren*, Math. Ann. **104** (1931), no. 1, 570–578. MR 1512685
51. Léon Van Hove, *Sur certaines représentations unitaires d'un groupe infini de transformations*, Acad. Roy. Belgique. Cl. Sci. Mém. Coll. in 8° **26** (1951), no. 6, 102. MR 15,198d
52. Dana P. Williams, *A (very) short course on c-star algebras*, https://math.dartmouth.edu/ dana/bookspapers/cstar.pdf, 2019.
53. Ke He Zhu, *An introduction to operator algebras*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1993. MR 1257406

A

Miscellaneous Background Results

A.1 Multiplicative System Theorems

Notation A.1 *Let Ω be a set and $\mathbb{H}$ be a subset of the bounded real valued functions on Ω. We say that $\mathbb{H}$ is* ***closed under bounded convergence*** *if; for every sequence, $\{f_n\}_{n=1}^{\infty} \subset \mathbb{H}$, satisfying:*

1. *there exists $M < \infty$ such that $|f_n(\omega)| \le M$ for all $\omega \in \Omega$ and $n \in \mathbb{N}$,*
2. *$f(\omega) := \lim_{n\to\infty} f_n(\omega)$ exists for all $\omega \in \Omega$, then $f \in \mathbb{H}$.*

Notation A.2 *For any σ-algebra, $\mathcal{B} \subset 2^{\Omega}$, let $\mathbb{B}(\Omega, \mathcal{B}; \mathbb{R})$ be the bounded $\mathcal{B}/\mathcal{B}_{\mathbb{R}}$-measurable functions from Ω to $\mathbb{R}$.*

Notation A.3 *If $\mathbb{M}$ is any subset of $\mathbb{B}\left(\Omega, 2^{\Omega}; \mathbb{R}\right)$, let $\mathbb{H}(\mathbb{M})$ denote the smallest subspace of bounded functions on Ω which contains $\mathbb{M} \cup \{1\}$. (As usual such a space exists by taking the intersection of all such spaces.)*

Definition A.4. *A subset, $\mathbb{M} \subset \mathbb{B}\left(\Omega, 2^{\Omega}; \mathbb{R}\right)$, is called a* ***multiplicative system*** *if $\mathbb{M}$ is closed under finite products, i.e. $f, g \in \mathbb{M}$, then $f \cdot g \in \mathbb{M}$.*

The following result may be found in Dellacherie [8, p. 14]. The style of proof given here may be found in Janson [23, Appendix A., p. 309].

Theorem A.5 (Dynkin's Multiplicative System Theorem). *Suppose that $\mathbb{H}$ is a vector subspace of bounded functions from Ω to $\mathbb{R}$ which contains the constant functions and is closed under bounded convergence. If $\mathbb{M} \subset \mathbb{H}$ is a multiplicative system, then $\mathbb{H}$ contains all bounded $\sigma(\mathbb{M})$ – measurable functions, i.e. $\mathbb{H}$ contains $\mathbb{B}(\Omega, \sigma(\mathbb{M}); \mathbb{R})$.*

Proof. We are going to in fact prove: if $\mathbb{M} \subset \mathbb{B}\left(\Omega, 2^{\Omega}; \mathbb{R}\right)$ is a multiplicative system, then $\mathbb{H}(\mathbb{M}) = \mathbb{B}(\Omega, \sigma(\mathbb{M}); \mathbb{R})$. This suffices to prove the theorem as $\mathbb{H}(\mathbb{M}) \subset \mathbb{H}$ is contained in $\mathbb{H}$ by very definition of $\mathbb{H}(\mathbb{M})$. To simplify notation let us now assume that $\mathbb{H} = \mathbb{H}(\mathbb{M})$. The remainder of the proof will be broken into five steps.

Step 1. ($\mathbb{H}$ is an algebra of functions.) For $f \in \mathbb{H}$, let $\mathbb{H}^f := \{g \in \mathbb{H} : gf \in \mathbb{H}\}$. The reader will now easily verify that $\mathbb{H}^f$ is a linear subspace of $\mathbb{H}$, $1 \in \mathbb{H}^f$, and $\mathbb{H}^f$ is closed under bounded convergence. Moreover if $f \in \mathbb{M}$, since $\mathbb{M}$ is a multiplicative system, $\mathbb{M} \subset \mathbb{H}^f$. Hence by the definition of $\mathbb{H}$, $\mathbb{H} = \mathbb{H}^f$, i.e. $fg \in \mathbb{H}$ for all $f \in \mathbb{M}$ and $g \in \mathbb{H}$. Having proved this it now follows for any $f \in \mathbb{H}$ that $\mathbb{M} \subset \mathbb{H}^f$ and therefore as before, $\mathbb{H}^f = \mathbb{H}$. Thus we may conclude that $fg \in \mathbb{H}$ whenever $f, g \in \mathbb{H}$, i.e. $\mathbb{H}$ is an algebra of functions.

Step 2. ($\mathcal{B} := \{A \subset \Omega : 1_A \in \mathbb{H}\}$ is a σ – algebra.) Using the fact that $\mathbb{H}$ is an algebra containing constants, the reader will easily verify that $\mathcal{B}$ is closed under complementation, finite intersections, and contains Ω, i.e. $\mathcal{B}$ is an algebra. Using the fact that $\mathbb{H}$ is closed under bounded convergence, it follows that $\mathcal{B}$ is closed under increasing unions and hence that $\mathcal{B}$ is σ – algebra.

Step 3. ($\mathbb{B}(\Omega, \mathcal{B}; \mathbb{R}) \subset \mathbb{H}$) Since $\mathbb{H}$ is a vector space and $\mathbb{H}$ contains 1_A for all $A \in \mathcal{B}$, $\mathbb{H}$ contains all $\mathcal{B}$ – measurable simple functions. Since every bounded $\mathcal{B}$ – measurable function may be written as a bounded limit of such simple functions, it follows that $\mathbb{H}$ contains all bounded $\mathcal{B}$ – measurable functions.

Step 4. ($\sigma(\mathbb{M}) \subset \mathcal{B}$.) Let $\varphi_n(x) = 0 \vee [(nx) \wedge 1]$ (see Figure A.1 below) so that $\varphi_n(x) \uparrow 1_{x>0}$. Given $f \in \mathbb{M}$ and $a \in \mathbb{R}$, let $F_n := \varphi_n(f - a)$ and $M := \sup_{\omega\in\Omega} |f(\omega) - a|$. By the Weierstrass approximation theorem, we may find polynomial functions, $p_l(x)$ such that $p_l \to \varphi_n$ uniformly on $[-M, M]$. Since p_l is a polynomial and $\mathbb{H}$ is an algebra, $p_l(f - a) \in \mathbb{H}$ for all l. Moreover, $p_l \circ (f - a) \to F_n$ uniformly as $l \to \infty$, from with it follows that $F_n \in \mathbb{H}$ for all n. Since, $F_n \uparrow 1_{\{f>a\}}$ it follows that $1_{\{f>a\}} \in \mathbb{H}$, i.e. $\{f > a\} \in \mathcal{B}$. As the sets $\{f > a\}$ with $a \in \mathbb{R}$ and $f \in \mathbb{M}$ generate $\sigma(\mathbb{M})$, it follows that $\sigma(\mathbb{M}) \subset \mathcal{B}$.

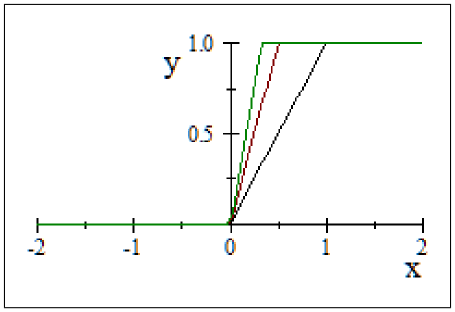

Fig. A.1. Plots of φ_1, φ_2 and φ_3 which are continuous functions used to approximate, $x \to 1_{x\ge0}$.

Step 5. $(\mathbb{H}(\mathbb{M}) = \mathbb{B}(\Omega, \sigma(\mathbb{M}); \mathbb{R}).)$ By step 4., $\sigma(\mathbb{M}) \subset \mathcal{B}$, and so $\mathbb{B}(\Omega, \sigma(\mathbb{M}); \mathbb{R}) \subset \mathbb{B}(\Omega, \mathcal{B}; \mathbb{R})$ which combined with step 3. shows,

$$\mathbb{B}(\Omega, \sigma(\mathbb{M}); \mathbb{R}) \subset \mathbb{B}(\Omega, \mathcal{B}; \mathbb{R}) \subset \mathbb{H}(\mathbb{M}).$$

However, we know that $\mathbb{B}(\Omega, \sigma(\mathbb{M}); \mathbb{R})$ is a subspace of bounded measurable functions containing $\mathbb{M}$ and therefore $\mathbb{H}(\mathbb{M}) \subset \mathbb{B}(\Omega, \sigma(\mathbb{M}); \mathbb{R})$ which suffices to complete the proof. ■

Corollary A.6. *Suppose $\mathbb{H}$ is a subspace of bounded real valued functions such that $1 \in \mathbb{H}$ and $\mathbb{H}$ is closed under bounded convergence. If $\mathcal{P} \subset 2^{\Omega}$ is a multiplicative class such that $1_A \in \mathbb{H}$ for all $A \in \mathcal{P}$, then $\mathbb{H}$ contains all bounded $\sigma(\mathcal{P})$ – measurable functions.*

Proof. Let $\mathbb{M} = \{1\} \cup \{1_A : A \in \mathcal{P}\}$. Then $\mathbb{M} \subset \mathbb{H}$ is a multiplicative system and the proof is completed with an application of Theorem A.5. ■

Example A.7. Suppose μ and ν are two probability measure on $(\Omega, \mathcal{B})$ such that

$$\int_{\Omega} f d\mu = \int_{\Omega} f d\nu \tag{A.1}$$

for all f in a multiplicative subset, $\mathbb{M}$, of bounded measurable functions on Ω. Then $\mu = \nu$ on $\sigma(\mathbb{M})$. Indeed, apply Theorem A.5 with $\mathbb{H}$ being the bounded measurable functions on Ω such that Eq. (A.1) holds. In particular if $\mathbb{M} = \{1\} \cup \{1_A : A \in \mathcal{P}\}$ with $\mathcal{P}$ being a multiplicative class we learn that $\mu = \nu$ on $\sigma(\mathbb{M}) = \sigma(\mathcal{P})$.

Exercise A.1. Let $\Omega := \{1, 2, 3, 4\}$ and $\mathbb{M} := \{1_A, 1_B\}$ where $A := \{1, 2\}$ and $B := \{2, 3\}$.

a) Show $\sigma(\mathbb{M}) = 2^{\Omega}$.
b) Find two distinct probability measures, μ and ν on 2^{Ω} such that $\mu(A) = \nu(A)$ and $\mu(B) = \nu(B)$, i.e. Eq. (A.1) holds for all $f \in \mathbb{M}$.

Moral: the assumption that $\mathbb{M}$ is multiplicative can not be dropped from Theorem A.5.

Proposition A.8. *Suppose μ and ν are two measures on $(\Omega, \mathcal{B})$, $\mathcal{P} \subset \mathcal{B}$ is a multiplicative system (i.e. closed under intersections as in Definition* **??***) such that $\mu(A) = \nu(A)$ for all $A \in \mathcal{P}$. If there exists $\Omega_n \in \mathcal{P}$ such that $\Omega_n \uparrow \Omega$ and $\mu(\Omega_n) = \nu(\Omega_n) < \infty$, then $\mu = \nu$ on $\sigma(\mathcal{P})$.*

Proof. Step 1. First assume that $\mu(\Omega) = \nu(\Omega) < \infty$ and then apply Example A.7 with $\mathbb{M} = \{1_A : A \in \mathcal{P}\}$ in order to find $\mu = \nu$ on $\sigma(\mathbb{M}) = \sigma(\mathcal{P})$.

Step 2. For the general case let $\mu_n(B) := \mu(B \cap \Omega_n)$ and $\nu_n(B) := \nu(B \cap \Omega_n)$ for all $B \in \mathcal{B}$. Then $\mu_n = \nu_n$ on $\mathcal{P}$ (because $\Omega_n \in \mathcal{P}$) and

$$\mu_n(\Omega) = \mu(\Omega_n) = \nu(\Omega_n) = \nu_n(\Omega).$$

Therefore by step 1, $\mu_n = \nu_n$ on $\sigma(\mathcal{P})$. Passing to the limit as $n \to \infty$ then shows

$$\begin{aligned} \mu(B) &= \lim_{n\to\infty} \mu(B \cap \Omega_n) = \lim_{n\to\infty} \mu_n(B) \\ &= \lim_{n\to\infty} \nu_n(B) = \lim_{n\to\infty} \nu(B \cap \Omega_n) = \nu(B) \end{aligned}$$

for all $B \in \sigma(\mathcal{P})$. ■

Here is a complex version of Theorem A.5.

Theorem A.9 (Complex Multiplicative System Theorem). *Suppose $\mathbb{H}$ is a complex linear subspace of the bounded complex functions on Ω, $1 \in \mathbb{H}$, $\mathbb{H}$ is closed under complex conjugation, and $\mathbb{H}$ is closed under bounded convergence. If $\mathbb{M} \subset \mathbb{H}$ is multiplicative system which is closed under conjugation, then $\mathbb{H}$ contains all bounded complex valued $\sigma(\mathbb{M})$-measurable functions.*

Proof. Let $\mathbb{M}_0 = \mathrm{span}_{\mathbb{C}}(\mathbb{M} \cup \{1\})$ be the complex span of $\mathbb{M}$. As the reader should verify, $\mathbb{M}_0$ is an algebra, $\mathbb{M}_0 \subset \mathbb{H}$, $\mathbb{M}_0$ is closed under complex conjugation and $\sigma(\mathbb{M}_0) = \sigma(\mathbb{M})$. Let

$$\begin{aligned} \mathbb{H}^{\mathbb{R}} &:= \{f \in \mathbb{H} : f \text{ is real valued}\} \text{ and} \\ \mathbb{M}_0^{\mathbb{R}} &:= \{f \in \mathbb{M}_0 : f \text{ is real valued}\}. \end{aligned}$$

Then $\mathbb{H}^{\mathbb{R}}$ is a real linear space of bounded real valued functions 1 which is closed under bounded convergence and $\mathbb{M}_0^{\mathbb{R}} \subset \mathbb{H}^{\mathbb{R}}$. Moreover, $\mathbb{M}_0^{\mathbb{R}}$ is a multiplicative system (as the reader should check) and therefore by Theorem A.5, $\mathbb{H}^{\mathbb{R}}$ contains all bounded $\sigma(\mathbb{M}_0^{\mathbb{R}})$ – measurable real valued functions. Since $\mathbb{H}$ and $\mathbb{M}_0$ are complex linear spaces closed under complex conjugation, for any $f \in \mathbb{H}$ or $f \in \mathbb{M}_0$, the functions $\mathrm{Re}\, f = \frac{1}{2}(f + \bar{f})$ and $\mathrm{Im}\, f = \frac{1}{2i}(f - \bar{f})$ are in $\mathbb{H}$ or $\mathbb{M}_0$ respectively. Therefore $\mathbb{M}_0 = \mathbb{M}_0^{\mathbb{R}} + i\mathbb{M}_0^{\mathbb{R}}$, $\sigma(\mathbb{M}_0^{\mathbb{R}}) = \sigma(\mathbb{M}_0) = \sigma(\mathbb{M})$, and $\mathbb{H} = \mathbb{H}^{\mathbb{R}} + i\mathbb{H}^{\mathbb{R}}$. Hence if $f : \Omega \to \mathbb{C}$ is a bounded $\sigma(\mathbb{M})$ – measurable function, then $f = \mathrm{Re}\, f + i\, \mathrm{Im}\, f \in \mathbb{H}$ since $\mathrm{Re}\, f$ and $\mathrm{Im}\, f$ are in $\mathbb{H}^{\mathbb{R}}$. ■

Lemma A.10. *If $-\infty < a < b < \infty$, there exists $f_n \in C_c(\mathbb{R}, [0, 1])$ such that $\lim_{n\to\infty} f_n = 1_{(a,b]}$.*

Proof. The reader should verify $\lim_{n\to\infty} f_n = 1_{(a,b]}$ where $f_n \in C_c(\mathbb{R}, [0, 1])$ is defined (for n sufficiently large) by

$$f_n(x) := \begin{cases} 0 & \text{on } (-\infty, a] \cup [b+\frac{1}{n}, \infty) \\ n(x-a) & \text{if } \quad a \le x \le a+\frac{1}{n} \\ 1 & \text{if } \quad a+\frac{1}{n} \le x \le b \\ 1-n(b-x) & \text{if } \quad b \le x \le b+\frac{1}{n} \end{cases}.$$

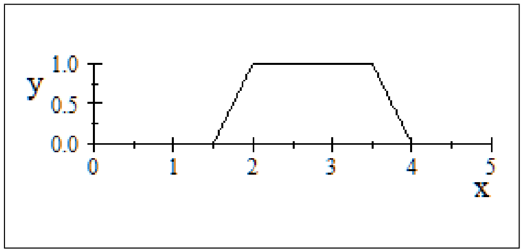

Fig. A.2. Here is a plot of $f_2(x)$ when $a = 1.5$ and $b = 3.5$.

■

Lemma A.11. *For each $\lambda > 0$, let $e_\lambda(x) := e^{i\lambda x}$. Then*

$$\mathcal{B}_{\mathbb{R}} = \sigma(e_\lambda : \lambda > 0) = \sigma\left(e_\lambda^{-1}(W) : \lambda > 0,\ W \in \mathcal{B}_{\mathbb{R}}\right).$$

Proof. Let $S^1 := \{z \in \mathbb{C} : |z| = 1\}$. For $-\pi < \alpha < \beta < \pi$ let

$$A(\alpha, \beta) := \left\{e^{i\theta} : \alpha < \theta < \beta\right\} = S^1 \cap \left\{re^{i\theta} : \alpha < \theta < \beta,\ r > 0\right\}$$

which is a measurable subset of $\mathbb{C}$ (why). Moreover we have $e_\lambda(x) \in A(\alpha, \beta)$ iff $\lambda x \in \sum_{n\in\mathbb{Z}} [(\alpha, \beta) + 2\pi n]$ and hence

$$e_\lambda^{-1}(A(\alpha, \beta)) = \sum_{n\in\mathbb{Z}} \left[\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right) + 2\pi\frac{n}{\lambda}\right] \in \sigma(e_\lambda : \lambda > 0).$$

Hence if $-\infty < a < b < \infty$ and $\lambda > 0$ sufficiently small so that $-\pi < \lambda a < \lambda b < \pi$, we have

$$e_\lambda^{-1}(A(\lambda a, \lambda b)) = \sum_{n\in\mathbb{Z}} \left[(a, b) + 2\pi\frac{n}{\lambda}\right]$$

and hence

$$(a, b) = \cap_{\lambda > 0} e_\lambda^{-1}(A(\lambda a, \lambda b)) \in \sigma(e_\lambda : \lambda > 0).$$

This shows $\mathcal{B}_{\mathbb{R}} \subset \sigma(e_\lambda : \lambda > 0)$. As e_λ is continuous and hence Borel measurable for all $\lambda > 0$ we automatically know that $\sigma(e_\lambda : \lambda > 0) \subset \mathcal{B}_{\mathbb{R}}$. ■

Remark A.12. A slight modification of the above proof actually shows if $\{\lambda_n\} \subset (0, \infty)$ with $\lim_{n\to\infty} \lambda_n = 0$, then $\sigma(e_{\lambda_n} : n \in \mathbb{N}) = \mathcal{B}_{\mathbb{R}}$.

Corollary A.13. *Each of the following σ – algebras on $\mathbb{R}^d$ are equal to $\mathcal{B}_{\mathbb{R}^d}$;*

1. $\mathcal{M}_1 := \sigma\left(\cup_{i=1}^n \{x \to f(x_i) : f \in C_c(\mathbb{R})\}\right)$,
2. $\mathcal{M}_2 := \sigma(x \to f_1(x_1) \dots f_d(x_d) : f_i \in C_c(\mathbb{R}))$
3. $\mathcal{M}_3 = \sigma\left(C_c\left(\mathbb{R}^d\right)\right)$, *and*
4. $\mathcal{M}_4 := \sigma\left(\left\{x \to e^{i\lambda\cdot x} : \lambda \in \mathbb{R}^d\right\}\right)$.

Proof. As the functions defining each $\mathcal{M}_i$ are continuous and hence Borel measurable, it follows that $\mathcal{M}_i \subset \mathcal{B}_{\mathbb{R}^d}$ for each i. So to finish the proof it suffices to show $\mathcal{B}_{\mathbb{R}^d} \subset \mathcal{M}_i$ for each i.

$\mathcal{M}_1$ case. Let $a, b \in \mathbb{R}$ with $-\infty < a < b < \infty$. By Lemma A.10, there exists $f_n \in C_c(\mathbb{R})$ such that $\lim_{n\to\infty} f_n = 1_{(a,b]}$. Therefore it follows that $x \to 1_{(a,b]}(x_i)$ is $\mathcal{M}_1$ – measurable for each i. Moreover if $-\infty < a_i < b_i < \infty$ for each i, then we may conclude that

$$x \to \prod_{i=1}^d 1_{(a_i,b_i]}(x_i) = 1_{(a_1,b_1]\times\dots\times(a_d,b_d]}(x)$$

is $\mathcal{M}_1$ – measurable as well and hence $(a_1, b_1] \times \dots \times (a_d, b_d] \in \mathcal{M}_1$. As such sets generate $\mathcal{B}_{\mathbb{R}^d}$ we may conclude that $\mathcal{B}_{\mathbb{R}^d} \subset \mathcal{M}_1$.

and therefore $\mathcal{M}_1 = \mathcal{B}_{\mathbb{R}^d}$.

$\mathcal{M}_2$ case. As above, we may find $f_{i,n} \to 1_{(a_i,b_i]}$ as $n \to \infty$ for each $1 \le i \le d$ and therefore,

$$1_{(a_1,b_1]\times\dots\times(a_d,b_d]}(x) = \lim_{n\to\infty} f_{1,n}(x_1) \dots f_{d,n}(x_d) \text{ for all } x \in \mathbb{R}^d.$$

This shows that $1_{(a_1,b_1]\times\dots\times(a_d,b_d]}$ is $\mathcal{M}_2$ – measurable and therefore $(a_1, b_1] \times \dots \times (a_d, b_d] \in \mathcal{M}_2$.

$\mathcal{M}_3$ case. This is easy since $\mathcal{B}_{\mathbb{R}^d} = \mathcal{M}_2 \subset \mathcal{M}_3 \subset \mathcal{B}_{\mathbb{R}^d}$.

$\mathcal{M}_4$ case. Let $\pi_j : \mathbb{R}^d \to \mathbb{R}$ be projection onto the j^{th} – factor, then for $\lambda > 0$, $e_\lambda \circ \pi_j(x) = e^{i\lambda x_j}$. It then follows that

$$\begin{aligned}\sigma(e_\lambda \circ \pi_j : \lambda > 0) &= \sigma\left((e_\lambda \circ \pi_j)^{-1}(W) : \lambda > 0, W \in \mathcal{B}_{\mathbb{C}}\right) \\ &= \sigma\left(\pi_j^{-1}\left(e_\lambda^{-1}(W)\right) : \lambda > 0, W \in \mathcal{B}_{\mathbb{C}}\right) \\ &= \pi_j^{-1}\left(\sigma\left(\left(e_\lambda^{-1}(W)\right) : \lambda > 0, W \in \mathcal{B}_{\mathbb{C}}\right)\right) = \pi_j^{-1}(\mathcal{B}_{\mathbb{R}})\end{aligned}$$

wherein we have used Lemma A.11 for the last equality. Since $\sigma(e_\lambda \circ \pi_j : \lambda > 0) \subset \mathcal{M}_4$ for each j we must have

$$\mathcal{B}_{\mathbb{R}^d} = \overbrace{\mathcal{B}_{\mathbb{R}} \otimes \dots \otimes \mathcal{B}_{\mathbb{R}}}^{d \text{ times}} = \sigma(\pi_j : 1 \le j \le d) \subset \mathcal{M}_4.$$

Alternative proof. By Lemma **??** below there exists $g_n \in$Trig$(\mathbb{R})$ such that $\lim_{n\to\infty} g_n = 1_{(a,b]}$. Since $x \to g_n(x_i)$ is in the span $\left\{x \to e^{i\lambda\cdot x} : \lambda \in \mathbb{R}^d\right\}$ for each n, it follows that $x \to 1_{(a,b]}(x_i)$ is $\mathcal{M}_4$ – measurable for all $-\infty < a < b < \infty$. Therefore, just as in the proof of case 1., we may now conclude that $\mathcal{B}_{\mathbb{R}^d} \subset \mathcal{M}_4$. ■

Corollary A.14. *Suppose that $\mathbb{H}$ is a subspace of complex valued functions on $\mathbb{R}^d$ which is closed under complex conjugation and bounded convergence. If $\mathbb{H}$ contains any one of the following collection of functions;*

1. $\mathbb{M} := \{x \to f_1(x_1)\dots f_d(x_d) : f_i \in C_c(\mathbb{R})\}$
2. $\mathbb{M} := C_c\left(\mathbb{R}^d\right)$, *or*
3. $\mathbb{M} := \left\{x \to e^{i\lambda\cdot x} : \lambda \in \mathbb{R}^d\right\}$

then $\mathbb{H}$ contains all bounded complex Borel measurable functions on $\mathbb{R}^d$.

Proof. Observe that if $f \in C_c(\mathbb{R})$ such that $f(x) = 1$ in a neighborhood of 0, then $f_n(x) := f(x/n) \to 1$ as $n \to \infty$. Therefore in cases 1. and 2., $\mathbb{H}$ contains the constant function, 1, since

$$1 = \lim_{n\to\infty} f_n(x_1)\dots f_n(x_d).$$

In case 3, $1 \in \mathbb{M} \subset \mathbb{H}$ as well. The result now follows from Theorem A.9 and Corollary A.13. ■

Proposition A.15 (Change of Variables Formula). *Suppose that $-\infty < a < b < \infty$ and $u : [a,b] \to \mathbb{R}$ is a continuously differentiable function which is **not necessarily** invertible. Let $[c,d] = u([a,b])$ where $c = \min u([a,b])$ and $d = \max u([a,b])$. (By the intermediate value theorem $u([a,b])$ is an interval.) Then for all bounded measurable functions, $f : [c,d] \to \mathbb{R}$ we have*

$$\int_{u(a)}^{u(b)} f(x)\,dx = \int_a^b f(u(t))\,\dot{u}(t)\,dt. \tag{A.2}$$

Moreover, Eq. (A.2) is also valid if $f : [c,d] \to \mathbb{R}$ is measurable and

$$\int_a^b |f(u(t))|\,|\dot{u}(t)|\,dt < \infty. \tag{A.3}$$

Proof. Let $\mathbb{H}$ denote the space of bounded measurable functions such that Eq. (A.2) holds. It is easily checked that $\mathbb{H}$ is a linear space closed under bounded convergence. Next we show that $\mathbb{M} = C([c,d],\mathbb{R}) \subset \mathbb{H}$ which coupled with Corollary A.14 will show that $\mathbb{H}$ contains all bounded measurable functions from $[c,d]$ to $\mathbb{R}$.

If $f : [c,d] \to \mathbb{R}$ is a continuous function and let F be an anti-derivative of f. Then by the fundamental theorem of calculus,

$$\begin{aligned}\int_a^b f(u(t))\,\dot{u}(t)\,dt &= \int_a^b F'(u(t))\,\dot{u}(t)\,dt \\ &= \int_a^b \frac{d}{dt}F(u(t))\,dt = F(u(t))\,|_a^b \\ &= F(u(b)) - F(u(a)) = \int_{u(a)}^{u(b)} F'(x)\,dx = \int_{u(a)}^{u(b)} f(x)\,dx.\end{aligned}$$

Thus $\mathbb{M} \subset \mathbb{H}$ and the first assertion of the proposition is proved.

Now suppose that $f : [c,d] \to \mathbb{R}$ is measurable and Eq. (A.3) holds. For $M < \infty$, let $f_M(x) = f(x)\cdot 1_{|f(x)|\le M}$ – a bounded measurable function. Therefore applying Eq. (A.2) with f replaced by $|f_M|$ shows,

$$\left|\int_{u(a)}^{u(b)} |f_M(x)|\,dx\right| = \left|\int_a^b |f_M(u(t))|\,\dot{u}(t)\,dt\right| \le \int_a^b |f_M(u(t))|\,|\dot{u}(t)|\,dt.$$

Using the MCT, we may let $M \uparrow \infty$ in the previous inequality to learn

$$\left|\int_{u(a)}^{u(b)} |f(x)|\,dx\right| \le \int_a^b |f(u(t))|\,|\dot{u}(t)|\,dt < \infty.$$

Now apply Eq. (A.2) with f replaced by f_M to learn

$$\int_{u(a)}^{u(b)} f_M(x)\,dx = \int_a^b f_M(u(t))\,\dot{u}(t)\,dt.$$

Using the DCT we may now let $M \to \infty$ in this equation to show that Eq. (A.2) remains valid. ■

Exercise A.2. Suppose that $u : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function such that $\dot{u}(t) \ge 0$ for all t and $\lim_{t\to\pm\infty} u(t) = \pm\infty$. Use the multiplicative system theorem to prove

$$\int_{\mathbb{R}} f(x)\,dx = \int_{\mathbb{R}} f(u(t))\,\dot{u}(t)\,dt \tag{A.4}$$

for all measurable functions $f : \mathbb{R} \to [0,\infty]$. In particular applying this result to $u(t) = at + b$ where $a > 0$ implies,

$$\int_{\mathbb{R}} f(x)\,dx = a\int_{\mathbb{R}} f(at+b)\,dt.$$

Definition A.16. *The* ***Fourier transform*** *or* ***characteristic function*** *of a finite measure, μ, on $\left(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}\right)$, is the function, $\hat{\mu} : \mathbb{R}^d \to \mathbb{C}$ defined by*

$$\hat{\mu}(\lambda) := \int_{\mathbb{R}^d} e^{i\lambda \cdot x} d\mu(x) \text{ for all } \lambda \in \mathbb{R}^d$$

Corollary A.17. *Suppose that μ and ν are two probability measures on $\left(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}\right)$. Then any one of the next three conditions implies that $\mu = \nu$;*

1. $\int_{\mathbb{R}^d} f_1(x_1) \dots f_d(x_d)\, d\nu(x) = \int_{\mathbb{R}^d} f_1(x_1) \dots f_d(x_d)\, d\mu(x)$ *for all* $f_i \in C_c(\mathbb{R})$.
2. $\int_{\mathbb{R}^d} f(x)\, d\nu(x) = \int_{\mathbb{R}^d} f(x)\, d\mu(x)$ *for all* $f \in C_c\left(\mathbb{R}^d\right)$.
3. $\hat{\nu} = \hat{\mu}$.

Item 3. asserts that the Fourier transform is injective.

Proof. Let $\mathbb{H}$ be the collection of bounded complex measurable functions from $\mathbb{R}^d$ to $\mathbb{C}$ such that

$$\int_{\mathbb{R}^d} f d\mu = \int_{\mathbb{R}^d} f d\nu. \tag{A.5}$$

It is easily seen that $\mathbb{H}$ is a linear space closed under complex conjugation and bounded convergence (by the DCT). Since $\mathbb{H}$ contains one of the multiplicative systems appearing in Corollary A.14, it contains all bounded Borel measurable functions form $\mathbb{R}^d \to \mathbb{C}$. Thus we may take $f = 1_A$ with $A \in \mathcal{B}_{\mathbb{R}^d}$ in Eq. (A.5) to learn, $\mu(A) = \nu(A)$ for all $A \in \mathcal{B}_{\mathbb{R}^d}$. ■

A.2 Weak, Weak*, and Strong topologies

Another collection of examples of topological vector spaces comes from putting different (weaker) topologies on familiar Banach spaces.

Definition A.18 (Weak and weak-* topologies). *Let X be a normed vector space and X^* its dual space (all continuous linear functionals on X).*

1. *The* ***weak topology*** *on X is the X^* topology of X, i.e. the smallest topology on X such that every element $f \in X^*$ is continuous. This topology is often denoted by $\sigma(X, X^*)$.*
2. *The* ***weak-$*$ topology*** *on X^* is the topology generated by X, i.e. the smallest topology on X^* such that the maps $f \in X^* \to f(x) \in \mathbb{C}$ are continuous for all $x \in X$. In other words it is the topology $\sigma(X^*, \hat{X})$ where $\hat{X}$ is the image of $X \ni x \to \hat{x} \in X^{**}$. [The weak topology on X^* is the topology generated by X^{**} which is may be finer than the weak-* topology on X^*.]*

Definition A.19 (Operator Topologies). *Let X and Y be be a normed vector spaces and $B(X,Y)$ the normed space of bounded linear transformations from X to Y.*

1. *The* ***strong operator topology (s.o.t.)*** *on $B(X,Y)$ is the smallest topology such that $T \in B(X,Y) \longrightarrow Tx \in Y$ is continuous for all $x \in X$.*
2. *The* ***weak operator topology (w.o.t.)*** *on $B(X,Y)$ is the smallest topology such that $T \in B(X,Y) \longrightarrow f(Tx) \in \mathbb{C}$ is continuous for all $x \in X$ and $f \in Y^*$.*

Remark A.20. Let us be a little more precise about the topologies described in the above definitions.

1. The **weak topology** on X has a neighborhood base at $x_0 \in X$ consisting of sets of the form
$$N = \cap_{i=1}^n \{x \in X : |f_i(x) - f_i(x_0)| < \varepsilon\}$$
where $f_i \in X^*$ and $\varepsilon > 0$.
2. The **weak-$*$ topology** on X^* has a neighborhood base at $f \in X^*$ consisting of sets of the form
$$N := \cap_{i=1}^n \{g \in X^* : |f(x_i) - g(x_i)| < \varepsilon\}$$
where $x_i \in X$ and $\varepsilon > 0$.
3. The **strong operator topology** on $B(X,Y)$ has a neighborhood base at $T \in X^*$ consisting of sets of the form
$$N := \cap_{i=1}^n \{S \in L(X,Y) : \|Sx_i - Tx_i\| < \varepsilon\}$$
where $x_i \in X$ and $\varepsilon > 0$.
4. The **weak operator topology** on $B(X,Y)$ has a neighborhood base at $T \in X^*$ consisting of sets of the form
$$N := \cap_{i=1}^n \{S \in L(X,Y) : |f_i(Sx_i - Tx_i)| < \varepsilon\}$$
where $x_i \in X$, $f_i \in X^*$ and $\varepsilon > 0$.
5. If we let τ_{op} – be the operator-norm topology, τ_s be strong operator topology, and τ_w be the weak operator topology on $B(X,Y)$, .then $\tau_w \subset \tau_s \subset \tau_{op}$. Consequently; if $\Gamma \subset B(X,Y)$ is a set, then $\overline{\Gamma}^{\tau_{op}} \subset \overline{\Gamma}^{\tau_s} \subset \overline{\Gamma}^{\tau_w}$ and in particular; a τ_w-closed set is a τ_s – closed set and a τ_s – closed set is a τ_{op} – closed set.

Lemma A.21. *Let us continue the same notation as in item 5. of Remark A.20. Then $A \in \overline{\Gamma}^{\tau_w}$ iff for every $\Lambda \subset_f X \times Y^*$, there exists $A_n \in \Gamma$ such that $\lim_{n\to\infty} f(A_n x) = f(Ax)$ for all $(f,x) \in \Lambda$ and similarly $A \in \overline{\Gamma}^{\tau_s}$ iff for every $\Lambda \subset_f X$, there exists $A_n \in \Gamma$ such that $\lim_{n\to\infty} A_n x = Ax$ for all $x \in \Lambda$. [Note well, the sequences $\{A_n\} \subset \Gamma$ used here are allowed to depend on Γ!]*

Proof. This follows directly from Proposition **??** and the definitions of the weak and strong operator topologies.

■

A.3 Quotient spaces, adjoints, and reflexivity

Definition A.22. *Let X and Y be Banach spaces and $A : X \to Y$ be a linear operator. The* ***transpose*** *of A is the linear operator $A^\dagger : Y^* \to X^*$ defined by $\left(A^\dagger f\right)(x) = f(Ax)$ for $f \in Y^*$ and $x \in X$. The* ***null space*** *of A is the subspace $\mathrm{Nul}\,(A) := \{x \in X : Ax = 0\} \subset X$. For $M \subset X$ and $N \subset X^*$ let*

$$M^0 := \{f \in X^* : f|_M = 0\} \text{ and}$$
$$N^\perp := \{x \in X : f(x) = 0 \text{ for all } f \in N\}.$$

Proposition A.23 (Basic properties of transposes and annihilators).

1. $\|A\| = \left\|A^\dagger\right\|$ *and* $A^{\dagger\dagger}\hat{x} = \widehat{Ax}$ *for all* $x \in X$.
2. M^0 *and* $N^\perp$ *are always closed subspaces of* X^* *and* X *respectively.*
3. $\left(M^0\right)^\perp = \bar{M}$.
4. $\bar{N} \subset \left(N^\perp\right)^0$ *with equality when X is reflexive. (See Exercise* **??***, Example* **??** *above which shows that* $\bar{N} \neq \left(N^\perp\right)^0$ *in general.)*
5. $\mathrm{Nul}\,(A) = \mathrm{Ran}(A^\dagger)^\perp$ *and* $\mathrm{Nul}(A^\dagger) = \mathrm{Ran}\,(A)^0$. *Moreover,* $\overline{\mathrm{Ran}\,(A)} = \mathrm{Nul}(A^\dagger)^\perp$ *and if X is reflexive, then* $\overline{\mathrm{Ran}(A^\dagger)} = \mathrm{Nul}\,(A)^0$.
6. *X is reflexive iff X^* is reflexive. More generally* $X^{***} = \widehat{X^*} \oplus \hat{X}^0$ *where*
$$\hat{X}^0 = \{\lambda \in X^{***} : \lambda(\hat{x}) = 0 \text{ for all } x \in X\}.$$

Proof.

1.
$$\|A\| = \sup_{\|x\|=1}\|Ax\| = \sup_{\|x\|=1}\sup_{\|f\|=1}|f(Ax)|$$
$$= \sup_{\|f\|=1}\sup_{\|x\|=1}\left|A^\dagger f(x)\right| = \sup_{\|f\|=1}\left\|A^\dagger f\right\| = \left\|A^\dagger\right\|.$$
2. This is an easy consequence of the assumed continuity off all linear functionals involved.
3. If $x \in M$, then $f(x) = 0$ for all $f \in M^0$ so that $x \in \left(M^0\right)^\perp$. Therefore $\bar{M} \subset \left(M^0\right)^\perp$. If $x \notin \bar{M}$, then there exists $f \in X^*$ such that $f|_M = 0$ while $f(x) \neq 0$, i.e. $f \in M^0$ yet $f(x) \neq 0$. This shows $x \notin \left(M^0\right)^\perp$ and we have shown $\left(M^0\right)^\perp \subset \bar{M}$.
4. It is again simple to show $N \subset \left(N^\perp\right)^0$ and therefore $\bar{N} \subset \left(N^\perp\right)^0$. Moreover, as above if $f \notin \bar{N}$ there exists $\psi \in X^{**}$ such that $\psi|_{\bar{N}} = 0$ while $\psi(f) \neq 0$. If X is reflexive, $\psi = \hat{x}$ for some $x \in X$ and since $g(x) = \psi(g) = 0$ for all $g \in \bar{N}$, we have $x \in N^\perp$. On the other hand, $f(x) = \psi(f) \neq 0$ so $f \notin \left(N^\perp\right)^0$. Thus again $\left(N^\perp\right)^0 \subset \bar{N}$.
5.
$$\begin{aligned}\mathrm{Nul}\,(A) &= \{x \in X : Ax = 0\} = \{x \in X : f(Ax) = 0\ \forall\ f \in X^*\}\\ &= \left\{x \in X : A^\dagger f(x) = 0\ \forall\ f \in X^*\right\}\\ &= \left\{x \in X : g(x) = 0\ \forall\ g \in \mathrm{Ran}(A^\dagger)\right\} = \mathrm{Ran}(A^\dagger)^\perp.\end{aligned}$$
Similarly,
$$\begin{aligned}\mathrm{Nul}(A^\dagger) &= \left\{f \in Y^* : A^\dagger f = 0\right\} = \left\{f \in Y^* : (A^\dagger f)(x) = 0\ \forall\ x \in X\right\}\\ &= \{f \in Y^* : f(Ax) = 0\ \forall\ x \in X\}\\ &= \left\{f \in Y^* : f|_{\mathrm{Ran}(A)} = 0\right\} = \mathrm{Ran}\,(A)^0.\end{aligned}$$
6. Let $\psi \in X^{***}$ and define $f_\psi \in X^*$ by $f_\psi(x) = \psi(\hat{x})$ for all $x \in X$ and set $\psi' := \psi - \hat{f}_\psi$. For $x \in X$ (so $\hat{x} \in X^{**}$) we have
$$\psi'(\hat{x}) = \psi(\hat{x}) - \hat{f}_\psi(\hat{x}) = f_\psi(x) - \hat{x}(f_\psi) = f_\psi(x) - f_\psi(x) = 0.$$
This shows $\psi' \in \hat{X}^0$ and we have shown $X^{***} = \widehat{X^*} + \hat{X}^0$. If $\psi \in \widehat{X^*} \cap \hat{X}^0$, then $\psi = \hat{f}$ for some $f \in X^*$ and $0 = \hat{f}(\hat{x}) = \hat{x}(f) = f(x)$ for all $x \in X$, i.e. $f = 0$ so $\psi = 0$. Therefore $X^{***} = \widehat{X^*} \oplus \hat{X}^0$ as claimed. If X is reflexive, then $\hat{X} = X^{**}$ and so $\hat{X}^0 = \{0\}$ showing $X^{***} = \widehat{X^*}$, i.e. X^* is reflexive. Conversely if X^* is reflexive we conclude that $\hat{X}^0 = \{0\}$ and therefore $X^{**} = \{0\}^\perp = \left(\hat{X}^0\right)^\perp = \hat{X}$, so that X is reflexive.
Alternative proof. Notice that $f_\psi = J^\dagger\psi$, where $J : X \to X^{**}$ is given by $Jx = \hat{x}$, and the composition
$$f \in X^* \xrightarrow{\hat{}} \hat{f} \in X^{***} \xrightarrow{J^\dagger} J^\dagger\hat{f} \in X^*$$
is the identity map since $\left(J^\dagger\hat{f}\right)(x) = \hat{f}(Jx) = \hat{f}(\hat{x}) = \hat{x}(f) = f(x)$ for all $x \in X$. Thus it follows that $X^* \xrightarrow{\hat{}} X^{***}$ is invertible iff $J^\dagger$ is its inverse which can happen iff $\mathrm{Nul}(J^\dagger) = \{0\}$. But as above $\mathrm{Nul}(J^\dagger) = \mathrm{Ran}\,(J)^0$ which will be zero iff $\overline{\mathrm{Ran}(J)} = X^{**}$ and since J is an isometry this is equivalent to saying $\mathrm{Ran}\,(J) = X^{**}$. So we have again shown X^* is reflexive iff X is reflexive.

■

Theorem A.24 (Banach Space Factor Theorem). *Let X be a Banach space, $M \subset X$ be a proper closed subspace, X/M the quotient space, $\pi : X \to X/M$ the projection map $\pi(x) = x + M$ for $x \in X$ and define the quotient norm on X/M by*

$$\|\pi(x)\|_{X/M} = \|x + M\|_{X/M} = \inf_{m \in M} \|x + m\|_X .$$

Then:

1. *$\|\cdot\|_{X/M}$ is a norm on X/M.*
2. *The projection map $\pi : X \to X/M$ is has norm 1, $\|\pi\| = 1$.*
3. *For all $a \in X$ and $\varepsilon > 0$, $\pi\left(B^X(a, \varepsilon)\right) = B^{X/M}(\pi(a), \varepsilon)$ and in particular π is an open mapping.*
4. *$(X/M, \|\cdot\|_{X/M})$ is a Banach space.*
5. *If Y is another normed space and $T : X \to Y$ is a bounded linear transformation such that $M \subset \mathrm{Nul}(T)$, then there exists a unique linear transformation $\hat{T} : X/M \to Y$ such that $T = \hat{T} \circ \pi$ and moreover $\|T\| = \left\|\hat{T}\right\|$.*
6. *The map,*

$$\left\{ \begin{array}{c} \text{closed subspaces} \\ \text{of } X \text{ containing } M \end{array} \right\} \ni N \to \pi(N) \in \left\{ \begin{array}{c} \text{closed subspaces} \\ \text{of } \pi(X/M) \end{array} \right\}$$

is a bijection. The inverse map is given by pulling back subspace of $\pi(X/M)$ by π^{-1}. [The word closed may be removed above and the result still holds as one learns in a linear algebra class.]

Proof. We take each item in turn.

1. Clearly $\|x + M\| \geq 0$ and if $\|x + M\| = 0$, then there exists $m_n \in M$ such that $\|x + m_n\| \to 0$ as $n \to \infty$, i.e. $x = -\lim_{n\to\infty} m_n \in \bar{M} = M$. Since $x \in M$, $x + M = 0 \in X/M$. If $c \in \mathbb{C}\setminus\{0\}$, $x \in X$, then

$$\|cx + M\| = \inf_{m \in M} \|cx + m\| = |c| \inf_{m \in M} \|x + m/c\| = |c| \, \|x + M\|$$

because m/c runs through M as m runs through M. Let $x_1, x_2 \in X$ and $m_1, m_2 \in M$ then

$$\|x_1 + x_2 + M\| \leq \|x_1 + x_2 + m_1 + m_2\| \leq \|x_1 + m_1\| + \|x_2 + m_2\|.$$

Taking infimums over $m_1, m_2 \in M$ then implies

$$\|x_1 + x_2 + M\| \leq \|x_1 + M\| + \|x_2 + M\|.$$

and we have completed the proof the $(X/M, \|\cdot\|)$ is a normed space.

2. Since $\|\pi(x)\| = \inf_{m \in M} \|x + m\| \leq \|x\|$ for all $x \in X$, $\|\pi\| \leq 1$. To see $\|\pi\| = 1$, let $x \in X \setminus M$ so that $\pi(x) \neq 0$. Given $\alpha \in (0, 1)$, there exists $m \in M$ such that

$$\|x + m\| \leq \alpha^{-1} \|\pi(x)\| .$$

Therefore,

$$\frac{\|\pi(x + m)\|}{\|x + m\|} = \frac{\|\pi(x)\|}{\|x + m\|} \geq \frac{\alpha \|x + m\|}{\|x + m\|} = \alpha$$

which shows $\|\pi\| \geq \alpha$. Since $\alpha \in (0, 1)$ is arbitrary we conclude that $\|\pi(x)\| = 1$.

3. Since $\|\pi\| < 1$ if $\varepsilon > 0$ then $\pi\left(B^X(0, \varepsilon)\right) \subset B^{X/M}(0, \varepsilon)$. Conversely if $y \in X$ and $\pi(y) \in B^{X/M}(0, \varepsilon)$ then there exists $m \in M$ so that $\|y + m\| < \varepsilon$, i.e. $y + m \in B^X(0, \varepsilon)$. Since $\pi(y) = \pi(y + m)$, this shows that $\pi(y) \in \pi\left(B^X(0, \varepsilon)\right)$ and so $\pi\left(B^X(0, \varepsilon)\right) = B^{X/M}(0, \varepsilon)$ for all $\varepsilon > 0$. For general $a \in X$ and $\varepsilon > 0$ we have

$$\begin{aligned} \pi\left(B^X(a, \varepsilon)\right) &= \pi\left(a + B^X(0, \varepsilon)\right) = \pi(a) + \pi\left(B^X(0, \varepsilon)\right) \\ &= \pi(a) + B^{X/M}(0, \varepsilon) = B^{X/M}(\pi(a), \varepsilon) . \end{aligned}$$

4. Let $\pi(x_n) \in X/M$ be a sequence such that $\sum \|\pi(x_n)\| < \infty$. As above there exists $m_n \in M$ such that $\|\pi(x_n)\| \geq \frac{1}{2}\|x_n + m_n\|$ and hence $\sum \|x_n + m_n\| \leq 2 \sum \|\pi(x_n)\| < \infty$. Since X is complete, $x := \sum_{n=1}^{\infty} (x_n + m_n)$ exists in X and therefore by the continuity of π,

$$\pi(x) = \sum_{n=1}^{\infty} \pi(x_n + m_n) = \sum_{n=1}^{\infty} \pi(x_n)$$

showing X/M is complete.

5. The existence of $\hat{T}$ is guaranteed by the "factor theorem" from linear algebra. Moreover $\left\|\hat{T}\right\| = \|T\|$ because

$$\|T\| = \left\|\hat{T} \circ \pi\right\| \leq \left\|\hat{T}\right\| \|\pi\| = \left\|\hat{T}\right\|$$

and

$$\begin{aligned} \left\|\hat{T}\right\| &= \sup_{x \notin M} \frac{\left\|\hat{T}(\pi(x))\right\|}{\|\pi(x)\|} = \sup_{x \notin M} \frac{\|Tx\|}{\|\pi(x)\|} \\ &\geq \sup_{x \notin M} \frac{\|Tx\|}{\|x\|} = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \|T\| . \end{aligned}$$

6. First we will shows that $\pi(N)$ is closed whenever N is a closed subspace of X containing M. To verify this, let $\{x_n\} \subset N$ be a sequence such that $\{\pi(x_n)\}_{n=1}^{\infty}$ is Cauchy in $\pi(X/M)$. As in the proof of item 3. we may find $m_n \in M$ such that $x = \lim_{n\to\infty}(x_n + m_n)$ exists with $x \in N$ as N is closed. Therefore

$$\pi(x) = \lim_{n\to\infty} \pi(x_n + m_n) = \lim_{n\to\infty} \pi(x_n) \in \pi(N)$$

which shows $\pi(N)$ is closed. Moreover, $x \in \pi^{-1}(\pi(N))$ iff $\pi(x) \in \pi(N)$ which happens iff $x+M \subset x+N$, i.e. iff $x \in N$. This show $\pi^{-1}(\pi(N)) = N$. Finally, if $\tilde{N}$ is a closed subspace of $\pi(X/M)$, then $N := \pi^{-1}\left(\tilde{N}\right)$ is a closed (π is continuous) subspace of X containing M such that $\pi(N) = \tilde{N}$.

■

Theorem A.25. *Let X be a Banach space. Then*

1. *Identifying X with $\hat{X} \subset X^{**}$, the weak $-*$ topology on X^{**} induces the weak topology on X. More explicitly, the map $x \in X \to \hat{x} \in \hat{X}$ is a homeomorphism when X is equipped with its weak topology and $\hat{X}$ with the relative topology coming from the weak-$*$ topology on X^{**}.*
2. *$\hat{X} \subset X^{**}$ is dense in the weak-$*$ topology on X^{**}.*
3. *Letting C and C^{**} be the closed unit balls in X and X^{**} respectively, then $\hat{C} := \{\hat{x} \in C^{**} : x \in C\}$ is dense in C^{**} in the weak $-*$ topology on X^{**}.*
4. *X is reflexive iff C is weakly compact.*

(See Definition A.19 for the topologies being used here.)

Proof.

1. The weak $-*$ topology on X^{**} is generated by
$$\left\{\hat{f} : f \in X^*\right\} = \{\psi \in X^{**} \to \psi(f) : f \in X^*\}.$$
So the induced topology on X is generated by
$$\{x \in X \to \hat{x} \in X^{**} \to \hat{x}(f) = f(x) : f \in X^*\} = X^*$$
and so the induced topology on X is precisely the weak topology.
2. A basic weak - $*$ neighborhood of a point $\lambda \in X^{**}$ is of the form
$$\mathcal{N} := \cap_{k=1}^{n} \{\psi \in X^{**} : |\psi(f_k) - \lambda(f_k)| < \varepsilon\} \tag{A.6}$$
for some $\{f_k\}_{k=1}^{n} \subset X^*$ and $\varepsilon > 0$. be given. We must now find $x \in X$ such that $\hat{x} \in \mathcal{N}$, or equivalently so that
$$|\hat{x}(f_k) - \lambda(f_k)| = |f_k(x) - \lambda(f_k)| < \varepsilon \text{ for } k = 1, 2, \dots, n. \tag{A.7}$$
In fact we will show there exists $x \in X$ such that $\lambda(f_k) = f_k(x)$ for $k = 1, 2, \dots, n$. To prove this stronger assertion we may, by discarding some of the f_k's if necessary, assume that $\{f_k\}_{k=1}^{n}$ is a linearly independent set. Since the $\{f_k\}_{k=1}^{n}$ are linearly independent, the map $x \in X \to (f_1(x), \dots, f_n(x)) \in \mathbb{C}^n$ is surjective (why) and hence there exists $x \in X$ such that
$$(f_1(x), \dots, f_n(x)) = Tx = (\lambda(f_1), \dots, \lambda(f_n)) \tag{A.8}$$
as desired.
3. Let $\lambda \in C^{**} \subset X^{**}$ and $\mathcal{N}$ be the weak - $*$ open neighborhood of λ as in Eq. (A.6). Working as before, given $\varepsilon > 0$, we need to find $x \in C$ such that Eq. (A.7). It will be left to the reader to verify that it suffices again to assume $\{f_k\}_{k=1}^{n}$ is a linearly independent set. (Hint: Suppose that $\{f_1, \dots, f_m\}$ were a maximal linearly dependent subset of $\{f_k\}_{k=1}^{n}$, then each f_k with $k > m$ may be written as a linear combination $\{f_1, \dots, f_m\}$.) As in the proof of item 2., there exists $x \in X$ such that Eq. (A.8) holds. The problem is that x may not be in C. To remedy this, let $N := \cap_{k=1}^{n} \mathrm{Nul}(f_k) = \mathrm{Nul}(T)$, $\pi : X \to X/N \cong \mathbb{C}^n$ be the projection map and $\bar{f}_k \in (X/N)^*$ be chosen so that $f_k = \bar{f}_k \circ \pi$ for $k = 1, 2, \dots, n$. Then we have produced $x \in X$ such that
$$(\lambda(f_1), \dots, \lambda(f_n)) = (f_1(x), \dots, f_n(x)) = (\bar{f}_1(\pi(x)), \dots, \bar{f}_n(\pi(x))).$$
Since $\left\{\bar{f}_1, \dots, \bar{f}_n\right\}$ is a basis for $(X/N)^*$ we find
$$\begin{aligned} \|\pi(x)\| &= \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{\left|\sum_{i=1}^{n} \alpha_i \bar{f}_i(\pi(x))\right|}{\left\|\sum_{i=1}^{n} \alpha_i \bar{f}_i\right\|} = \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{\left|\sum_{i=1}^{n} \alpha_i \lambda(f_i)\right|}{\left\|\sum_{i=1}^{n} \alpha_i f_i\right\|} \\ &= \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{\left|\lambda(\sum_{i=1}^{n} \alpha_i f_i)\right|}{\left\|\sum_{i=1}^{n} \alpha_i f_i\right\|} \\ &\le \|\lambda\| \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{\left\|\sum_{i=1}^{n} \alpha_i f_i\right\|}{\left\|\sum_{i=1}^{n} \alpha_i f_i\right\|} = 1. \end{aligned}$$
Hence we have shown $\|\pi(x)\| \le 1$ and therefore for any $\alpha > 1$ there exists $y = x + n \in X$ such that $\|y\| < \alpha$ and $(\lambda(f_1), \dots, \lambda(f_n)) = (f_1(y), \dots, f_n(y))$. Hence
$$|\lambda(f_i) - f_i(y/\alpha)| \le \left|f_i(y) - \alpha^{-1} f_i(y)\right| \le (1 - \alpha^{-1}) |f_i(y)|$$
which can be arbitrarily small (i.e. less than ε) by choosing α sufficiently close to 1.

4. Let $\hat{C} := \{\hat{x} : x \in C\} \subset C^{**} \subset X^{**}$. If X is reflexive, $\hat{C} = C^{**}$ is weak - $*$ compact and hence by item 1., C is weakly compact in X. Conversely if C is weakly compact, then $\hat{C} \subset C^{**}$ is weak $-*$ compact being the continuous image of a continuous map. Since the weak $-*$ topology on X^{**} is Hausdorff, it follows that $\hat{C}$ is weak $-*$ closed and so by item 3, $C^{**} = \overline{\hat{C}}^{\text{weak}-*} = \hat{C}$. So if $\lambda \in X^{**}$, $\lambda/\|\lambda\| \in C^{**} = \hat{C}$, i.e. there exists $x \in C$ such that $\hat{x} = \lambda/\|\lambda\|$. This shows $\lambda = (\|\lambda\| x)\hat{}$ and therefore $\hat{X} = X^{**}$.

■

A.4 Rayleigh Quotient

Theorem A.26 (Rayleigh quotient). *If H is a Hilbert space and $T \in B(H)$ is a bounded self-adjoint operator, then*

$$M := \sup_{f \neq 0} \frac{|\langle Tf, f\rangle|}{\|f\|^2} = \|T\| \quad \left(= \sup_{f\neq 0} \frac{\|Tf\|}{\|f\|}\right).$$

Moreover, if there exists a non-zero element $f \in H$ such that

$$\frac{|\langle Tf, f\rangle|}{\|f\|^2} = \|T\|,$$

then f is an eigenvector of T with $Tf = \lambda f$ and $\lambda \in \{\pm\|T\|\}$.

Proof. First proof. Applying Eq. (B.5) with $Q(f,g) = \langle Tf, g\rangle$ and Eq. (B.4) with $Q(f,g) = \langle f, g\rangle$ along with the Cauchy-Schwarz inequality implies,

$$\begin{aligned}4\operatorname{Re}\langle Tf, g\rangle &= \langle T(f+g), (f+g)\rangle - \langle T(f-g), (f-g)\rangle \\ &\leq M\left[\|f+g\|^2 + \|f-g\|^2\right] = 2M\left[\|f\|^2 + \|g\|^2\right].\end{aligned}$$

Replacing f by $e^{i\theta} f$ where θ is chosen so that $e^{i\theta}\langle Tf, g\rangle = |\langle Tf, g\rangle|$ then shows

$$4|\langle Tf, g\rangle| \leq 2M\left[\|f\|^2 + \|g\|^2\right]$$

and therefore,

$$\|T\| = \sup_{\|f\|=\|g\|=1} |\langle f, Tg\rangle| \leq M$$

and since it is clear $M \leq \|T\|$ we have shown $M = \|T\|$.

If $f \in H \setminus \{0\}$ and $\|T\| = |\langle Tf, f\rangle| / \|f\|^2$ then, using Schwarz's inequality,

$$\|T\| = \frac{|\langle Tf, f\rangle|}{\|f\|^2} \leq \frac{\|Tf\|}{\|f\|} \leq \|T\|. \tag{A.9}$$

This implies $|\langle Tf, f\rangle| = \|Tf\| \|f\|$ and forces equality in Schwarz's inequality. So by Theorem **??**, Tf and f are linearly dependent, i.e. $Tf = \lambda f$ for some $\lambda \in \mathbb{C}$. Substituting this into (A.9) shows that $|\lambda| = \|T\|$. Since T is self-adjoint,

$$\lambda \|f\|^2 = \langle \lambda f, f\rangle = \langle Tf, f\rangle = \langle f, Tf\rangle = \langle f, \lambda f\rangle = \bar{\lambda}\langle f, f\rangle = \bar{\lambda}\|f\|^2,$$

which implies that $\lambda \in \mathbb{R}$ and therefore, $\lambda \in \{\pm\|T\|\}$.

Second proof. By the spectral theorem for bounded operators of Chapter **??** below, it suffices to prove the theorem in the case where $T = M_g \in B(H)$ where $H = L^2(\Omega, \mu)$, $(\Omega, \mathcal{F}, \mu)$ is a finite measure space, and $g : \Omega \to \mathbb{R}$ is a bounded measurable function. In this case,

$$|\langle Tf, f\rangle| = \left|\int_\Omega g|f|^2 d\mu\right| \leq \|g\|_{L^\infty(\mu)} \int_\Omega |f|^2 d\mu = \|g\|_{L^\infty(\mu)} \|f\|^2_{L^2(\mu)}.$$

If $m < \|g\|_{L^\infty(\mu)} = \|T\|_{op}$ then we can choose $f = 1_A$ and $\varepsilon \in \{\pm 1\}$ so that $\mu(A) > 0$ and $\varepsilon g 1_A \geq m 1_A$. For this f it follows that

$$|\langle Tf, f\rangle| = \int_A \varepsilon g d\mu \geq m \cdot \mu(A) = m\|f\|^2_{L^2(\mu)}.$$

Combining these last two assertions shows

$$m \leq \sup_{\|f\|\neq 0} \frac{|\langle Tf, f\rangle|}{\|f\|^2} \leq \|T\|_{op}$$

which completes this proof as $m < \|T\|_{op}$ was arbitrary.

■

B

Spectral Theorem (Compact Operator Case)

Before giving the general spectral theorem for bounded self-adjoint operators in the next chapter, we pause to consider the special case of "compact" operators. The theory in this setting looks very much like the finite dimensional matrix case.

B.1 Basics of Compact Operators

Definition B.1 (Compact Operator). *Let $A : X \to Y$ be a bounded operator between two Banach spaces. Then A is **compact** if $A\left[B_X(0,1)\right]$ is precompact in Y or equivalently for any $\{x_n\}_{n=1}^\infty \subset X$ such that $\|x_n\| \le 1$ for all n the sequence $y_n := Ax_n \in Y$ has a convergent subsequence.*

Definition B.2. *A bounded operator $A : X \to Y$ is said to have **finite rank** if $\mathrm{Ran}\,(A) \subset Y$ is finite dimensional.*

The following result is a simple consequence of Theorem **??** and Corollary **??**.

Corollary B.3. *If $A : X \to Y$ is a finite rank operator, then A is compact. In particular if either $\dim(X) < \infty$ or $\dim(Y) < \infty$ then any bounded operator $A : X \to Y$ is finite rank and hence compact.*

Theorem B.4. *Let X and Y be Banach spaces and $\mathcal{K} := \mathcal{K}(X, Y)$ denote the compact operators from X to Y. Then $\mathcal{K}(X, Y)$ is a norm-closed subspace of $B\left(X, Y\right)$. In particular, operator norm limits of finite rank operators are compact.*

Proof. Using the sequential definition of compactness it is easily seen that $\mathcal{K}$ is a vector subspace of $B\left(X, Y\right)$. To finish the proof, we must show that $K \in B\left(X, Y\right)$ is compact if there exists $K_n \in \mathcal{K}(X, Y)$ such that $\lim_{n\to\infty} \|K_n - K\|_{op} = 0$.

First Proof. Let $U := B_0\left(1\right)$ be the unit ball in X. Given $\varepsilon > 0$, choose $N = N(\varepsilon)$ such that $\|K_N - K\| \le \varepsilon$. Using the fact that $K_N U$ is precompact, choose a finite subset $\Lambda \subset U$ such that $K_N U \subset \cup_{\sigma\in\Lambda} B_{K_N\sigma}\left(\varepsilon\right)$. Then given $y = Kx \in KU$ we have $K_N x \in B_{K_N\sigma}\left(\varepsilon\right)$ for some $\sigma \in \Lambda$ and for this σ;

$$\begin{aligned}\|y - K_N\sigma\| &= \|Kx - K_N\sigma\| \\ &\le \|Kx - K_N x\| + \|K_N x - K_N\sigma\| < \varepsilon\,\|x\| + \varepsilon < 2\varepsilon.\end{aligned}$$

This shows $KU \subset \cup_{\sigma\in\Lambda} B_{K_N\sigma}\left(2\varepsilon\right)$ and therefore is KU is 2ε – bounded for all $\varepsilon > 0$, i.e. KU is totally bounded and hence precompact.

Second Proof. Suppose $\{x_n\}_{n=1}^\infty$ is a bounded sequence in X. By compactness, there is a subsequence $\left\{x_n^1\right\}_{n=1}^\infty$ of $\{x_n\}_{n=1}^\infty$ such that $\left\{K_1 x_n^1\right\}_{n=1}^\infty$ is convergent in Y. Working inductively, we may construct subsequences

$$\{x_n\}_{n=1}^\infty \supset \left\{x_n^1\right\}_{n=1}^\infty \supset \left\{x_n^2\right\}_{n=1}^\infty \cdots \supset \{x_n^m\}_{n=1}^\infty \supset \dots$$

such that $\{K_m x_n^m\}_{n=1}^\infty$ is convergent in Y for each m. By the usual Cantor's diagonalization procedure, let $\sigma_n := x_n^n$, then $\{\sigma_n\}_{n=1}^\infty$ is a subsequence of $\{x_n\}_{n=1}^\infty$ such that $\{K_m\sigma_n\}_{n=1}^\infty$ is convergent for all m. Since

$$\begin{aligned}\|K\sigma_n - K\sigma_l\| &\le \|(K - K_m)\,\sigma_n\| + \|K_m(\sigma_n - \sigma_l)\| + \|(K_m - K)\,\sigma_l)\| \\ &\le 2\,\|K - K_m\| + \|K_m(\sigma_n - \sigma_l)\|\,,\end{aligned}$$

$$\lim\sup_{n,l\to\infty} \|K\sigma_n - K\sigma_l\| \le 2\,\|K - K_m\| \to 0 \text{ as } m \to \infty,$$

which shows $\{K\sigma_n\}_{n=1}^\infty$ is Cauchy and hence convergent. ■

Example B.5. Let $X = \ell^2 = Y$ and $\lambda_n \in \mathbb{C}$ such that $\lim_{n\to\infty} \lambda_n = 0$, then $A : X \to Y$ defined by $(Ax)(n) = \lambda_n x(n)$ is compact. To verify this claim, for each $m \in \mathbb{N}$ let $(A_m x)(n) = \lambda_n x(n) 1_{n\le m}$. In matrix language,

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots \\ 0 & \lambda_2 & 0 & \cdots \\ 0 & 0 & \lambda_3 & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix} \text{ and } A_m = \begin{pmatrix} \lambda_1 & 0 & \cdots & & & \\ 0 & \lambda_2 & 0 & \cdots & & \\ \vdots & \ddots & \ddots & \ddots & & \\ & & 0 & \lambda_m & 0 & \cdots \\ & & \cdots & 0 & 0 & \ddots \\ & & & & \ddots & \ddots \end{pmatrix}.$$

Then A_m is finite rank and $\|A - A_m\|_{op} = \max_{n>m} |\lambda_n| \to 0$ as $m \to \infty$. The claim now follows from Theorem B.4.

We will see more examples of compact operators below in Section B.4 and Exercise **??** below.

Lemma B.6. *If $X \xrightarrow{A} Y \xrightarrow{B} Z$ are bounded operators between Banach spaces such the either A or B is compact then the composition $BA : X \to Z$ is also compact. In particular if $\dim X = \infty$ and $A \in L(X,Y)$ is an invertible operator such that*[1] $A^{-1} \in L(Y,X)$ *, then A is **not** compact.*

Proof. Let $B_X(0,1)$ be the open unit ball in X. If A is compact and B is bounded, then $BA(B_X(0,1)) \subset B(\overline{AB_X(0,1)})$ which is compact since the image of compact sets under continuous maps are compact. Hence we conclude that $\overline{BA(B_X(0,1))}$ is compact, being the closed subset of the compact set $B(\overline{AB_X(0,1)})$. If A is continuous and B is compact, then $A(B_X(0,1))$ is a bounded set and so by the compactness of B, $BA(B_X(0,1))$ is a precompact subset of Z, i.e. BA is compact.

Alternatively: Suppose that $\{x_n\}_{n=1}^\infty \subset X$ is a bounded sequence. If A is compact, then $y_n := Ax_n$ has a convergent subsequence, $\{y_{n_k}\}_{k=1}^\infty$. Since B is continuous it follows that $z_{n_k} := By_{n_k} = BAx_{n_k}$ is a convergent subsequence of $\{BAx_n\}_{n=1}^\infty$. Similarly if A is bounded and B is compact then $y_n = Ax_n$ defines a bounded sequence inside of Y. By compactness of B, there is a subsequence $\{y_{n_k}\}_{k=1}^\infty$ for which $\{BAx_{n_k} = By_{n_k}\}_{k=1}^\infty$ is convergent in Z.

For the second statement, if A were compact then $I_X := A^{-1}A$ would be compact as well. As I_X takes the unit ball to the unit ball, the identity is compact iff $\dim X < \infty$. ∎

Corollary B.7. *Let X be a Banach space and $\mathcal{K}(X) := \mathcal{K}(X,X)$. Then $\mathcal{K}(X)$ is a norm-closed ideal of $L(X)$ which contains I_X iff $\dim X < \infty$.*

Lemma B.8. *Suppose that $T, T_n \in L(X,Y)$ for $n \in \mathbb{N}$ where X and Y are normed spaces. If $T_n \xrightarrow{s} T$, $M = \sup_n \|T_n\| < \infty$,*[2] *and $x_n \to x$ in X as $n \to \infty$, then $T_n x_n \to Tx$ in Y as $n \to \infty$. Moreover if $K \subset X$ is a compact set then*

$$\lim_{n\to\infty} \sup_{x\in K} \|Tx - T_n x\| = 0. \tag{B.1}$$

Proof. 1. We have,

$$\begin{aligned}\|Tx - T_n x_n\| &\le \|Tx - T_n x\| + \|T_n x - T_n x_n\| \\ &\le \|Tx - T_n x\| + M\|x - x_n\| \to 0 \text{ as } n \to \infty.\end{aligned}$$

2. For sake of contradiction, suppose that

[1] Later we will see that A being one to one and onto automatically implies that A^{-1} is bounded by the open mapping Theorem **??**.

[2] If X and Y are Banach spaces, the uniform boundedness principle shows that $T_n \xrightarrow{s} T$ automatically implies $\sup_n \|T_n\| < \infty$.

$$\limsup_{n\to\infty} \sup_{x\in K} \|Tx - T_n x\| = \varepsilon > 0.$$

In this case we can find $\{n_k\}_{k=1}^\infty \subset \mathbb{N}$ and $x_{n_k} \in K$ such that $\|Tx_{n_k} - T_{n_k}x_{n_k}\| \ge \varepsilon/2$. Since K is compact, by passing to a subsequence if necessary, we may assume $\lim_{k\to\infty} x_{n_k} = x$ exists in K. On the other hand by part **1.** we know that

$$\lim_{k\to\infty} \|Tx_{n_k} - T_{n_k}x_{n_k}\| = \left\| \lim_{k\to\infty} Tx_{n_k} - \lim_{k\to\infty} T_{n_k}x_{n_k} \right\| = \|Tx - Tx\| = 0.$$

2 alternate proof. Given $\varepsilon > 0$, there exists $\{x_1, \dots, x_N\} \subset K$ such that $K \subset \cup_{l=1}^N B_{x_l}(\varepsilon)$. If $x \in K$, choose l such that $x \in B_{x_l}(\varepsilon)$ in which case,

$$\begin{aligned}\|Tx - T_n x\| &\le \|Tx - Tx_l\| + \|Tx_l - T_n x_l\| + \|T_n x_l - T_n x\| \\ &\le \left(\|T\|_{op} + M\right)\varepsilon + \|Tx_l - T_n x_l\|\end{aligned}$$

and therefore it follows that

$$\sup_{x\in K} \|Tx - T_n x\| \le \left(\|T\|_{op} + M\right)\varepsilon + \max_{1\le l\le N} \|Tx_l - T_n x_l\|$$

and therefore,

$$\limsup_{n\to\infty} \sup_{x\in K} \|Tx - T_n x\| \le \left(\|T\|_{op} + M\right)\varepsilon.$$

As $\varepsilon > 0$ was arbitrary we conclude that Eq. (B.1) holds. ∎

B.2 Compact Operators on Hilbert spaces

For the rest of this section, let H and B be Hilbert spaces and $U := \{x \in H : \|x\| < 1\}$ be the **open unit ball** in H.

Proposition B.9. *A bounded operator $K : H \to B$ is compact iff there exists finite rank operators, $K_n : H \to B$, such that $\|K - K_n\| \to 0$ as $n \to \infty$.*

Proof. Suppose that $K : H \to B$. Since $\overline{K(U)}$ is compact it contains a countable dense subset and from this it follows that $\overline{K(H)}$ is a separable subspace of B. Let $\{\varphi_\ell\}$ be an orthonormal basis for $\overline{K(H)} \subset B$ and

$$P_n y = \sum_{\ell=1}^n \langle y, \varphi_\ell \rangle \varphi_\ell$$

be the orthogonal projection of y onto $\operatorname{span}\{\varphi_\ell\}_{\ell=1}^n$. Then $\lim_{n\to\infty}\|P_n y - y\| = 0$ for all $y\in\overline{K(H)}$. Define $K_n := P_n K$ – a finite rank operator on H. It then follows that

$$\begin{aligned}\limsup_{n\to\infty}\|K-K_n\| &= \limsup_{n\to\infty}\sup_{x\in U}\|Kx-K_nx\| \\ &= \limsup_{n\to\infty}\sup_{x\in U}\|(I-P_n)Kx\| \\ &\le \limsup_{n\to\infty}\sup_{y\in\overline{K(U)}}\|(I-P_n)y\| = 0\end{aligned}$$

by Lemma B.8 along with the facts that $\overline{K(U)}$ is compact and $P_n \xrightarrow{s} I$. The converse direction follows from Corollary B.3 and Theorem B.4. ■

Corollary B.10. *If K is compact then so is K^*.*

Proof. First Proof. Let $K_n = P_nK$ be as in the proof of Proposition B.9, then $K_n^* = K^*P_n$ is still finite rank. Furthermore, using Proposition **??**,

$$\|K^* - K_n^*\| = \|K-K_n\| \to 0 \text{ as } n\to\infty$$

showing K^* is a limit of finite rank operators and hence compact.

Second Proof. Let $\{x_n\}_{n=1}^\infty$ be a bounded sequence in B, then

$$\|K^*x_n - K^*x_m\|^2 = \langle x_n - x_m, KK^*(x_n-x_m)\rangle \le 2C\|KK^*(x_n-x_m)\| \tag{B.2}$$

where C is a bound on the norms of the x_n. Since $\{K^*x_n\}_{n=1}^\infty$ is also a bounded sequence, by the compactness of K there is a subsequence $\{x_n'\}$ of the $\{x_n\}$ such that KK^*x_n' is convergent and hence by Eq. (B.2), so is the sequence $\{K^*x_n'\}$. ■

Example B.11. Let $(X,\mathcal{B},\mu)$ be a σ-finite measure spaces whose $\sigma-$algebra is countably generated by sets of finite measure. If $k\in L^2(X\times X,\mu\otimes\mu)$, then $K: L^2(\mu)\to L^2(\mu)$ defined by

$$Kf(x) := \int_X k(x,y)f(y)\,d\mu(y)$$

is a compact operator.

Proof. First observe that

$$|Kf(x)|^2 \le \|f\|^2\int_X |k(x,y)|^2\,d\mu(y)$$

and hence

$$\|Kf\|^2 \le \|f\|^2\int_{X\times X}|k(x,y)|^2\,d\mu(x)\,d\mu(y)$$

from which it follows that $\|K\|_{op}\le\|k\|_{L^2(\mu\otimes\mu)}$.

Now let $\{\psi_n\}_{n=1}^\infty$ be an orthonormal basis for $L^2(X,\mu)$ and let

$$k_N(x,y) := \sum_{m,n=1}^N \langle k,\psi_m\otimes\psi_n\rangle\,\psi_m\otimes\psi_n$$

where $f\otimes g(x,y) := f(x)g(y)$. Then

$$K_Nf(x) := \int_X k_N(x,y)f(y)\,d\mu(y) = \sum_{m,n=1}^N\langle k,\psi_m\otimes\psi_n\rangle\langle f,\bar\psi_n\rangle\psi_m$$

is a finite rank and hence compact operator. Since

$$\|K-K_N\|_{op}\le\|k-k_N\|_{L^2(\mu\otimes\mu)}\to 0 \text{ as } N\to\infty$$

it follows that K is compact as well. ■

B.3 The Spectral Theorem for Self Adjoint Compact Operators

For the rest of this section, $K\in\mathcal{K}(H) := \mathcal{K}(H,H)$ will be a self-adjoint compact operator or **S.A.C.O.** for short. Because of Proposition B.9, we might expect compact operators to behave very much like finite dimensional matrices. This is typically the case as we will see below.

Example B.12 (Model S.A.C.O.). Let $H=\ell_2$ and K be the diagonal matrix

$$K = \begin{pmatrix}\lambda_1 & 0 & 0 & \cdots\\ 0 & \lambda_2 & 0 & \cdots\\ 0 & 0 & \lambda_3 & \cdots\\ \vdots & \vdots & \ddots & \ddots\end{pmatrix},$$

where $\lim_{n\to\infty}|\lambda_n| = 0$ and $\lambda_n\in\mathbb{R}$. Then K is a self-adjoint compact operator. This assertion was proved in Example B.5.

The main theorem (Theorem B.15) of this subsection states that up to unitary equivalence, Example B.12 is essentially the most general example of an S.A.C.O. Before stating and proving this theorem we will require the following results.

Lemma B.13. *Let $Q : H \times H \to \mathbb{C}$ be a symmetric sesquilinear form on H where Q is symmetric means $\overline{Q(h,k)} = Q(k,h)$ for all $h,k \in H$. Letting $Q(h) := Q(h,h)$, then for all $h,k \in H$,*

$$Q(h+k) = Q(h) + Q(k) + 2\operatorname{Re} Q(h,k), \tag{B.3}$$

$$Q(h+k) + Q(h-k) = 2Q(h) + 2Q(k), \text{ and} \tag{B.4}$$

$$Q(h+k) - Q(h-k) = 4\operatorname{Re} Q(h,k). \tag{B.5}$$

Proof. The simple proof is left as an exercise to the reader. ∎

Exercise B.1 (This may be skipped). Suppose that $A : H \to H$ is a bounded self-adjoint operator on H. Show;

1. $f(x) := \langle Ax, x\rangle \in \mathbb{R}$ for all $x \in H$.
2. If there exists $x_0 \in H$ with $\|x_0\| = 1$ such that
$$\lambda_0 := \sup_{\|x\|=1} \langle Ax, x\rangle = \langle Ax_0, x_0\rangle$$
then $Ax_0 = \lambda_0 x_0$. **Hint:** Given $y \in H$ let $c(t) := \frac{x_0+ty}{\|x_0+ty\|_H}$ for t near 0. Then apply the first derivative test to the function $g(t) = \langle Ac(t), c(t)\rangle$.
3. If we further assume that A is compact, then A has at least one eigenvector.

Proposition B.14. *Let K be a S.A.C.O., then either $\lambda = \|K\|$ or $\lambda = -\|K\|$ is an eigenvalue of K.*

Proof. (For those who have done Exercise B.1, that exercise along with Theorem A.26 constitutes a proof.) Without loss of generality we may assume that K is non-zero since otherwise the result is trivial. By Theorem A.26, there exists $u_n \in H$ such that $\|u_n\| = 1$ and

$$\frac{|\langle u_n, Ku_n\rangle|}{\|u_n\|^2} = |\langle u_n, Ku_n\rangle| \longrightarrow \|K\| \text{ as } n \to \infty. \tag{B.6}$$

By passing to a subsequence if necessary, we may assume that $\lambda := \lim_{n\to\infty}\langle u_n, Ku_n\rangle$ exists and $\lambda \in \{\pm\|K\|\}$. By passing to a further subsequence if necessary, we may assume, using the compactness of K, that Ku_n is convergent as well. We now compute:

$$\begin{aligned} 0 \le \|Ku_n - \lambda u_n\|^2 &= \|Ku_n\|^2 - 2\lambda\langle Ku_n, u_n\rangle + \lambda^2 \\ &\le \lambda^2 - 2\lambda\langle Ku_n, u_n\rangle + \lambda^2 \\ &\to \lambda^2 - 2\lambda^2 + \lambda^2 = 0 \text{ as } n \to \infty. \end{aligned}$$

Hence

$$Ku_n - \lambda u_n \to 0 \text{ as } n \to \infty \tag{B.7}$$

and therefore

$$u := \lim_{n\to\infty} u_n = \frac{1}{\lambda}\lim_{n\to\infty} Ku_n$$

exists. By the continuity of the inner product, $\|u\| = 1 \neq 0$. By passing to the limit in Eq. (B.7) we find that $Ku = \lambda u$. ∎

Theorem B.15 (Compact Operator Spectral Theorem). *Suppose that $K : H \to H$ is a non-zero S.A.C.O., then*

1. *there exists at least one eigenvalue $\lambda \in \{\pm\|K\|\}$.*
2. *There are at most countably many* ***non-zero*** *eigenvalues, $\{\lambda_n\}_{n=1}^N$, where $N = \infty$ is allowed. (Unless K is finite rank (i.e. $\dim \operatorname{Ran}(K) < \infty$), N will be infinite.)*
3. *The λ_n's (including multiplicities) may be arranged so that $|\lambda_n| \ge |\lambda_{n+1}|$ for all n. If $N = \infty$ then $\lim_{n\to\infty}|\lambda_n| = 0$. (In particular any eigenspace for K with* ***non-zero*** *eigenvalue is finite dimensional.)*
4. *The eigenvectors $\{\varphi_n\}_{n=1}^N$ can be chosen to be an O.N. set such that $H = \overline{span\{\varphi_n\}} \overset{\perp}{\oplus} \operatorname{Nul}(K)$.*
5. *Using the $\{\varphi_n\}_{n=1}^N$ above,*
$$Kf = \sum_{n=1}^N \lambda_n\langle f, \varphi_n\rangle\varphi_n \text{ for all } f \in H. \tag{B.8}$$
6. *The spectrum of K is $\sigma(K) = \{0\} \cup \{\lambda_n : n < N+1\}$ if $\dim H = \infty$, otherwise $\sigma(K) = \{\lambda_n : n \le N\}$ with $N \le \dim H$.*

Proof. We will find λ_n's and φ_n's recursively. Let $\lambda_1 \in \{\pm\|K\|\}$ and $\varphi_1 \in H$ such that $K\varphi_1 = \lambda_1\varphi_1$ as in Proposition B.14.

Take $M_1 = \operatorname{span}(\varphi_1)$ so $K(M_1) \subset M_1$. By Lemma 3.25, $KM_1^\perp \subset M_1^\perp$. Define $K_1 : M_1^\perp \to M_1^\perp$ via $K_1 = K|_{M_1^\perp}$. Then K_1 is again a compact operator. If $K_1 = 0$, we are done. If $K_1 \neq 0$, by Proposition B.14 there exists $\lambda_2 \in \{\pm\|K_1\|\}$ and $\varphi_2 \in M_1^\perp$ such that $\|\varphi_2\| = 1$ and $K_1\varphi_2 = K\varphi_2 = \lambda_2\varphi_2$. Let $M_2 := \operatorname{span}(\varphi_1, \varphi_2)$.

Again $K(M_2) \subset M_2$ and hence $K_2 := K|_{M_2^\perp} : M_2^\perp \to M_2^\perp$ is compact and if $K_2 = 0$ we are done. When $K_2 \neq 0$, we apply Proposition B.14 again to find $\lambda_3 \in \{\pm\|K\|_2\}$ and $\varphi_3 \in M_2^\perp$ such that $\|\varphi_3\| = 1$ and $K_2\varphi_3 = K\varphi_3 = \lambda_3\varphi_3$.

Continuing this way indefinitely or until we reach a point where $K_n = 0$, we construct a sequence $\{\lambda_n\}_{n=1}^N$ of eigenvalues and orthonormal eigenvectors $\{\varphi_n\}_{n=1}^N$ such that $|\lambda_n| \ge |\lambda_{n+1}|$ with the further property that

$$|\lambda_n| = \sup_{\varphi\perp\{\varphi_1,\varphi_2,\dots\varphi_{n-1}\}} \frac{\|K\varphi\|}{\|\varphi\|}. \tag{B.9}$$

When $N<\infty$, the remaining results in the theorem are easily verified. So from now on let us assume that $N=\infty$.

If $\varepsilon := \lim_{n\to\infty}|\lambda_n| > 0$, then $\left\{\lambda_n^{-1}\varphi_n\right\}_{n=1}^\infty$ is a bounded sequence in H. Hence, by the compactness of K, there exists a subsequence $\{n_k : k\in\mathbb{N}\}$ of $\mathbb{N}$ such that $\left\{\varphi_{n_k} = \lambda_{n_k}^{-1}K\varphi_{n_k}\right\}_{k=1}^\infty$ is a convergent. However, since $\{\varphi_{n_k}\}_{k=1}^\infty$ is an orthonormal set, this is impossible and hence we must conclude that $\varepsilon := \lim_{n\to\infty}|\lambda_n| = 0$.

Let $M := \operatorname{span}\{\varphi_n\}_{n=1}^\infty$. Then $K(M)\subset M$ and hence, by Lemma 3.25, $K(M^\perp)\subset M^\perp$. Using Eq. (B.9),

$$\|K|_{M^\perp}\| \le \left\|K|_{M_n^\perp}\right\| = |\lambda_n| \longrightarrow 0 \text{ as } n\to\infty$$

showing $K|M^\perp\equiv 0$. Define P_0 to be orthogonal projection onto $M^\perp$. Then for $f\in H$,

$$f = P_0f + (1-P_0)f = P_0f + \sum_{n=1}^\infty \langle f,\varphi_n\rangle\varphi_n$$

and

$$Kf = KP_0f + K\sum_{n=1}^\infty\langle f,\varphi_n\rangle\varphi_n = \sum_{n=1}^\infty \lambda_n\langle f,\varphi_n\rangle\varphi_n$$

which proves Eq. (B.8).

Since $\{\lambda_n\}_{n=1}^\infty\subset\sigma(K)$ and $\sigma(K)$ is closed, it follows that $0\in\sigma(K)$ and hence $\{\lambda_n\}_{n=1}^\infty\cup\{0\}\subset\sigma(K)$. Suppose that $z\notin\{\lambda_n\}_{n=1}^\infty\cup\{0\}$ and let d be the distance between z and $\{\lambda_n\}_{n=1}^\infty\cup\{0\}$. Notice that $d>0$ because $\lim_{n\to\infty}\lambda_n = 0$.

A few simple computations show that:

$$(K-zI)f = \sum_{n=1}^\infty\langle f,\varphi_n\rangle(\lambda_n - z)\varphi_n - zP_0f,$$

$(K-z)^{-1}$ exists,

$$(K-zI)^{-1}f = \sum_{n=1}^\infty\langle f,\varphi_n\rangle(\lambda_n-z)^{-1}\varphi_n - z^{-1}P_0f,$$

and

$$\begin{aligned}\|(K-zI)^{-1}f\|^2 &= \sum_{n=1}^\infty|\langle f,\varphi_n\rangle|^2\frac{1}{|\lambda_n-z|^2} + \frac{1}{|z|^2}\|P_0f\|^2\\ &\le\left(\frac{1}{d}\right)^2\left(\sum_{n=1}^\infty|\langle f,\varphi_n\rangle|^2 + \|P_0f\|^2\right) = \frac{1}{d^2}\|f\|^2.\end{aligned}$$

We have thus shown that $(K-zI)^{-1}$ exists, $\|(K-zI)^{-1}\|\le d^{-1}<\infty$ and hence $z\notin\sigma(K)$. ∎

Theorem B.16 (Structure of Compact Operators). *Let $K : H\to B$ be a compact operator. Then there exists $N\in\mathbb{N}\cup\{\infty\}$, orthonormal subsets $\{\varphi_n\}_{n=1}^N\subset H$ and $\{\psi_n\}_{n=1}^N\subset B$ and a sequence $\{\alpha_n\}_{n=1}^N\subset\mathbb{R}_+$ such that $\alpha_1\ge\alpha_2\ge\dots$ (with $\lim_{n\to\infty}\alpha_n = 0$ if $N=\infty$), $\|\psi_n\|\le 1$ for all n and*

$$Kf = \sum_{n=1}^N\alpha_n\langle f,\varphi_n\rangle\psi_n \text{ for all } f\in H. \tag{B.10}$$

Proof. Since K^*K is a self-adjoint compact operator, Theorem B.15 implies there exists an orthonormal set $\{\varphi_n\}_{n=1}^N\subset H$ and positive numbers $\{\lambda_n\}_{n=1}^N$ such that

$$K^*K\psi = \sum_{n=1}^N\lambda_n\langle\psi,\varphi_n\rangle\varphi_n \text{ for all } \psi\in H.$$

Let A be the positive square root of K^*K defined by

$$A\psi := \sum_{n=1}^N\sqrt{\lambda_n}\langle\psi,\varphi_n\rangle\varphi_n \text{ for all } \psi\in H.$$

A simple computation shows, $A^2 = K^*K$, and therefore,

$$\begin{aligned}\|A\psi\|^2 &= \langle A\psi,A\psi\rangle = \left\langle\psi,A^2\psi\right\rangle\\ &= \langle\psi,K^*K\psi\rangle = \langle K\psi,K\psi\rangle = \|K\psi\|^2\end{aligned}$$

for all $\psi\in H$. Hence we may define a unitary operator, $u:\overline{\operatorname{Ran}(A)}\to\overline{\operatorname{Ran}(K)}$ by the formula

$$uA\psi = K\psi \text{ for all } \psi\in H.$$

We then have

$$K\psi = uA\psi = \sum_{n=1}^N\sqrt{\lambda_n}\langle\psi,\varphi_n\rangle u\varphi_n \tag{B.11}$$

which proves the result with $\psi_n := u\varphi_n$ and $\alpha_n = \sqrt{\lambda_n}$.

It is instructive to find ψ_n explicitly and to verify Eq. (B.11) by brute force. Since $\varphi_n = \lambda_n^{-1/2}A\varphi_n$,

$$\psi_n = \lambda_n^{-1/2}uA\varphi_n = \lambda_n^{-1/2}K\varphi_n$$

and

$$\langle K\varphi_n,K\varphi_m\rangle = \langle\varphi_n,K^*K\varphi_m\rangle = \lambda_n\delta_{mn}.$$

This verifies that $\{\psi_n\}_{n=1}^N$ is an orthonormal set. Moreover,

$$\sum_{n=1}^{N}\sqrt{\lambda_n}\langle\psi,\varphi_n\rangle\psi_n=\sum_{n=1}^{N}\sqrt{\lambda_n}\langle\psi,\varphi_n\rangle\lambda_n^{-1/2}K\varphi_n$$
$$=K\sum_{n=1}^{N}\langle\psi,\varphi_n\rangle\varphi_n=K\psi$$

since $\sum_{n=1}^{N}\langle\psi,\varphi_n\rangle\varphi_n=P\psi$ where P is orthogonal projection onto $\mathrm{Nul}(K)^\perp$.

Second Proof. Let $K=u|K|$ be the polar decomposition of K. Then $|K|$ is self-adjoint and compact, by Corollary **??** below, and hence by Theorem B.15 there exists an orthonormal basis $\{\varphi_n\}_{n=1}^{N}$ for $\mathrm{Nul}(|K|)^\perp=\mathrm{Nul}(K)^\perp$ such that $|K|\varphi_n=\lambda_n\varphi_n$, $\lambda_1\geq\lambda_2\geq\dots$ and $\lim_{n\to\infty}\lambda_n=0$ if $N=\infty$. For $f\in H$,

$$Kf=u|K|\sum_{n=1}^{N}\langle f,\varphi_n\rangle\varphi_n=\sum_{n=1}^{N}\langle f,\varphi_n\rangle u|K|\varphi_n=\sum_{n=1}^{N}\lambda_n\langle f,\varphi_n\rangle u\varphi_n$$

which is Eq. (B.10) with $\psi_n:=u\varphi_n$. ■

Exercise B.2 (Continuation of Example ??). Let $H:=L^2([0,1],m)$, $k(x,y):=\min(x,y)$ for $x,y\in[0,1]$ and define $K:H\to H$ by

$$Kf(x)=\int_0^1 k(x,y)f(y)\,dy.$$

From Example B.11 we know that K is a compact operator[3] on H. Since k is real and symmetric, it is easily seen that K is self-adjoint. Show:

1. If $g\in C^2([0,1])$ with $g(0)=0=g'(1)$, then $Kg''=-g$. Use this to conclude $\langle Kf|g''\rangle=-\langle f|g\rangle$ for all $g\in C_c^\infty((0,1))$ and consequently that $\mathrm{Nul}(K)=\{0\}$.
2. Now suppose that $f\in H$ is an eigenvector of K with eigenvalue $\lambda\neq0$. Show that there is a version[4] of f which is in $C([0,1])\cap C^2((0,1))$ and this version, still denoted by f, solves
$$\lambda f''=-f \text{ with } f(0)=f'(1)=0. \tag{B.12}$$
where $f'(1):=\lim_{x\uparrow1}f'(x)$.
3. Use Eq. (B.12) to find all the eigenvalues and eigenfunctions of K.
4. Use the results above along with the spectral Theorem B.15, to show
$$\left\{\sqrt{2}\sin\left(\left(n+\frac{1}{2}\right)\pi x\right):n\in\mathbb{N}_0\right\}$$
is an orthonormal basis for $L^2([0,1],m)$ with $\lambda_n=\left[\left(n+\frac{1}{2}\right)\pi\right]^{-2}$.

[3] See Exercise B.3 from which it will follow that K is a Hilbert Schmidt operator and hence compact.

[4] A measurable function g is called a version of f iff $g=f$ a.e..

5. Repeat this problem in the case that $k(x,y)=\min(x,y)-xy$. In this case you should find that Eq. (B.12) is replaced by
$$\lambda f''=-f \text{ with } f(0)=f(1)=0$$
from which one finds;
$$\left\{f_n:=\sqrt{2}\sin(n\pi x):n\in\mathbb{N}\right\}$$
is an orthonormal basis of eigenvectors of K with corresponding eigenvalues; $\lambda_n=(n\pi)^{-2}$.
6. Use the result of the last part to show,
$$\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}.$$
Hint: First show
$$k(x,y)=\sum_{n=1}^{\infty}\lambda_n f_n(x)f_n(y) \text{ for a.e. } (x,y).$$
Then argue the above equation holds for **every** $(x,y)\in[0,1]^2$. Finally take $y=x$ in the above equation and integrate to arrive at the desired result.

Note: for a wide reaching generalization of this exercise the reader should consult Conway [7, Section II.6 (p.49-54)].

Worked Solution to Exercise (B.2). Let $I=[0,1]$ below.

1. Suppose that $g\in C^2([0,1])$ with $g(0)=0=g'(1)$, then
$$Kg''(x)=\int_0^1 x\wedge y g''(y)\,dy=\int_0^x yg''(y)\,dy+x\int_x^1 g''(y)\,dy$$
$$=-\int_0^x g'(y)\,dy+yg'(y)|_0^x+x(g'(1)-g'(x))$$
$$=-g(x)+g(0)=-g(x).$$
Thus if $g\in C_c^2((0,1))$ we have
$$\langle Kf|g''\rangle=\langle f|Kg''\rangle=-\langle f|g\rangle.$$
In particular if $Kf=0$, this implies that $\int_I f(x)\bar{g}(x)\,dx=0$ for all $g\in C_c^2((0,1))$. Since $C_c^\infty((0,1))$ is dense in $L^2([0,1],m)$ we may choose $g_n\in C_c^2((0,1))$ such that $g_n\to f$ in L^2 as $n\infty$ and therefore
$$0=\lim_{n\to\infty}\int_I f(x)\bar{g}_n(x)\,dx=\int_I|f|^2\,dm.$$
This shows that $f=0$ a.e.

2. If, for a.e. x,
$$\lambda f(x) = Kf(x) = \int_I x \wedge y f(y)\,dy =: F(x)$$
then F is continuous and $F(0) = 0$. Hence $\lambda^{-1}F$ is a continuous version of f. We now re-define f to be $\lambda^{-1}F$. Since
$$f(x) = \lambda^{-1}\int_I x \wedge y f(y)\,dy = \lambda^{-1}\left(\int_0^x y f(y)\,dy + x\int_x^1 f(y)\,dy\right)$$
it follows that $f \in C^1([0,1])$ and
$$f'(x) = \lambda^{-1}\left(xf(x) - xf(x) + \int_x^1 f(y)\,dy\right) = \lambda^{-1}\int_x^1 f(y)\,dy.$$
From this it follows that $f \in C([0,1]) \cap C^2((0,1))$ and that $f'' = -\lambda^{-1} f$ and $f'(1) = 0$.
3. By writing out all of the solutions to Eq. (B.12) we find the only possibilities are
$$f_n(x) = \sin\left(\left(n + \frac{1}{2}\right)\pi x\right) \text{ for } n \in \mathbb{N}$$
with corresponding eigenvalues being $\lambda_n = \left[\left(n + \frac{1}{2}\right)\pi\right]^{-2}$. Notice that if $f'' = -\lambda^{-1}f$ and f satisfies the required boundary conditions, then it follows from the computations in part 1. that
$$-f = Kf'' = K\left(-\lambda^{-1}f\right) = -\lambda^{-1}Kf$$
and therefore,
$$Kf = \lambda f.$$
4. By the spectral Theorem B.15, we must have that $\left\{\frac{f_n}{\|f_n\|_2} : n \in \mathbb{N}\right\}$ is an orthonormal basis for L^2. Since
$$\|f_n\|_2^2 = \int_0^1 \sin^2\left(\left(n + \frac{1}{2}\right)\pi x\right)dx = \int_0^1\left(\frac{1}{2} - \frac{1}{2}\cos\left[(2n+1)\pi x\right]\right)dx = \frac{1}{2}$$
we find $\left\{\sqrt{2}\sin\left(\left(n + \frac{1}{2}\right)\pi x\right) : n \in \mathbb{N}\right\}$ is an orthonormal basis of eigenvectors for H.
Shorter solution. For $f \in L^2(m)$, let
$$F(x) := Kf(x) = \int_I x \wedge y f(y)\,dy = \int_0^x y f(y)\,dy + x\int_x^1 f(y)\,dy.$$
Observe that F is continuous and in fact absolutely continuous, $F(0) = 0$ and
$$F'(x) = xf(x) + \int_x^1 f(y)\,dy - xf(x) = \int_x^1 f(y)\,dy \text{ a.e. } x.$$
If $F := Kf = 0$ then $F' = 0$ a.e. and therefore $\int_x^1 f(y)\,dy = 0$ for all x. Differentiating this equation shows $0 = -f(x)$ a.e. and hence $f = 0$ and therefore $\mathrm{Nul}(K) = 0$.
If $F = Kf = \lambda f$ for some $\lambda \neq 0$ then we learn f has an absolutely continuous version and from the previous equations we find
$$f(0) = 0,\ f'(1) = 0, \text{ and } \lambda f'' = F'' = -f.$$
Thus the eigenfunctions of this equation must be of the form $f(x) = c\sin(kx)$ with k chosen so that $0 = f'(1) = ck\cos(k)$, i.e. $k = \left(n + \frac{1}{2}\right)\pi$.
5. **Modification for Dirichlet Boundary Conditions.** If $k(x,y) = x \wedge y - xy$ instead, then we have
$$\begin{aligned} F(x) &= Kf(x) = \int_0^x y f(y)\,dy + x\int_x^1 f(y)\,dy - x\int_0^1 y f(y)\,dy, \\ F'(x) &= \int_x^1 f(y)\,dy - \int_0^1 y f(y)\,dy, \text{ and} \\ F''(x) &= -f(y). \end{aligned}$$
Thus again $\mathrm{Nul}(K) = \{0\}$ and everything goes through as before except that now $F(0) = 0$ and $F(1) = 0$. Thus the eigenfunctions are of the form $f(x) = c\sin kx$ with k chosen so that $0 = f(1) = c\sin k$. Thus we must have $k = n\pi$ now so that $f_n(x) = c_n \sin n\pi x$. As $\lambda_n f_n'' = -f_n$ we learn that $\lambda_n (n\pi)^2 = -1$ so that
$$\lambda_n = \frac{1}{(n\pi)^2}$$
in this case.
6. We know that $\{f_m \otimes f_n\}_{m,n=1}^\infty$ is an orthonormal basis for $L^2\left(I^2, m \otimes m\right)$. Since
$$\begin{aligned} \langle k, f_m \otimes f_n\rangle &= \int_{I^2} k(x,y) f_m(x) f_n(y)\,dx dy \\ &= \langle Kf_n, f_m\rangle = \lambda_n \langle f_n, f_m\rangle = \lambda_n \delta_{m,n}, \end{aligned}$$
we find
$$k = \sum_{m,n=1}^\infty \lambda_n \delta_{m,n} f_m \otimes f_n = \sum_{n=1}^\infty \lambda_n f_n \otimes f_n \ \ m \otimes m - \text{a.e.}$$
As both sides of the previous equation are continuous, we may conclude that

$$k(x,y)=\sum_{n=1}^{\infty}\lambda_n f_n(x) f_n(y) \text{ for } \mathbf{every}\ x,y\in I.$$

Thus it follows that

$$k(x,x)=\sum_{n=1}^{\infty}\lambda_n f_n^2(x)$$

and then integrating this equation shows

$$\sum_{n=1}^{\infty}\lambda_n=\int_I k(x,x)\,dx=\int_0^1\left(x-x^2\right)dx=\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$$

and hence it follows that

$$\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}.$$

B.4 Hilbert Schmidt Operators

In this section H and B will be Hilbert spaces.

Proposition B.17. *Let H and B be a separable Hilbert spaces, $K:H\to B$ be a bounded linear operator, $\{e_n\}_{n=1}^{\infty}$ and $\{u_m\}_{m=1}^{\infty}$ be orthonormal basis for H and B respectively. Then:*

1. $\sum_{n=1}^{\infty}\|Ke_n\|^2=\sum_{m=1}^{\infty}\|K^*u_m\|^2$ *allowing for the possibility that the sums are infinite. In particular the* ***Hilbert Schmidt norm*** *of K,*
$$\|K\|_{HS}^2:=\sum_{n=1}^{\infty}\|Ke_n\|^2,$$
is well defined independent of the choice of orthonormal basis $\{e_n\}_{n=1}^{\infty}$. We say $K:H\to B$ is a ***Hilbert Schmidt operator*** *if $\|K\|_{HS}<\infty$ and let $HS(H,B)$ denote the space of Hilbert Schmidt operators from H to B.*
2. *For all $K\in L(H,B)$, $\|K\|_{HS}=\|K^*\|_{HS}$ and*
$$\|K\|_{HS}\ge\|K\|_{op}:=\sup\{\|Kh\|:h\in H \text{ such that } \|h\|=1\}.$$
3. *The set $HS(H,B)$ is a subspace of $L(H,B)$ (the bounded operators from $H\to B$), $\|\cdot\|_{HS}$ is a norm on $HS(H,B)$ for which $(HS(H,B),\|\cdot\|_{HS})$ is a Hilbert space, and the corresponding inner product is given by*
$$\langle K_1|K_2\rangle_{HS}=\sum_{n=1}^{\infty}\langle K_1e_n|K_2e_n\rangle. \tag{B.13}$$
4. *If $K:H\to B$ is a bounded finite rank operator, then K is Hilbert Schmidt.*
5. *Let $P_Nx:=\sum_{n=1}^{N}\langle x|e_n\rangle e_n$ be orthogonal projection onto $\mathrm{span}\{e_n:n\le N\}\subset H$ and for $K\in HS(H,B)$, let $K_N:=KP_N$. Then*
$$\|K-K_N\|_{op}^2\le\|K-K_N\|_{HS}^2\to 0 \text{ as } N\to\infty,$$
which shows that finite rank operators are dense in $(HS(H,B),\|\cdot\|_{HS})$. In particular of $HS(H,B)\subset\mathcal{K}(H,B)$ – the space of compact operators from $H\to B$.
6. *If Y is another Hilbert space and $A:Y\to H$ and $C:B\to Y$ are bounded operators, then*
$$\|KA\|_{HS}\le\|K\|_{HS}\|A\|_{op} \text{ and } \|CK\|_{HS}\le\|K\|_{HS}\|C\|_{op},$$
in particular $HS(H,H)$ is an ideal in $L(H)$.

Proof. Items 1. and 2. By Parseval's equality and Fubini's theorem for sums,

$$\begin{aligned}\sum_{n=1}^{\infty}\|Ke_n\|^2&=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}|\langle Ke_n|u_m\rangle|^2\\&=\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}|\langle e_n|K^*u_m\rangle|^2=\sum_{m=1}^{\infty}\|K^*u_m\|^2.\end{aligned}$$

This proves $\|K\|_{HS}$ is well defined independent of basis and that $\|K\|_{HS}=\|K^*\|_{HS}$. For $x\in H\setminus\{0\}$, $x/\|x\|$ may be taken to be the first element in an orthonormal basis for H and hence

$$\left\|K\frac{x}{\|x\|}\right\|\le\|K\|_{HS}.$$

Multiplying this inequality by $\|x\|$ shows $\|Kx\|\le\|K\|_{HS}\|x\|$ and hence $\|K\|_{op}\le\|K\|_{HS}$.

Item 3. For $K_1,K_2\in L(H,B)$,

$$\begin{aligned}\|K_1+K_2\|_{HS}&=\sqrt{\sum_{n=1}^{\infty}\|K_1e_n+K_2e_n\|^2}\\&\le\sqrt{\sum_{n=1}^{\infty}[\|K_1e_n\|+\|K_2e_n\|]^2}\\&=\|\{\|K_1e_n\|+\|K_2e_n\|\}_{n=1}^{\infty}\|_{\ell_2}\\&\le\|\{\|K_1e_n\|\}_{n=1}^{\infty}\|_{\ell_2}+\|\{\|K_2e_n\|\}_{n=1}^{\infty}\|_{\ell_2}\\&=\|K_1\|_{HS}+\|K_2\|_{HS}.\end{aligned}$$

From this triangle inequality and the homogeneity properties of $\|\cdot\|_{HS}$, we now easily see that $HS(H,B)$ is a subspace of $L(H,B)$ and $\|\cdot\|_{HS}$ is a norm on $HS(H,B)$. Since

$$\begin{aligned}\sum_{n=1}^{\infty}|\langle K_1e_n|K_2e_n\rangle| &\le \sum_{n=1}^{\infty}\|K_1e_n\|\,\|K_2e_n\| \\ &\le \sqrt{\sum_{n=1}^{\infty}\|K_1e_n\|^2}\sqrt{\sum_{n=1}^{\infty}\|K_2e_n\|^2} = \|K_1\|_{HS}\,\|K_2\|_{HS},\end{aligned}$$

the sum in Eq. (B.13) is well defined and is easily checked to define an inner product on $HS(H,B)$ such that $\|K\|_{HS}^2 = \langle K|K\rangle_{HS}$.

The proof that $\left(HS(H,B), \|\cdot\|_{HS}^2\right)$ is complete is very similar to the proof of Theorem **??**. Indeed, suppose $\{K_m\}_{m=1}^{\infty}$ is a $\|\cdot\|_{HS}$ – Cauchy sequence in $HS(H,B)$. Because $L(H,B)$ is complete, there exists $K\in L(H,B)$ such that $\|K-K_m\|_{op}\to 0$ as $m\to\infty$. Thus, making use of Fatou's Lemma **??**,

$$\begin{aligned}\|K-K_m\|_{HS}^2 &= \sum_{n=1}^{\infty}\|(K-K_m)\,e_n\|^2 \\ &= \sum_{n=1}^{\infty}\liminf_{l\to\infty}\|(K_l-K_m)\,e_n\|^2 \\ &\le \liminf_{l\to\infty}\sum_{n=1}^{\infty}\|(K_l-K_m)\,e_n\|^2 \\ &= \liminf_{l\to\infty}\|K_l-K_m\|_{HS}^2 \to 0 \text{ as } m\to\infty.\end{aligned}$$

Hence $K\in HS(H,B)$ and $\lim_{m\to\infty}\|K-K_m\|_{HS}^2 = 0$.

Item 4. Since $\mathrm{Nul}(K^*)^{\perp} = \overline{\mathrm{Ran}\,(K)} = \mathrm{Ran}\,(K)$,

$$\|K\|_{HS}^2 = \|K^*\|_{HS}^2 = \sum_{n=1}^{N}\|K^*v_n\|_H^2 < \infty$$

where $N := \dim\mathrm{Ran}\,(K)$ and $\{v_n\}_{n=1}^{N}$ is an orthonormal basis for $\mathrm{Ran}\,(K) = K\,(H)$.

Item 5. Simply observe,

$$\|K-K_N\|_{op}^2 \le \|K-K_N\|_{HS}^2 = \sum_{n>N}\|Ke_n\|^2 \to 0 \text{ as } N\to\infty.$$

Item 6. For $C\in L(B,Y)$ and $K\in L(H,B)$ then

$$\|CK\|_{HS}^2 = \sum_{n=1}^{\infty}\|CKe_n\|^2 \le \|C\|_{op}^2\sum_{n=1}^{\infty}\|Ke_n\|^2 = \|C\|_{op}^2\,\|K\|_{HS}^2$$

and for $A\in L\,(Y,H)$,

$$\|KA\|_{HS} = \|A^*K^*\|_{HS} \le \|A^*\|_{op}\,\|K^*\|_{HS} = \|A\|_{op}\,\|K\|_{HS}.$$

■

Remark B.18. The separability assumptions made in Proposition B.17 are unnecessary. In general, we define

$$\|K\|_{HS}^2 = \sum_{e\in\beta}\|Ke\|^2$$

where $\beta\subset H$ is an orthonormal basis. The same proof of Item 1. of Proposition B.17 shows $\|K\|_{HS}$ is well defined and $\|K\|_{HS} = \|K^*\|_{HS}$. If $\|K\|_{HS}^2 < \infty$, then there exists a countable subset $\beta_0\subset\beta$ such that $Ke=0$ if $e\in\beta\setminus\beta_0$. Let $H_0 := \overline{\mathrm{span}(\beta_0)}$ and $B_0 := \overline{K(H_0)}$. Then $K\,(H)\subset B_0$, $K|_{H_0^{\perp}} = 0$ and hence by applying the results of Proposition B.17 to $K|_{H_0} : H_0\to B_0$ one easily sees that the separability of H and B are unnecessary in Proposition B.17.

Example B.19. Let (X,μ) be a measure space, $H = L^2(X,\mu)$ and

$$k(x,y) := \sum_{i=1}^{n} f_i\,(x)\,g_i\,(y)$$

where

$$f_i, g_i \in L^2(X,\mu) \text{ for } i=1,\dots,n.$$

Define

$$(Kf)\,(x) = \int_X k(x,y)f\,(y)\,d\mu\,(y),$$

then $K : L^2(X,\mu)\to L^2(X,\mu)$ is a finite rank operator and hence Hilbert Schmidt.

Exercise B.3. Suppose that (X,μ) is a σ–finite measure space such that $H = L^2(X,\mu)$ is separable and $k : X\times X\to\mathbb{R}$ is a measurable function, such that

$$\|k\|_{L^2(X\times X,\mu\otimes\mu)}^2 := \int_{X\times X}|k(x,y)|^2 d\mu\,(x)\,d\mu\,(y) < \infty.$$

Define, for $f\in H$,

$$Kf\,(x) = \int_X k(x,y)f\,(y)\,d\mu\,(y),$$

when the integral makes sense. Show:

1. $Kf(x)$ is defined for μ–a.e. x in X.
2. The resulting function Kf is in H and $K : H \to H$ is linear.
3. $\|K\|_{HS} = \|k\|_{L^2(X\times X,\mu\otimes\mu)} < \infty$. (This implies $K \in HS(H,H)$.)

Exercise B.4 (Converse to Exercise B.3). Suppose that (X,μ) is a σ–finite measure space such that $H = L^2(X,\mu)$ is separable and $K : H \to H$ is a Hilbert Schmidt operator. Show there exists $k \in L^2(X \times X, \mu\otimes\mu)$ such that K is the integral operator associated to k, i.e.

$$Kf(x) = \int_X k(x,y) f(y)\, d\mu(y). \tag{B.14}$$

In fact you should show

$$k(x,y) := \sum_{n=1}^{\infty} \left(\left(\overline{K^*\varphi_n}\right)(y)\right) \varphi_n(x) \;\; \left(L^2(\mu\otimes\mu) \text{ – convergent sum}\right) \tag{B.15}$$

where $\{\varphi_n\}_{n=1}^{\infty}$ is any orthonormal basis for H.

C

Trace Class & Fredholm Operators

al?

In this section H and B will be Hilbert spaces. Typically H and B will be separable, but we will not assume this until it is needed later.

C.1 Trace Class Operators

See B. Simon [44] for more details and ideals of compact operators.

Theorem C.1. *Let $A \in B(H)$ be a non-negative operator, $\{e_n\}_{n=1}^\infty$ be an orthonormal basis for H and*

$$\operatorname{tr}(A) := \sum_{n=1}^{\infty} \langle Ae_n | e_n \rangle .$$

Then $\operatorname{tr}(A) = \left\| \sqrt{A} \right\|_{HS}^2 \in [0, \infty]$ is well defined independent of the choice of orthonormal basis for H. Moreover if $\operatorname{tr}(A) < \infty$, then A is a compact operator.

Proof. Let $B := \sqrt{A}$, then

$$\operatorname{tr}(A) = \sum_{n=1}^{\infty} \langle Ae_n | e_n \rangle = \sum_{n=1}^{\infty} \left\langle B^2 e_n | e_n \right\rangle = \sum_{n=1}^{\infty} \langle Be_n | Be_n \rangle = \|B\|_{HS}^2 .$$

This shows $\operatorname{tr}(A)$ is well defined and that $\operatorname{tr}(A) = \left\| \sqrt{A} \right\|_{HS}^2$. If $\operatorname{tr}(A) < \infty$ then $\sqrt{A}$ is Hilbert Schmidt and hence compact. Therefore $A = \left(\sqrt{A}\right)^2$ is compact as well. ∎

Definition C.2. *An operator $A \in L(H, B)$ is **trace class** if $\operatorname{tr}(|A|) = \operatorname{tr}(\sqrt{A^* A}) < \infty$.*

Proposition C.3. *If $A \in L(H, B)$ is trace class then A is compact.*

Proof. By the polar decomposition Theorem **??**, $A = u|A|$ where u is a partial isometry and by Corollary **??** $|A|$ is also compact. Therefore A is compact as well. ∎

Proposition C.4. *If $A \in L(B)$ is trace class and $\{e_n\}_{n=1}^\infty$ is an orthonormal basis for H, then*

$$\operatorname{tr}(A) := \sum_{n=1}^{\infty} \langle Ae_n | e_n \rangle$$

is absolutely convergent and the sum is independent of the choice of orthonormal basis for H.

Proof. Let $A = u|A|$ be the polar decomposition of A and $\{\phi_n\}_{n=1}^\infty$ be an orthonormal basis of eigenvectors for $\operatorname{Nul}(|A|)^\perp = \operatorname{Nul}(A)^\perp$ such that

$$|A| \phi_m = \lambda_m \phi_m$$

with $\lambda_m \downarrow 0$ and $\sum_{m=1}^\infty \lambda_m < \infty$. Then

$$\begin{aligned}
\sum_n |\langle Ae_n | e_n \rangle| &= \sum_n |\langle |A| e_n | u^* e_n \rangle| = \sum_n \left| \sum_m \langle |A| e | \phi_m \rangle \langle \phi_m | u^* e_n \rangle \right| \\
&= \sum_n \left| \sum_m \lambda_m \langle e_n | \phi_m \rangle \langle \phi_m | u^* e_n \rangle \right| \\
&\le \sum_m \lambda_m \sum_n |\langle e_n | \phi_m \rangle \langle u\phi_m | e_n \rangle| \\
&= \sum_m \lambda_m |\langle \phi_m | u\phi_m \rangle| \le \sum_m \lambda_m < \infty.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\sum_n \langle Ae_n | e_n \rangle &= \sum_n \langle |A| e_n | u^* e_n \rangle = \sum_n \sum_m \lambda_m \langle e_n | \phi_m \rangle \langle \phi_m | u^* e_n \rangle \\
&= \sum_m \lambda_m \sum_n \langle u\phi_m | e_n \rangle \langle e_n | \phi_m \rangle \\
&= \sum_m \lambda_m \langle u\phi_m | \phi_m \rangle
\end{aligned}$$

showing $\sum_n \langle Ae_n | e_n \rangle = \sum_m \lambda_m \langle u\phi_m | \phi_m \rangle$ which proves $\operatorname{tr}(A)$ is well defined independent of basis. ∎

Remark C.5. Suppose K is a compact operator written in the form

$$Kf = \sum_{n=1}^{N} \lambda_n \langle f|\phi_n\rangle \psi_n \text{ for all } f \in H. \tag{C.1}$$

where $\{\phi_n\}_{n=1}^{\infty} \subset H$, $\{\psi_n\}_{n=1}^{\infty} \subset B$ are bounded sets and $\lambda_n \in \mathbb{C}$ such that $\sum_{n=1}^{\infty} |\lambda_n| < \infty$. Then K is trace class and

$$\operatorname{tr}(K) = \sum_{n=1}^{N} \lambda_n \langle \psi_n|\phi_n\rangle.$$

BRUCE STOP

Indeed, $K^* g = \sum_{n=1}^{N} \bar{\lambda}_n (g, \psi_n) \phi_n$ and hence

$$K^* K f = \sum_{n=1}^{N} \bar{\lambda}_n \langle Kf|\psi_n\rangle \phi_n$$

$$Kf = \sum_{n=1}^{N} \lambda_n \langle f|\phi_n\rangle \psi_n \text{ for all } f \in H. \tag{C.2}$$

We will say $K \in \mathcal{K}(H)$ is **trace class** if

$$\operatorname{tr}(\sqrt{K^* K}) := \sum_{n=1}^{N} \lambda_n < \infty$$

in which case we define

$$\operatorname{tr}(K) = \sum_{n=1}^{N} \lambda_n \langle \psi_n|\phi_n\rangle.$$

Notice that if $\{e_m\}_{m=1}^{\infty}$ is any orthonormal basis in H (or for the $\overline{\operatorname{Ran}(K)}$ if H is not separable) then

$$\begin{aligned}\sum_{m=1}^{M} \langle K e_m|e_m\rangle &= \sum_{m=1}^{M} \left\langle \sum_{n=1}^{N} \lambda_n \langle e_m|\phi_n\rangle \psi_n | e_m \right\rangle = \sum_{n=1}^{N} \lambda_n \sum_{m=1}^{M} \langle e_m|\phi_n\rangle \langle \psi_n|e_m\rangle \\ &= \sum_{n=1}^{N} \lambda_n \langle P_M \psi_n|\phi_n\rangle\end{aligned}$$

where P_M is orthogonal projection onto $\operatorname{Span}(e_1, \dots, e_M)$. Therefore by dominated convergence theorem ,

$$\begin{aligned}\sum_{m=1}^{\infty} \langle K e_m|e_m\rangle &= \lim_{M\to\infty} \sum_{n=1}^{N} \lambda_n \langle P_M \psi_n|\phi_n\rangle = \sum_{n=1}^{N} \lambda_n \lim_{M\to\infty} \langle P_M \psi_n|\phi_n\rangle \\ &= \sum_{n=1}^{N} \lambda_n \langle \psi_n|\phi_n\rangle = \operatorname{tr}(K).\end{aligned}$$

C.2 Fredholm Operators

Lemma C.6. *Let $M \subset H$ be a closed subspace and $V \subset H$ be a finite dimensional subspace. Then $M + V$ is closed as well. In particular if* $\operatorname{codim}(M) := \dim(H/M) < \infty$ *and $W \subset H$ is a subspace such that $M \subset W$, then W is closed and* $\operatorname{codim}(W) < \infty$.

Proof. Let $P : H \to M$ be orthogonal projection and let $V_0 := (I - P)V$. Since $\dim(V_0) \le \dim(V) < \infty$, V_0 is still closed. Also it is easily seen that $M + V = M \overset{\perp}{\oplus} V_0$ from which it follows that $M + V$ is closed because $\{z_n = m_n + v_n\} \subset M \overset{\perp}{\oplus} V_0$ is convergent iff $\{m_n\} \subset M$ and $\{v_n\} \subset V_0$ are convergent. If $\operatorname{codim}(M) < \infty$ and $M \subset W$, there is a finite dimensional subspace $V \subset H$ such that $W = M + V$ and so by what we have just proved, W is closed as well. It should also be clear that $\operatorname{codim}(W) \le \operatorname{codim}(M) < \infty$. ■

Lemma C.7. *If $K : H \to B$ is a finite rank operator, then there exists $\{\phi_n\}_{n=1}^{k} \subset H$ and $\{\psi_n\}_{n=1}^{k} \subset B$ such that*

1. *$Kx = \sum_{n=1}^{k} \langle x|\phi_n\rangle \psi_n$ for all $x \in H$.*
2. *$K^* y = \sum_{n=1}^{k} \langle y|\psi_n\rangle \phi_n$ for all $y \in B$, in particular K^* is still finite rank.*
   For the next two items, further assume $B = H$.
3. *$\dim \operatorname{Nul}(I + K) < \infty$.*
4. *$\dim \mathit{coker}(I + K) < \infty$, $\operatorname{Ran}(I + K)$ is closed and*

$$\operatorname{Ran}(I + K) = \operatorname{Nul}(I + K^*)^{\perp}.$$

Proof.

1. Choose $\{\psi_n\}_1^k$ to be an orthonormal basis for $\operatorname{Ran}(K)$. Then for $x \in H$,

$$Kx = \sum_{n=1}^{k} \langle Kx|\psi_n\rangle \psi_n = \sum_{n=1}^{k} \langle x|K^* \psi_n\rangle \psi_n = \sum_{n=1}^{k} \langle x|\phi_n\rangle \psi_n$$

   where $\phi_n := K^* \psi_n$.
2. Item 2. is a simple computation left to the reader.
3. Since $\operatorname{Nul}(I + K) = \{x \in H \mid x = -Kx\} \subset \operatorname{Ran}(K)$ it is finite dimensional.

4. Since $x = (I + K)x \in \operatorname{Ran}(I + K)$for $x \in \operatorname{Nul}(K)$, $\operatorname{Nul}(K) \subset \operatorname{Ran}(I + K)$. Since $\{\phi_1, \phi_2, \ldots, \phi_k\}^\perp \subset \operatorname{Nul}(K)$, $H = \operatorname{Nul}(K) + \operatorname{span}(\{\phi_1, \phi_2, \ldots, \phi_k\})$ and thus $\operatorname{codim}(\operatorname{Nul}(K)) < \infty$. From these comments and Lemma C.6, $\operatorname{Ran}(I + K)$ is closed and $\operatorname{codim}(\operatorname{Ran}(I + K)) \leq \operatorname{codim}(\operatorname{Nul}(K)) < \infty$. The assertion that $\operatorname{Ran}(I+K) = \operatorname{Nul}(I+K^*)^\perp$ is a consequence of Lemma 3.25 below.

∎

Definition C.8. *A bounded operator $F : H \to B$ is **Fredholm**iff the $\dim \operatorname{Nul}(F) < \infty$, $\dim coker(F) < \infty$ and $\operatorname{Ran}(F)$ is closed in B. (Recall: $coker(F) := B/\operatorname{Ran}(F)$.) The **index**of F is the integer,*

$$\begin{aligned} index(F) &= \dim \operatorname{Nul}(F) - \dim coker(F) &\quad (C.3)\\ &= \dim \operatorname{Nul}(F) - \dim \operatorname{Nul}(F^*). &\quad (C.4)\end{aligned}$$

Notice that equations (C.3) and (C.4) are the same since, (using $\operatorname{Ran}(F)$ is closed)

$$B = \operatorname{Ran}(F) \oplus \operatorname{Ran}(F)^\perp = \operatorname{Ran}(F) \oplus \operatorname{Nul}(F^*)$$

so that $\operatorname{coker}(F) = B/\operatorname{Ran}(F) \cong \operatorname{Nul}(F^*)$.

Lemma C.9. *The requirement that $\operatorname{Ran}(F)$ is closed in Definition C.8 is redundant.*

Proof. By restricting F to $\operatorname{Nul}(F)^\perp$, we may assume without loss of generality that $\operatorname{Nul}(F) = \{0\}$. Assuming $\dim \operatorname{coker}(F) < \infty$, there exists a finite dimensional subspace $V \subset B$ such that $B = \operatorname{Ran}(F) \oplus V$. Since V is finite dimensional, V is closed and hence $B = V \oplus V^\perp$. Let $\pi : B \to V^\perp$ be the orthogonal projection operator onto $V^\perp$ and let $G := \pi F : H \to V^\perp$ which is continuous, being the composition of two bounded transformations. Since G is a linear isomorphism, as the reader should check, the open mapping theorem implies the inverse operator $G^{-1} : V^\perp \to H$ is bounded. Suppose that $h_n \in H$ is a sequence such that $\lim_{n\to\infty} F(h_n) =: b$ exists in B. Then by composing this last equation with π, we find that $\lim_{n\to\infty} G(h_n) = \pi(b)$ exists in $V^\perp$. Composing this equation with G^{-1} shows that $h := \lim_{n\to\infty} h_n = G^{-1}\pi(b)$ exists in H. Therefore, $F(h_n) \to F(h) \in \operatorname{Ran}(F)$, which shows that $\operatorname{Ran}(F)$ is closed. ∎

Remark C.10. It is essential that the subspace $M := \operatorname{Ran}(F)$ in Lemma C.9 is the image of a bounded operator, for it is not true that every finite codimensional subspace M of a Banach space B is necessarily closed. To see this suppose that B is a separable infinite dimensional Banach space and let $A \subset B$ be an **algebraic** basis for B, which exists by a Zorn's lemma argument. Since $\dim(B) = \infty$ and B is complete, A must be uncountable. Indeed, if A were countable we could write $B = \cup_{n=1}^\infty B_n$ where B_n are finite dimensional (necessarily closed) subspaces of B. This shows that B is the countable union of nowhere dense closed subsets which violates the Baire Category theorem.

By separability of B, there exists a countable subset $A_0 \subset A$ such that the closure of $M_0 := \operatorname{span}(A_0)$ is equal to B. Choose $x_0 \in A \setminus A_0$, and let $M := \operatorname{span}(A \setminus \{x_0\})$. Then $M_{\mathbf{0}} \subset M$ so that $B = \bar{M}_0 = \bar{M}$, while $\operatorname{codim}(M) = 1$. Clearly this M can not be closed.

Example C.11. Suppose that H and B are finite dimensional Hilbert spaces and $F : H \to B$ is Fredholm. Then

$$\operatorname{index}(F) = \dim(B) - \dim(H). \quad (C.5)$$

The formula in Eq. (C.5) may be verified using the rank nullity theorem,

$$\dim(H) = \dim \operatorname{Nul}(F) + \dim \operatorname{Ran}(F),$$

and the fact that

$$\dim(B/\operatorname{Ran}(F)) = \dim(B) - \dim \operatorname{Ran}(F).$$

Theorem C.12. *A bounded operator $F : H \to B$ is Fredholm iff there exists a bounded operator $A : B \to H$ such that $AF - I$ and $FA - I$ are both compact operators. (In fact we may choose A so that $AF - I$ and $FA - I$ are both finite rank operators.)*

Proof. ($\Rightarrow$) Suppose F is Fredholm, then $F : \operatorname{Nul}(F)^\perp \to \operatorname{Ran}(F)$ is a bijective bounded linear map between Hilbert spaces. (Recall that $\operatorname{Ran}(F)$ is a closed subspace of B and hence a Hilbert space.) Let $\tilde{F}$ be the inverse of this map—a bounded map by the open mapping theorem. Let $P : H \to \operatorname{Ran}(F)$ be orthogonal projection and set $A := \tilde{F}P$. Then $AF - I = \tilde{F}PF - I = \tilde{F}F - I = -Q$ where Q is the orthogonal projection onto $\operatorname{Nul}(F)$. Similarly, $FA - I = F\tilde{F}P - I = -(I - P)$. Because $I - P$ and Q are finite rank projections and hence compact, both $AF - I$ and $FA - I$ are compact. ($\Leftarrow$) We first show that the operator $A : B \to H$ may be modified so that $AF - I$ and $FA - I$ are both finite rank operators. To this end let $G := AF - I$ (G is compact) and choose a finite rank approximation G_1 to G such that $G = G_1 + \mathcal{E}$ where $\|\mathcal{E}\| < 1$. Define $A_L : B \to H$ to be the operator $A_L := (I + \mathcal{E})^{-1}A$. Since $AF = (I + \mathcal{E}) + G_1$,

$$A_L F = (I + \mathcal{E})^{-1}AF = I + (I + \mathcal{E})^{-1}G_1 = I + K_L$$

where K_L is a finite rank operator. Similarly there exists a bounded operator $A_R : B \to H$ and a finite rank operator K_R such that $FA_R = I + K_R$. Notice that $A_L F A_R = A_R + K_L A_R$ on one hand and $A_L F A_R = A_L + A_L K_R$ on the other. Therefore, $A_L - A_R = A_L K_R - K_L A_R =: S$ is a finite rank operator. Therefore $FA_L = F(A_R + S) = I + K_R + FS$, so that $FA_L - I = K_R - FS$ is still a finite rank operator. Thus we have shown that there exists a bounded operator $\tilde{A} : B \to H$ such that $\tilde{A}F - I$ and $F\tilde{A} - I$ are both finite rank operators. We now assume that A is chosen such that $AF - I = G_1$, $FA - I = G_2$ are finite rank. Clearly $\mathrm{Nul}(F) \subset \mathrm{Nul}(AF) = \mathrm{Nul}(I + G_1)$ and $\mathrm{Ran}(F) \supseteq \mathrm{Ran}(FA) = \mathrm{Ran}(I + G_2)$. The theorem now follows from Lemma C.6 and Lemma C.7. ■

Corollary C.13. *If $F : H \to B$ is Fredholm then F^* is Fredholm and index$(F) = -$index(F^*).*

Proof. Choose $A : B \to H$ such that both $AF - I$ and $FA - I$ are compact. Then $F^*A^* - I$ and $A^*F^* - I$ are compact which implies that F^* is Fredholm. The assertion, $\mathrm{index}(F) = -\mathrm{index}(F^*)$, follows directly from Eq. (C.4). ■

Lemma C.14. *A bounded operator $F : H \to B$ is Fredholm if and only if there exists orthogonal decompositions $H = H_1 \oplus H_2$ and $B = B_1 \oplus B_2$ such that*

1. *H_1 and B_1 are closed subspaces,*
2. *H_2 and B_2 are finite dimensional subspaces, and*
3. *F has the block diagonal form*

$$F = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} : \begin{matrix} H_1 \\ \oplus \\ H_2 \end{matrix} \longrightarrow \begin{matrix} B_1 \\ \oplus \\ B_2 \end{matrix} \tag{C.6}$$

with $F_{11} : H_1 \to B_1$ being a bounded invertible operator.

Furthermore, given this decomposition, index$(F) = \dim(H_2) - \dim(B_2)$.

Proof. If F is Fredholm, set $H_1 = \mathrm{Nul}(F)^\perp$, $H_2 = \mathrm{Nul}(F)$, $B_1 = \mathrm{Ran}(F)$, and $B_2 = \mathrm{Ran}(F)^\perp$. Then $F = \begin{pmatrix} F_{11} & 0 \\ 0 & 0 \end{pmatrix}$, where $F_{11} := F|_{H_1} : H_1 \to B_1$ is invertible. For the converse, assume that F is given as in Eq. (C.6). Let $A := \begin{pmatrix} F_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ then

$$AF = \begin{pmatrix} I & F_{11}^{-1}F_{12} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} + \begin{pmatrix} 0 & F_{11}^{-1}F_{12} \\ 0 & -I \end{pmatrix},$$

so that $AF - I$ is finite rank. Similarly one shows that $FA - I$ is finite rank, which shows that F is Fredholm. Now to compute the index of F, notice that $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathrm{Nul}(F)$ iff

$$\begin{aligned} F_{11}x_1 + F_{12}x_2 &= 0 \\ F_{21}x_1 + F_{22}x_2 &= 0 \end{aligned}$$

which happens iff $x_1 = -F_{11}^{-1}F_{12}x_2$ and $(-F_{21}F_{11}^{-1}F_{12} + F_{22})x_2 = 0$. Let $D := (F_{22} - F_{21}F_{11}^{-1}F_{12}) : H_2 \to B_2$, then the mapping

$$x_2 \in \mathrm{Nul}(D) \to \begin{pmatrix} -F_{11}^{-1}F_{12}x_2 \\ x_2 \end{pmatrix} \in \mathrm{Nul}(F)$$

is a linear isomorphism of vector spaces so that $\mathrm{Nul}(F) \cong \mathrm{Nul}(D)$. Since

$$F^* = \begin{pmatrix} F_{11}^* & F_{21}^* \\ F_{12}^* & F_{22}^* \end{pmatrix} \quad \begin{matrix} B_1 \\ \oplus \\ B_2 \end{matrix} \longrightarrow \begin{matrix} H_1 \\ \oplus \\ H_2 \end{matrix},$$

similar reasoning implies $\mathrm{Nul}(F^*) \cong \mathrm{Nul}(D^*)$. This shows that $\mathrm{index}(F) = \mathrm{index}(D)$. But we have already seen in Example C.11 that $\mathrm{index}(D) = \dim H_2 - \dim B_2$. ■

Proposition C.15. *Let F be a Fredholm operator and K be a compact operator from $H \to B$. Further assume $T : B \to X$ (where X is another Hilbert space) is also Fredholm. Then*

1. *the Fredholm operators form an open subset of the bounded operators. Moreover if $\mathcal{E} : H \to B$ is a bounded operator with $\|\mathcal{E}\|$ sufficiently small we have index(F) =index$(F + \mathcal{E})$.*
2. *$F + K$ is Fredholm and index(F) = index$(F + K)$.*
3. *TF is Fredholm and index(TF) = index(T) + index(F)*

Proof.

1. We know F may be written in the block form given in Eq. (C.6) with $F_{11} : H_1 \to B_1$ being a bounded invertible operator. Decompose $\mathcal{E}$ into the block form as
$$\mathcal{E} = \begin{pmatrix} \mathcal{E}_{11} & \mathcal{E}_{12} \\ \mathcal{E}_{21} & \mathcal{E}_{22} \end{pmatrix}$$
and choose $\|\mathcal{E}\|$ sufficiently small such that $\|\mathcal{E}_{11}\|$ is sufficiently small to guarantee that $F_{11} + \mathcal{E}_{11}$ is still invertible. (Recall that the invertible operators form an open set.) Thus $F + \mathcal{E} = \begin{pmatrix} F_{11} + \mathcal{E}_{11} & * \\ * & * \end{pmatrix}$ has the block form of a Fredholm operator and the index may be computed as:
$$\mathrm{index}(F + \mathcal{E}) = \dim H_2 - \dim B_2 = \mathrm{index}(F).$$
2. Given $K : H \to B$ compact, it is easily seen that $F + K$ is still Fredholm. Indeed if $A : B \to H$ is a bounded operator such that $G_1 := AF - I$ and $G_2 := FA - I$ are both compact, then $A(F + K) - I = G_1 + AK$ and $(F + K)A - I = G_2 + KA$ are both compact. Hence $F + K$ is Fredholm by Theorem C.12. By item 1., the function $f(t) := \mathrm{index}(F + tK)$ is a continuous locally constant function of $t \in \mathbb{R}$ and hence is constant. In particular, $\mathrm{index}(F + K) = f(1) = f(0) = \mathrm{index}(F)$.

3. It is easily seen, using Theorem C.12 that the product of two Fredholm operators is again Fredholm. So it only remains to verify the index formula in item 3. For this let $H_1 := \mathrm{Nul}(F)^\perp$, $H_2 := \mathrm{Nul}(F)$, $B_1 := \mathrm{Ran}(T) = T(H_1)$, and $B_2 := \mathrm{Ran}(T)^\perp = \mathrm{Nul}(T^*)$. Then F decomposes into the block form:
$$F = \begin{pmatrix} \tilde{F} & 0 \\ 0 & 0 \end{pmatrix} : \begin{matrix} H_1 \\ \oplus \\ H_2 \end{matrix} \longrightarrow \begin{matrix} B_1 \\ \oplus \\ B_2 \end{matrix},$$
where $\tilde{F} = F|_{H_1} : H_1 \to B_1$ is an invertible operator. Let $Y_1 := T(B_1)$ and $Y_2 := Y_1^\perp = T(B_1)^\perp$. Notice that $Y_1 = T(B_1) = TQ(B_1)$, where $Q : B \to B_1 \subset B$ is orthogonal projection onto B_1. Since B_1 is closed and B_2 is finite dimensional, Q is Fredholm. Hence TQ is Fredholm and $Y_1 = TQ(B_1)$ is closed in Y and is of finite codimension. Using the above decompositions, we may write T in the block form:
$$T = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} : \begin{matrix} B_1 \\ \oplus \\ B_2 \end{matrix} \longrightarrow \begin{matrix} Y_1 \\ \oplus \\ Y_2 \end{matrix}.$$
Since $R = \begin{pmatrix} 0 & T_{12} \\ T_{21} & T_{22} \end{pmatrix} : B \to Y$ is a finite rank operator and hence $RF : H \to Y$ is finite rank, $\mathrm{index}(T - R) = \mathrm{index}(T)$ and $\mathrm{index}(TF - RF) = \mathrm{index}(TF)$. Hence without loss of generality we may assume that T has the form $T = \begin{pmatrix} \tilde{T} & 0 \\ 0 & 0 \end{pmatrix}$, ($\tilde{T} = T_{11}$) and hence
$$TF = \begin{pmatrix} \tilde{T}\tilde{F} & 0 \\ 0 & 0 \end{pmatrix} : \begin{matrix} H_1 \\ \oplus \\ H_2 \end{matrix} \longrightarrow \begin{matrix} Y_1 \\ \oplus \\ Y_2 \end{matrix}.$$
We now compute the $\mathrm{index}(T)$. Notice that $\mathrm{Nul}(T) = \mathrm{Nul}(\tilde{T}) \oplus B_2$ and $\mathrm{Ran}(T) = \tilde{T}(B_1) = Y_1$. So
$$\mathrm{index}(T) = \mathrm{index}(\tilde{T}) + \dim(B_2) - \dim(Y_2).$$
Similarly,
$$\mathrm{index}(TF) = \mathrm{index}(\tilde{T}\tilde{F}) + \dim(H_2) - \dim(Y_2),$$
and as we have already seen
$$\mathrm{index}(F) = \dim(H_2) - \dim(B_2).$$
Therefore,

$$\mathrm{index}(TF) - \mathrm{index}(T) - \mathrm{index}(F) = \mathrm{index}(\tilde{T}\tilde{F}) - \mathrm{index}(\tilde{T}).$$
Since $\tilde{F}$ is invertible, $\mathrm{Ran}(\tilde{T}) = \mathrm{Ran}(\tilde{T}\tilde{F})$ and $\mathrm{Nul}(\tilde{T}) \cong \mathrm{Nul}(\tilde{T}\tilde{F})$. Thus $\mathrm{index}(\tilde{T}\tilde{F}) - \mathrm{index}(\tilde{T}) = 0$ and the theorem is proved.

■

C.3 Tensor Product Spaces

References for this section are Reed and Simon [36] (Volume 1, Chapter VI.5), Simon [45], and Schatten [42]. See also Reed and Simon [35] (Volume 2 § IX.4 and §XIII.17).

Let H and K be separable Hilbert spaces and $H \otimes K$ will denote the usual Hilbert completion of the algebraic tensors $H \otimes_f K$. Recall that the inner product on $H \otimes K$ is determined by $\langle h \otimes k | h' \otimes k' \rangle = \langle h | h' \rangle \langle k | k' \rangle$. The following proposition is well known.

Proposition C.16 (Structure of $H \otimes K$). *There is a bounded linear map $T : H \otimes K \to B_{anti}(K, H)$ (the space of bounded anti-linear maps from K to H) determined by*
$$T(h \otimes k)k' := \langle k | k' \rangle h \text{ for all } k, k' \in K \text{ and } h \in H.$$
Moreover $T(H \otimes K) = HS(K, H)$ — the Hilbert Schmidt operators[1] from K to H. The map $T : H \otimes K \to HS(K, H)$ is unitary equivalence of Hilbert spaces. Finally, any $A \in H \otimes K$ may be expressed as
$$A = \sum_{n=1}^{\infty} \lambda_n h_n \otimes k_n, \tag{C.7}$$
where $\{h_n\}$ and $\{k_n\}$ are orthonormal sets in H and K respectively and $\{\lambda_n\} \subset \mathbb{R}$ such that $\|A\|^2 = \sum |\lambda_n|^2 < \infty$.

Proof. Let $A := \sum a_{ji} h_j \otimes k_i$, where $\{h_i\}$ and $\{k_j\}$ are orthonormal bases for H and K respectively and $\{a_{ji}\} \subset \mathbb{R}$ such that $\|A\|^2 = \sum |a_{ji}|^2 < \infty$. Then evidently, $T(A)k := \sum a_{ji} h_j \langle k_i | k \rangle$ and
$$\begin{aligned} \|T(A)k\|^2 &= \sum_j |\sum_i a_{ji} \langle k_i | k \rangle|^2 \le \sum_j \sum_i |a_{ji}|^2 |\langle k_i | k \rangle|^2 \\ &\le \sum_j \sum_i |a_{ji}|^2 \|k\|^2. \end{aligned}$$

[1] Don't we need to use the anti-linear HS operators here. Perhaps we should use the opposite Hilbert space instead somewhere.

Thus $T : H \otimes K \to B(K, H)$ is bounded. Moreover,

$$\|T(A)\|_{HS}^2 := \sum \|T(A)k_i\|^2 = \sum_{ij} |a_{ji}|^2 = \|A\|^2,$$

which proves the T is an isometry. We will now prove that T is surjective and at the same time prove Eq. (C.7). To motivate the construction, suppose that $Q = T(A)$ where A is given as in Eq. (C.7). Then

$$Q^*Q = T\left(\sum_{n=1}^{\infty} \lambda_n k_n \otimes h_n\right) T\left(\sum_{n=1}^{\infty} \lambda_n h_n \otimes k_n\right) = T\left(\sum_{n=1}^{\infty} \lambda_n^2 k_n \otimes k_n\right).$$

That is $\{k_n\}$ is an orthonormal basis for $(\operatorname{Nul} Q^*Q)^{\perp}$ with $Q^*Qk_n = \lambda_n^2 k_n$. Also $Qk_n = \lambda_n h_n$, so that $h_n = \lambda_n^{-1} Qk_n$. We will now reverse the above argument. Let $Q \in HS(K, H)$. Then Q^*Q is a self-adjoint compact operator on K. Therefore there is an orthonormal basis $\{k_n\}_{n=1}^{\infty}$ for the $(\operatorname{Nul} Q^*Q)^{\perp}$ which consists of eigenvectors of Q^*Q. Let $\lambda_n \in (0, \infty)$ such that $Q^*Qk_n = \lambda_n^2 k_n$ and set $h_n = \lambda_n^{-1} Qk_n$. Notice that

$$\begin{aligned}\langle h_n | h_m \rangle &= \left\langle \lambda_n^{-1} Qk_n | \lambda_m^{-1} Qk_m \right\rangle \\ &= \left\langle \lambda_n^{-1} k_n | \lambda_m^{-1} Q^*Qk_m \right\rangle = \left\langle \lambda_n^{-1} k | \lambda_m^{-1} \lambda_m^2 k_m \right\rangle = \delta_{mn},\end{aligned}$$

so that $\{h_n\}$ is an orthonormal set in H. Define

$$A = \sum_{n=1}^{\infty} \lambda_n h_n \otimes k_n$$

and notice that $T(A)k_n = \lambda_n h_n = Qk_n$ for all n and $T(A)k = 0$ for all $k \in \operatorname{Nul} Q = \operatorname{Nul} Q^*Q$. That is $T(A) = Q$. Therefore T is surjective and Eq. (C.7) holds. ∎

Notation C.17 *In the future we will identify $A \in H \otimes K$ with $T(A) \in HS(K, H)$ and drop T from the notation. So that with this notation we have $(h \otimes k)k' = \langle k | k' \rangle h$.*

Let $A \in H \otimes H$, we set $\|A\|_1 := \operatorname{tr}\sqrt{A^*A} := \operatorname{tr}\sqrt{T(A)^*T(A)}$ and we let

$$H \otimes_1 H := \{A \in H \otimes H : \|A\|_1 < \infty\}.$$

We will now compute $\|A\|_1$ for $A \in H \otimes H$ described as in Eq. (C.7). First notice that $A^* = \sum_{n=1}^{\infty} \lambda_n k_n \otimes h_n$ and

$$A^*A = \sum_{n=1}^{\infty} \lambda_n^2 k_n \otimes k_n.$$

Hence $\sqrt{A^*A} = \sum_{n=1}^{\infty} |\lambda_n| k_n \otimes k_n$ and hence $\|A\|_1 = \sum_{n=1}^{\infty} |\lambda_n|$. Also notice that $\|A\|^2 = \sum_{n=1}^{\infty} |\lambda_n|^2$ and $\|A\|_{op} = \max_n |\lambda_n|$. Since

$$\|A\|_1^2 = \{\sum_{n=1}^{\infty} |\lambda_n|\}^2 \geq \sum_{n=1}^{\infty} |\lambda_n|^2 = \|A\|^2,$$

we have the following relations among the various norms,

$$\|A\|_{op} \leq \|A\| \leq \|A\|_1. \tag{C.8}$$

Proposition C.18. *There is a continuous linear map $C : H \otimes_1 H \to \mathbb{R}$ such that $C(h \otimes k) = (h, k)$ for all $h, k \in H$. If $A \in H \otimes_1 H$, then*

$$CA = \sum \langle e_m \otimes e_m | A \rangle, \tag{C.9}$$

where $\{e_m\}$ is any orthonormal basis for H. Moreover, if $A \in H \otimes_1 H$ is positive, i.e. $T(A)$ is a non-negative operator, then $\|A\|_1 = CA$.

Proof. Let $A \in H \otimes_1 H$ be given as in Eq. (C.7) with $\sum_{n=1}^{\infty} |\lambda_n| = \|A\|_1 < \infty$. Then define $CA := \sum_{n=1}^{\infty} \lambda_n (h_n, k_n)$ and notice that $|CA| \leq \sum |\lambda_n| = \|A\|_1$, which shows that C is a contraction on $H \otimes_1 H$. (Using the universal property of $H \otimes_f H$ it is easily seen that C is well defined.) Also notice that for $M \in \mathbb{Z}_+$ that

$$\sum_{m=1}^{M} \langle e_m \otimes e_m | A \rangle = \sum_{n=1}^{\infty} \sum_{m=1}^{M} (e_m \otimes e_m, \lambda_n h_n \otimes k_n,), \tag{C.10}$$

$$= \sum_{n=1}^{\infty} \lambda_n \langle P_M h | k_n \rangle, \tag{C.11}$$

where P_M denotes orthogonal projection onto $\operatorname{span}\{e_m\}_{m=1}^{M}$. Since $|\lambda_n \langle P_M h_n | k_n \rangle| \leq |\lambda_n|$ and $\sum_{n=1}^{\infty} |\lambda_n| = \|A\|_1 < \infty$, we may let $M \to \infty$ in Eq. (C.11) to find that

$$\sum_{m=1}^{\infty} \langle e_m \otimes e_m | A \rangle = \sum_{n=1}^{\infty} \lambda_n \langle h_n | k_n \rangle = CA.$$

This proves Eq. (C.9). For the final assertion, suppose that $A \geq 0$. Then there is an orthonormal basis $\{k_n\}_{n=1}^{\infty}$ for the $(\operatorname{Nul} A)^{\perp}$ which consists of eigenvectors of A. That is $A = \sum \lambda_n k_n \otimes k_n$ and $\lambda_n \geq 0$ for all n. Thus $CA = \sum \lambda_n$ and $\|A\|_1 = \sum \lambda_n$.

Proposition C.19 (Noncommutative Fatou' s Lemma). *Let A_n be a sequence of positive operators on a Hilbert space H and $A_n \to A$ weakly as $n \to \infty$, then*

$$trA \le \liminf_{n\to\infty} trA_n. \quad (C.12)$$

Also if $A_n \in H \otimes_1 H$ *and* $A_n \to A$ *in* $B(H)$, *then*

$$\|A\|_1 \le \liminf_{n\to\infty} \|A_n\|_1. \quad (C.13)$$

■

Proof. Let A_n be a sequence of positive operators on a Hilbert space H and $A_n \to A$ weakly as $n \to \infty$ and $\{e_k\}_{k=1}^\infty$ be an orthonormal basis for H. Then by Fatou's lemma for sums,

$$\begin{aligned}\mathrm{tr}A &= \sum_{k=1}^\infty \langle Ae_k|e_k\rangle = \sum_{k=1}^\infty \lim_{n\to\infty} \langle A_n e_k|e_k\rangle \\ &\le \liminf_{n\to\infty} \sum_{k=1}^\infty \langle A_n e_k|e_k\rangle = \liminf_{n\to\infty} \mathrm{tr}A_n.\end{aligned}$$

Now suppose that $A_n \in H \otimes_1 H$ and $A_n \to A$ in $B(H)$. Then by Proposition **??**, $|A_n| \to |A|$ in $B(H)$ as well. Hence by Eq. (C.12), $\|$

$$A\|_1 := \mathrm{tr}\,|A| \le \liminf_{n\to\infty} \mathrm{tr}|A_n| \le \liminf_{n\to\infty} \|A_n\|_1.$$

■

Proposition C.20. *Let* X *be a Banach space,* $B : H \times K \to X$ *be a bounded bi-linear form, and*

$$\|B\| := \sup\{|B(h,k)| : \|h\|\|k\| \le 1\}.$$

Then there is a unique bounded linear map $\tilde{B} : H \otimes_1 K \to X$ *such that* $\tilde{B}(h \otimes k) = B(h,k)$. *Moreover* $\|\tilde{B}\|_{op} = \|\tilde{B}\|$.

Proof. Let $A = \sum_{n=1}^\infty \lambda_n h_n \otimes k_n \in H \otimes_1 K$ as in Eq. (C.7). Clearly, if $\tilde{B}$ is to exist we must have $\tilde{B}(A) := \sum_{n=1}^\infty \lambda_n B(h_n, k_n)$. Notice that

$$\sum_{n=1}^\infty |\lambda_n||B(h_n,k_n)| \le \sum_{n=1}^\infty |\lambda_n|\|B\| = \|A\|_1 \cdot \|B\|.$$

This shows that $\tilde{B}(A)$ is well defined and that $\|\tilde{B}\|_{op} \le \|\tilde{B}\|$. The opposite inequality follows from the trivial computation:

$$\begin{aligned}\|B\| &= \sup\{|B(h,k)| : \|h\|\|k\| = 1\} \\ &= \sup\{|\tilde{B}(h\otimes k)| : \|h \otimes_1 k\|_1 = 1\} \le \|\tilde{B}\|_{op}.\end{aligned}$$

■

Lemma C.21. *Suppose that* $P \in B(H)$ *and* $Q \in B(K)$, *then* $P \otimes Q : H \otimes K \to H \otimes K$ *is a bounded operator. Moreover,* $P \otimes Q(H \otimes_1 K) \subset H \otimes_1 K$ *and we have the norm equalities*

$$\|P \otimes Q\|_{B(H\otimes K)} = \|P\|_{B(H)}\|Q\|_{B(K)}$$

and

$$\|P \otimes Q\|_{B(H\otimes_1 K)} = \|P\|_{B(H)}\|Q\|_{B(K)}.$$

Proof. We will give essentially the same proof of $\|P \otimes Q\|_{B(H\otimes K)} = \|P\|_{B(H)}\|Q\|_{B(K)}$ as the proof on p. 299 of Reed and Simon [36]. Let $A \in H \otimes K$ as in Eq. (C.7). Then

$$(P \otimes I)A = \sum_{n=1}^\infty \lambda_n Ph_n \otimes k_n$$

and hence

$$(P \otimes I)A\{(P \otimes I)A\}^* = \sum_{n=1}^\infty \lambda_n^2 Ph_n \otimes Ph_n.$$

Therefore,

$$\begin{aligned}\|(P \otimes I)A\|^2 &= \mathrm{tr}(P \otimes I)A\{(P \otimes I)A\}^* \\ &= \sum_{n=1}^\infty \lambda_n^2 (Ph_n, Ph_n) \le \|P\|^2 \sum_{n=1}^\infty \lambda_n^2 \\ &= \|P\|^2\|A\|_1^2,\end{aligned}$$

which shows that Thus $\|P \otimes I\|_{B(H\otimes K)} \le \|P\|$. By symmetry, $\|I \otimes Q\|_{B(H\otimes K)} \le \|Q\|$. Since $P \otimes Q = (P \otimes I)(I \otimes Q)$, we have

$$\|P \otimes Q\|_{B(H\otimes K)} \le \|P\|_{B(H)}\|Q\|_{B(K)}.$$

The reverse inequality is easily proved by considering $P \otimes Q$ on elements of the form $h \otimes k \in H \otimes K$. Now suppose that $A \in H \otimes_1 K$ as in Eq. (C.7). Then

$$\begin{aligned}\|(P \otimes Q)A\|_1 &\le \sum_{n=1}^\infty |\lambda_n|\|Ph_n \otimes Qk_n\|_1 \\ &\le \|P\|\|Q\| \sum_{n=1}^\infty |\lambda_n| = \|P\|\|Q\|\|A\|,\end{aligned}$$

which shows that

$$\|P \otimes Q\|_{B(H\otimes_1 K)} \le \|P\|_{B(H)}\|Q\|_{B(K)}.$$

Again the reverse inequality is easily proved by considering $P \otimes Q$ on elements of the form $h \otimes k \in H \otimes_1 K$. ■

Lemma C.22. *Suppose that P_m and Q_m are orthogonal projections on H and K respectively which are strongly convergent to the identity on H and K respectively. Then $P_m \otimes Q_m : H \otimes_1 K \to H \otimes_1 K$ also converges strongly to the identity in $H \otimes_1 K$.*

Proof. Let $A = \sum_{n=1}^{\infty} \lambda_n h_n \otimes k_n \in H \otimes_1 K$ as in Eq. (C.7). Then

$$\begin{aligned}
&\|P_m \otimes Q_m A - A\|_1 \\
&\quad \le \sum_{n=1}^{\infty} |\lambda_n| \|P_m h_n \otimes Q_m k_n - h_n \otimes k_n\|_1 \\
&\quad = \sum_{n=1}^{\infty} |\lambda_n| \|(P_m h_n - h_n) \otimes Q_m k_n + h_n \otimes (Q_m k_n - k_n)\|_1 \\
&\quad \le \sum_{n=1}^{\infty} |\lambda_n| \{\|P_m h_n - h_n\| \|Q_m k_n\| + \|h_n\| \|Q_m k_n - k_n\|\} \\
&\quad \le \sum_{n=1}^{\infty} |\lambda_n| \{\|P_m h_n - h_n\| + \|Q_m k_n - k_n\|\} \to 0 \text{ as } m \to \infty
\end{aligned}$$

by the dominated convergence theorem. ■