Bruce K. Driver

Functional Analysis Tools with Examples

January 21, 2020 *File:241Functional˙2020s.tex*

Contents

Part I

Background

1

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 [18–21] and especially [74]. 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,*

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

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

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

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 $\{U \ne 0\}$ is a null set and therefore $\|u(\omega)\| = 0$ if ω is not in the null set, $E \cup \{U \ne 0\}$.

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

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

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

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

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

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

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

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

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 $w \in D(z_0, \rho)$,*

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

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

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

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

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

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 $t \ge 0$.

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

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

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 $\sum_{j=1}^\infty \mu(\{\|G_{j+1} - G_j\| > \varepsilon_j\}) \le \sum_{j=1}^\infty \varepsilon_j < \infty$.

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

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

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

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 $j \in \mathbb{N}$,

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

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

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

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,

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

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,

$$\|G_j - F\|_p^p = \int_\Omega \liminf_{k \to \infty} \|G_j - G_k\|^p d\mu \leq \liminf_{k \to \infty} \int_\Omega \|G_j - G_k\|^p d\mu$$
$$= \liminf_{k \to \infty} \|G_j - G_k\|_p^p \to 0 \text{ as } j \to \infty.$$

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

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

∎

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, $\mu(F \neq 0) < \infty$ 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 $1 \leq p < \infty$ 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

$$B_n^\varepsilon := \left\{x \in X : \|x - x_n\| \leq \min\left(\varepsilon, \frac{1}{2}\|x_n\|\right)\right\}$$

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 $\|y - x_n\| \leq \varepsilon$ and $\frac{1}{2}\|x_n\| \leq \|y\| \leq \frac{3}{2}\|x_n\|$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

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

Putting these two estimates together shows,

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

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 $\|F_\varepsilon^N\| \leq \|F_\varepsilon\| \leq 2\|F\|$ for all N. Therefore another application of the dominated convergence theorem implies, $\lim_{N \to \infty} \|F_\varepsilon^N - F_\varepsilon\|_{L^p(\mu; X)} = 0$. 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 $\|F(\omega) - x_n\| \le \frac{1}{2}\|x_n\|$ and hence,

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

From this inequality we see that $\varphi(F(\omega)) \ge \frac{1}{2}\|x_n\| > 0$ 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* $F \in \mathcal{S}(\mu; X)$,

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

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

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:

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

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* $\|Tz\| \le C\|z\|$ *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*

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

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* $F \in L^1(\Omega, \mathcal{F}, \mu; X)$,

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

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

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

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

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

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,*

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

In other words, $y \in Y$ is in $\operatorname{essran}_\mu(q)$ 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, $X_0 := \overline{\operatorname{span}(\operatorname{essran}_\mu(F))}$. 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 $\|u_n\|_X \leq g$ 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,*

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

Proof. Since $\|u(\omega)\|_X = \lim_{n\to\infty} \|u_n(\omega)\| \leq g(\omega)$ for a.e. ω, it follows that $u \in L^1(\mu, X)$. Moreover, $\|u - u_n\|_X \leq 2g$ 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

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

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,

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

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

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

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

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

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

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

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,
$$\begin{aligned}g(x,y) &:= \sum_{i=1}^\infty |\langle D(y) v, e_i\rangle| \, |\langle C(x) e_i, w\rangle| \\ &\le \sqrt{\sum_{i=1}^\infty |\langle D(y) v, e_i\rangle|^2 \sum_{i=1}^\infty |\langle C(x) e_i, w\rangle|^2} \\ &= \sqrt{\|D(y) v\|^2 \cdot \|C(x)^* w\|^2} \\ &\le \|D(y)\|_{op} \|C^*(x)\|_{op} \|v\| \, \|w\| \\ &\quad = \|D(y)\|_{op} \|C(x)\|_{op} \|v\| \, \|w\|,\end{aligned}$$
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 $\mu -$ 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 [41, 42].For general Banach and C^*-algebra stuff have a look at [38, 80]. Also see the lecture notes in [47, 79]. Putnam's file looked quite good. For a very detailed statements see [11, 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

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

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

$$\|(a,\alpha)\| = \sup\{\|ab + \alpha b\| : b \in \mathcal{A} \text{ with } \|b\| \le 1\} \tag{2.3}$$

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 $L_a b = ab$ for all $a, b \in \mathcal{A}$. Indeed, $\|L_a b\| = \|ab\| \le \|a\|\|b\|$ 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

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

and so taking the sup of this expression over $\|b\| \le 1$ 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 $a \in \mathcal{A}$,

$$\|a\| = \left\|a\,\frac{a^*}{\|a\|}\right\| \le \sup\{\|ab\| : b \in \mathcal{A} \text{ with } \|b\| \le 1\} \le \|a\|$$

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

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

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

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

Proof. By the assumptions and Proposition 2.12, both x and $1+x^{-1}h$ are invertible with

$$\left\|1+x^{-1}h\right\| \leq \frac{1}{1-\|x^{-1}h\|}.$$

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

$$\begin{aligned} \|R_N\| &\leq \left\|x^{-1}\right\| \left\|\left(1+x^{-1}h\right)^{-1}\right\| \left\|\left(x^{-1}h\right)^{N+1}\right\| \\ &\leq \left\|\left(x^{-1}h\right)^{N+1}\right\| \left\|x^{-1}\right\| \frac{1}{1-\|x^{-1}h\|}. \end{aligned}$$

■

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 $\left\|x^{-1}h\right\| \leq \left\|x^{-1}\right\| \|h\| < 1$. 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

$$\left\|(x+h)^{-1} - x^{-1}\right\| \leq \left\|x^{-1}\right\| \left\|(x+h)^{-1}\right\| \|h\| \leq \left\|x^{-1}\right\|^2 \cdot \frac{\|h\|}{1-\|x^{-1}h\|} \to 0 \text{ as } h \to 0.$$

■

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 $\|e^{ta}\| \leq e^{|t|\|a\|}$.
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*

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

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

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

Moreover,

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

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 $\mathrm{ad}_a \in B(\mathcal{A})$ be defined by $\mathrm{ad}_a b = ab - ba$.*

Notice that

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

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

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

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

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

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

Thus it follows that

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

From this we conclude,

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

■

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,

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

where

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

because

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

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*

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

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*

$$u^A(t,s) := \left((\mathcal{I} - \Lambda_s)^{-1}\mathbf{1}\right)(t) = \sum_{n=0}^\infty u_n^A(t,s) \tag{2.20}$$

$$= \mathbf{1} + \sum_{n=1}^\infty \int_s^t d\tau_n \int_s^{\tau_n} d\tau_{n-1} \cdots \int_s^{\tau_2} d\tau_1 A(\tau_n)\dots A(\tau_1). \tag{2.21}$$

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. ∎

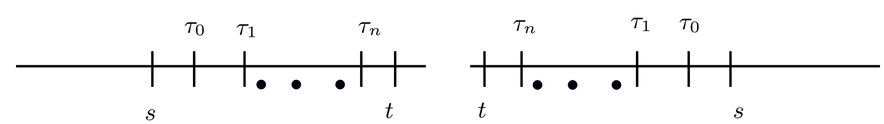

Fig. 2.1. This figures shows how to find the new limits of integration when $t > s$ and $t < s$ respectively.

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

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

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

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

where by summing Eq. (2.24),

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

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

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

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, $\ln(e^{tC}) := \ln(1+A(t))$ 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 $\ln(e^{tC}) = tC$ 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 [11]. [That this result holds under the additional assumption that $\mathcal{B}$ is commutative and "symmetric" is contained in Theorem **??** 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.31, 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 [9, 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 $B(L^2(\mu))$. Then $\mathcal{B}$ is a C^*-subalgebra of $B(L^2(\mu))$ 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 $T_1, \ldots, T_n \in B(H)$, let $\mathcal{A}(T_1, \ldots, T_n)$ be the smallest subalgebra of $B(H)$ containing $\{T_1, \ldots, T_n\}$, i.e. $\mathcal{A}$ consists of linear combination of words in $\{T_1, \ldots, T_n\}$. With this notation, $\mathcal{A}(T_1, \ldots, T_n, T_1^*, \ldots, T_n^*)$ is the smallest $*$ -sub-algebra of $B(H)$ which contains $\{T_1, \ldots, T_n\}$. We let

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

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

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

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

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

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

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

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

For most of this chapter we will mostly interested in the commutative $*$-subalgebra, $\mathcal{A}(I, T)$ 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, $\{f_j\}_{j=1}^n \subset \mathcal{B}$, $\mathbf{f} = (f_1, \ldots, f_n) : \Omega \to \mathbb{C}^n$, and* $\operatorname{essran}_\mu(\mathbf{f})$ *be the essential range of* $\mathbf{f}$ *(see Definition 1.32). Then* $\hat{\mathbf{f}} : C(\operatorname{essran}_\mu(\mathbf{f})) \to L^{\infty}(\mu)$ *defined by* $\hat{\mathbf{f}}(\psi) = \psi(\mathbf{f})$ *for all* $\psi \in C(\operatorname{essran}_\mu(\mathbf{f}))$ *is an isometric C^*-isomorphism onto* $C^*(\mathbf{f}, 1)$.

Proof. Let us first show that

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

It is clear that $\|\psi(\mathbf{f})\|_{\infty} \le \|\psi\|_{C(\operatorname{essran}_\mu(\mathbf{f}))}$. If $M < \|\psi\|_{C(\operatorname{essran}_\mu(\mathbf{f}))}$, then there exists $\mathbf{z} \in \operatorname{essran}_\mu(\mathbf{f})$ 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 $M < \|\psi\|_{C(\operatorname{essran}_\mu(\mathbf{f}))}$ was arbitrary, it follows that $\|\psi(\mathbf{f})\|_{\infty} \ge \|\psi\|_{C(\operatorname{essran}_\mu(\mathbf{f}))}$ and Eq. (2.41) is proved.

Let $\mathcal{B}_0 := \hat{\mathbf{f}}(C(\operatorname{essran}_\mu(\mathbf{f})))$ 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 $\psi \in C(\operatorname{essran}_\mu(\mathbf{f}))$, there exists $p_k \in \mathbb{C}[z_1, \ldots, z_n, \bar{z}_1, \ldots, \bar{z}_n]$ such that

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

Using

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

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

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

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

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

and that

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

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

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

and hence $u(t) = B$ for all t, i.e. $e^{wN} B e^{-wM} = B$ 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 $\mathbf{T} \cup \mathbf{T}^* := \{T_j, T_j^*\}_{j=1}^n$ 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 **??** 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 **??** 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 **??** below for some more information about the symmetry condition on a Banach algebra. See Exercise **??** 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 **??**.

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

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

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

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
$$\frac{d}{dt}T_t S_\varepsilon = \frac{1}{\varepsilon}[T_{t+\varepsilon} - T_t] \text{ for } t > 0 \text{ and}$$
$$\frac{d}{dt}|_{0+} T_t S_\varepsilon := \lim_{t \downarrow 0}\left(\frac{T_t - I}{t}\right)S_\varepsilon = \frac{1}{\varepsilon}[T_\varepsilon - I].$$
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 3.40 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

$$\begin{aligned}
\sum_{j=0}^{n} p_j I \lambda^j &= (\lambda I - B) \sum_{k=0}^{n-1} \lambda^k C_k \\
&= \sum_{k=0}^{n-1} \lambda^{k+1} C_k - \sum_{k=0}^{n-1} \lambda^k B C_k \\
&= \sum_{k=1}^{n} \lambda^k C_{k-1} - \sum_{k=0}^{n-1} \lambda^k B C_k \\
&= \lambda^n C_{n-1} + \sum_{k=1}^{n-1} \lambda^k \left[C_{k-1} - B C_k\right] - B C_0.
\end{aligned}$$

Comparing coefficients of λ^j then implies,

$$\begin{aligned}
p_n I &= C_{n-1}, \\
p_k I &= \left[C_{k-1} - B C_k\right] \text{ for } 1 \le k \le n-1, \\
p_0 I &= -B C_0
\end{aligned}$$

and hence

$$\begin{aligned}
B^n p_n I &= B^n C_{n-1}, \\
B^k p_k I &= B^k \left[C_{k-1} - B C_k\right] \text{ for } 1 \le k \le n-1, \\
p_0 I &= -B C_0.
\end{aligned}$$

Summing these identities then shows,

$$\begin{aligned}
p(B) = p(P) I &= B^n C_{n-1} + \sum_{k=1}^{n-1} B^k \left[C_{k-1} - B C_k\right] - B C_0 \\
&= B^n C_{n-1} + \sum_{k=1}^{n-1} B^k C_{k-1} - \sum_{k=1}^{n-1} B^{k+1} C_k - B C_0 \\
&= \sum_{k=1}^{n} B^k C_{k-1} - \sum_{k=0}^{n-1} B^{k+1} C_k = \mathbf{0}.
\end{aligned}$$

■

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 $0 \notin \mathrm{Ran}(f) = f(X)$ 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. $0 \notin \mathrm{Ran}(f)$. Conversely if $0 \notin \mathrm{Ran}(f)$, 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

$$\frac{1}{f} = \|\cdot\|_\infty - \lim_{n\to\infty} p_n(f, \bar{f}) \implies \frac{1}{f} \in C^*(f,1)$$

■

We now are going to take $X = S = \{z \in \mathbb{C} : |z| = 1\}$ in the next couple of results.

Theorem 3.12. *Let $\mathcal{B} = C(S^1; \mathbb{C})$ 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

$$\mathcal{A} = \left\{ f \in \mathcal{B} : \int_{-\pi}^{\pi} f(e^{i\theta}) e^{in\theta} d\theta = 0 \text{ for all } n \in \mathbb{N} \right\}. \tag{3.1}$$

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

$$\int_{-\pi}^{\pi} p(e^{i\theta}) e^{in\theta} d\theta = \sum_{k=0}^{n} p_k \int_{-\pi}^{\pi} e^{ik\theta} e^{in\theta} d\theta = 0 \text{ for all } n \in \mathbb{N}$$

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

$$p_k := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) e^{-ik\theta} d\theta \text{ for all } k \in \mathbb{Z}$$

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 $k \in \mathbb{N}$ becasue $f \in \mathcal{A}$. By the theory of the Fourier series (using the Féjer kernel) we know that

$$q_N(z) := \frac{1}{2N+1} \sum_{n=0}^{N} p_n(z) \to f(z) \text{ uniformly in } z,$$

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. 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 $D = D(0,1)$. ■

Remark 3.13. Notice that $\mathcal{B} = C(S^1; \mathbb{C}) = C^*(u,1)$ 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 [9], 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 **??** below.]*

Proof. We know that $\sigma_{\mathcal{B}}(u) = u(S^1) = S^1$ 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

$$p_n(z) \overset{\text{unif.}}{\longrightarrow} v_\lambda(z) = \frac{1}{z-\lambda} \text{ as } n \to \infty.$$

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*

$$\operatorname{essran}_\mu(q) = \{w \in \mathbb{C} : \mu(q^{-1}(D(w,\varepsilon))) > 0 \text{ for all } \varepsilon > 0\}.$$

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 $0 \notin \operatorname{essran}_\mu(f)$.*

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*

$$\operatorname{essran}_\mu(f) = \sigma_{L^\infty(\mu)}(f) = \sigma_{C^*(f,1)}(f). \tag{3.2}$$

Proof. We start with the proof of the first equality in Eq. (3.2). If $\lambda \notin \operatorname{essran}_\mu(f)$ iff there exists $\varepsilon > 0$ so that $\mu(\{|f-\lambda| < \varepsilon\}) = 0$. Thus if $\lambda \notin \operatorname{essran}_\mu(f)$, 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 $\lambda \notin \operatorname{essran}_\mu(f)$.

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 $\lambda \notin \sigma_{L^\infty(\mu)}(f) = \operatorname{essran}_\mu(f)$ 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 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 $B(L^p(\mu))$.

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 $\mathbf{q} = (q_1, \dots, q_n)$ 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

$$C(\operatorname{essran}_\mu(\mathbf{q})) \ni f \to M_{f\circ\mathbf{q}} \in \mathcal{B} \subset B\left(L^2(\mu)\right)$$

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, $S(\omega_1, \omega_2, \dots) = (0, \omega_1, \omega_2, \dots)$. 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 $\sigma_{ap}(S) \subset \mathbb{C} \setminus D$.

Moreover we can find eigenvectors of S^* as follows;

$$(S^* - \lambda)(\omega_1, \omega_2, \dots) = (\omega_2 - \lambda\omega_1, \omega_3 - \lambda\omega_2, \dots)$$

which is zero when $\omega_2 = \lambda\omega_1$, $\omega_3 = \lambda\omega_2 = \lambda^2\omega_1, \dots\ \omega_n = \lambda^{n-1}\omega_1$. Therefore we have produced an eigenvector, namely

$$S^*(1, \lambda, \lambda^2, \dots) = \lambda(1, \lambda, \lambda^2, \dots).$$

Since $(1, \lambda, \lambda^2, \dots) \in \ell^2$ if $|\lambda| < 1$, it follows that $D \subset \sigma_{ev}(S^*) \subset \sigma_{ap}(S^*)$ and since $\sigma_{ap}(S^*)$ is closed, $\bar{D} \subset \sigma_{ap}(S^*) \subset \sigma(S^*) \subset \bar{D}$, i.e.

$$\sigma_{ap}(S^*) = \sigma(S^*) = \bar{D} = \sigma(S).$$

We may now further conclude that $\sigma_{ap}(S) \subset \bar{D} \setminus D = S^1$ and in particular $\sigma_{ap}(S) \subsetneqq \sigma(S)$.

Notice that for $\lambda \in S^1$ that $\omega^N := (1, \lambda, \lambda^2, \dots \lambda^N, 0, 0, \dots)$ satisfy $\|\omega^N\|_{\ell^2}^2 = N+1$ while

$$S^*\omega^N - \lambda\omega^N = \lambda\omega^{N-1} - \lambda\omega^N = -\lambda^{N+1}e_{N+1}.$$

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 $\|\omega^N/\sqrt{N+1}\|_{\ell^2} = 1$, which shows directly that $S^1 \subset \sigma_{ap}(S^*)$.

Exercise 3.2. Continuing then notation used in Example 3.24, show $\sigma_{ap}(S) = S^1$.

3.1.4 Spectrum of Normal Operators

Lemma 3.25. *If H and K be Hilbert spaces and $A \in L(H,K)$, then;*

1. $\mathrm{Nul}(A^*) = \mathrm{Ran}(A)^{\perp}$, *and*
2. $\overline{\mathrm{Ran}(A)} = \mathrm{Nul}(A^*)^{\perp}$,
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 $y \in \mathrm{Nul}(A^*) \iff A^*y = 0 \iff \langle y, Ah\rangle = \langle 0, h\rangle = 0$ for all $h \in H \iff y \in \mathrm{Ran}(A)^{\perp}$.

2. By Exercise **??**, $\overline{\mathrm{Ran}(A)} = \mathrm{Ran}(A)^{\perp\perp}$, and so $\overline{\mathrm{Ran}(A)} = \mathrm{Ran}(A)^{\perp\perp} = \mathrm{Nul}(A^*)^{\perp}$.

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 separable 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. ∎

Lemma 3.30. *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,

$$\begin{aligned}\|(B + iC)\psi\|^2 &= \|B\psi\|^2 + \|C\psi\|^2 + 2\operatorname{Re}\langle B\psi, iC\psi\rangle \\ &= \|B\psi\|^2 + \|C\psi\|^2 + 2\operatorname{Im}\langle CB\psi, \psi\rangle \\ &= \|B\psi\|^2 + \|C\psi\|^2\end{aligned}$$

where the last equality follows from Lemma 3.27 because,

$$(CB)^* = B^*C^* = BC = CB.$$

∎

Remark 3.31. 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 3.30,

$$\|A\psi\|^2 = \frac{1}{4}\left[\|(A + A^*)\psi\|^2 + \|(A - A^*)\psi\|^2\right]$$

which is symmetric under the interchange of A with A^*.

Corollary 3.32. *If $A \in B(H)$ is a normal operator and $\lambda \in \sigma_{ev}(A)$ then;*

1. $\operatorname{Nul}(A) = \operatorname{Nul}(A^*)$ *and*
2. $\operatorname{Nul}(A - \lambda) = \operatorname{Nul}(A^* - \bar{\lambda})$, *i.e. $Au = \lambda u$ iff $A^*u = \bar{\lambda}u$.*

Lemma 3.33. *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 $\lambda \notin \sigma_{ap}(M)$, 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,

$$\operatorname{Nul}(A - \lambda I) = \{0\} = \operatorname{Nul}\left((A - \lambda I)^*\right).$$

By Corollary 2.10, $\operatorname{Ran}(A - \lambda I)$ is closed. Using these comments along with Lemma 3.25 allows us to conclude,

$$\operatorname{Ran}(A - \lambda I) = \overline{\operatorname{Ran}(A - \lambda I)} = \operatorname{Nul}\left((A - \lambda I)^*\right)^\perp = \{0\}^\perp = H$$

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). If $\lambda \notin \sigma(A)$ then $A - \lambda I$ and $(A - \lambda I)^*$ are both invertible and hence so is $(A - \lambda)^*(A - \lambda)$, i.e. $0 \notin \sigma\left((A - \lambda)^*(A - \lambda)\right)$. Conversely if $0 \notin \sigma\left((A - \lambda)^*(A - \lambda)\right)$, then $T = [(A - \lambda)^*(A - \lambda)]^{-1}$ exists. With this notation and using the fact that A is normal gives,

$$\begin{gathered} I = T(A - \lambda)^*(A - \lambda) \text{ and} \\ (A - \lambda)(A - \lambda)^*T = (A - \lambda)^*(A - \lambda)T = I.\end{gathered}$$

These equations show that $A - \lambda I$ has both a left and a right inverse and hence $(A - \lambda)$ is invertible, i.e. $\lambda \notin \sigma(A)$. ∎

Example 3.34. 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 $0 \notin \sigma(S^*S)$.

Note: if $A, B \in B(H)$ are commuting operators so that AB is invertible then both A and B are invertible. Indeed,

$$\left[(AB)^{-1} A\right] B = (AB)^{-1} AB = I$$

and

$$B\left[A(AB)^{-1}\right] = AB(AB)^{-1} = I$$

which shows B has both a left and a right inverse and hence is invertible. This example shows that we can not drop the assumption that $[A, B] = 0$ in this last assertion.

Lemma 3.35. *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 3.30.] Equation (3.7) along with Lemma 3.33 shows that $\lambda \notin \sigma(A)$ if $\beta \neq 0$, i.e. $\sigma(A) \subset \mathbb{R}$. ∎

Remark 3.36. 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 3.37. *The **resolvent** (operators) of a is the function,*

$$\rho(a) \ni \lambda \to R_\lambda = (a-\lambda)^{-1} \in \mathcal{A}_{inv}.$$

Lemma 3.38 (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 3.39. *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 3.40. *Let $\mathcal{A}$ be a complex Banach algebra with unit. Then $\sigma(x) \neq \emptyset$ for every $x \in \mathcal{A}$.*

Proof. Suppose $\sigma(x)$ is empty. Then for any $\xi \in \mathcal{A}^*$, $\lambda \to \xi((x-\lambda)^{-1})$ is an entire function which vanishes as $\lambda \to \infty$. So by Liouville's theorem, $\xi[(x-\lambda)^{-1}] \equiv 0$ for all $\xi \in \mathcal{A}^*$ and $\lambda \in \mathbb{C}$. Taking $\lambda = 0$, we conclude that $x^{-1} = 0$ which is impossible. [Again we could avoid applying linear functionals simply by making use of the Cauchy integral formula (see Theorem 1.10) to reprove Liouville's theorem in the Banach space context.] ∎

Remark 3.41. 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 $i = 1, 2, \ldots, n$ 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.

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. Given $z_0 \in \mathbb{C}$, factor $p(\lambda) - z_0$ as

$$p(\lambda) - z_0 = \alpha(\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$$

where $\alpha \in \mathbb{C}^\times$ and $\{\lambda_i\}_{i=1}^n \subset \mathbb{C}$ are the solutions (with multiplicity) to $p(\lambda) = z_0$. Since

$$p(a) - z_0 = \alpha(a - \lambda_1)\cdots(a - \lambda_n)$$

we may conclude using Remark 3.41 that $z_0 \in \sigma(p(a))$ iff $\lambda_i \in \sigma(a)$ for some i, i.e. iff $z_0 = p(\lambda)$ for some $\lambda \in \sigma(a)$, i.e. iff $z_0 \in p(\sigma(a))$. ■

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 3.39, 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,

$$r(a)^n = r(a^n) \le \|a^n\| \implies r(a) \le \|a^n\|^{1/n}.$$

Passing to the limit as $n \to \infty$ in this inequality shows

$$r(a) \le \liminf_{n\to\infty} \|a^n\|^{1/n}. \tag{3.11}$$

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,

$$|\lambda| \limsup_{n\to\infty} \|a^n\|^{1/n} = \limsup_{n\to\infty} \|(\lambda a)^n\|^{1/n} \le 1 \ \forall\ |\lambda| < 1/r(a)$$

and hence $\limsup_{n\to\infty} \|a^n\|^{1/n} \le r(a)$ which along with Eq. (3.11) completes the proof. ■

Theorem 3.46 (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 $\mathcal{A} = \mathbb{C} \cdot 1_{\mathcal{A}}$.*

[1] Recall that $\mathcal{A}$ is a division algebra iff every non-zero element is invertible.

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. $\mathcal{A} = \mathbb{C} \cdot 1$. ■

Exercise 3.3 (Compare with Proposition ??). Let $\mathcal{B}$ be a complex Banach algebra with unit, then for any $a, b \in \mathcal{B}$ which commute, show;

1. $r(ab) \le r(a) r(b)$ and
2. $r(a+b) \le r(a) + r(b)$.

Proposition 3.47 (Optional). *If $a \in \mathcal{A}$ and $\lambda \in \rho(a)$, then*

$$\left\| (a-\lambda)^{-1} \right\| \ge r\left((a-\lambda)^{-1} \right) \ge \frac{1}{\operatorname{dist}(\lambda, \sigma(a))}.$$

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

$$\operatorname{dist}(\lambda, \sigma(a)) \ge \frac{1}{r\left((a-\lambda)^{-1} \right)} \iff r\left((a-\lambda)^{-1} \right) \ge \frac{1}{\operatorname{dist}(\lambda, \sigma(a))}.$$

■

Functional Calculus I

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 $f(T)$ for almost any $T \in \mathcal{B}$ provided that f is analytic on an open neighborhood of the spectrum of T. 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 Functional Calculus

Let $\mathcal{B}$ be a unital Banach algebra and $T \in \mathcal{B}$. Suppose that $\sigma(T)$ 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.

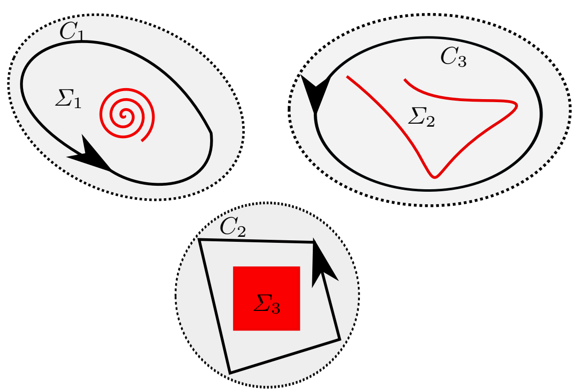

Fig. 4.1. The spectrum of T is in red, the counter clockwise contours are in black, and Ω is the union of the grey sets.

Given a holomorphic function, f, on Ω we let

$$f(T) := \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-T}dz := \sum_{k=1}^n \frac{1}{2\pi i}\oint_{C_k} \frac{f(z)}{z-T}dz,$$

where $\frac{1}{z-T} := (z-T)^{-1}$ and $C = \cup_{k=1}^n C_k$.

Let us observe that $f(T)$ 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

$$\ell(f(T)) = \frac{1}{2\pi i}\oint_C f(z)\,\ell\left((z-T)^{-1}\right)dz$$

where $f(z)\,\ell\left((z-T)^{-1}\right)$ is a holomorphic function on $\Omega \setminus \sigma(T)$. Therefore $\frac{1}{2\pi i}\oint_C f(z)\,\ell\left((z-T)^{-1}\right)dz$ remains constant over deformations of C which remain in $\Omega \setminus \sigma(T)$. As ℓ is arbitrary it follows that $\frac{1}{2\pi i}\oint_C \frac{f(z)}{z-T}dz$ remains constant over such deformations as well.

Theorem 4.1. *The map $H(\Omega) \ni f \to f(T) \in B(X)$ is an algebra homomorphism satisfying the consistency criteria; if $f(z) = \sum_{m=0}^N a_m z^m$ is a polynomial then*

$$f(T) = \sum_{m=0}^N a_m T^m.$$

More generally, $\rho > 0$ is chosen so that $\|T\| < \rho$ and $f \in H(D(0,\rho))$, then

$$f(T) = \sum_{m=0}^\infty \frac{f^m(0)}{m!}T^m.$$

Proof. It is clear that $H(\Omega) \ni f \to f(T) \in B(X)$ 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

$$\begin{aligned} f(T)\,g(T) &= \left(\frac{1}{2\pi i}\right)^2 \sum_{k,l=1}^n \oint_{C_k} \frac{f(z)}{z-T}dz \oint_{\tilde{C}_l} \frac{g(\zeta)}{\zeta-T}d\zeta \\ &= \left(\frac{1}{2\pi i}\right)^2 \sum_{k,l=1}^n \oint_{C_k} dz \oint_{\tilde{C}_l} d\zeta \frac{f(z)}{z-T}\frac{g(\zeta)}{\zeta-T}. \\ &=: \left(\frac{1}{2\pi i}\right)^2 \sum_{k,l=1}^n A_{kl} \end{aligned}$$

Using the resolvent formula,

$$\frac{1}{z-T}-\frac{1}{\zeta-T}=\frac{\zeta-z}{(z-T)(\zeta-T)},$$

we find

$$\begin{aligned}A_{kl}&:=\oint_{C_k}dz\oint_{\tilde{C}_l}d\zeta\frac{f(z)}{z-T}\frac{g(\zeta)}{\zeta-T}\\&=\oint_{C_k}dz\oint_{\tilde{C}_l}d\zeta f(z)\,g(\zeta)\,\frac{1}{\zeta-z}\left(\frac{1}{z-T}-\frac{1}{\zeta-T}\right)\end{aligned}$$

If $k\neq l$ then for $z\in C_k$, $\zeta\to g(\zeta)\frac{1}{\zeta-z}\frac{1}{z-T}$ is analytic for ζ inside $\tilde{C}_l$ and therefore

$$\oint_{\tilde{C}_l}d\zeta g(\zeta)\,\frac{1}{\zeta-z}\frac{1}{z-T}=0. \tag{4.1}$$

Similarly if $\zeta\in\tilde{C}_l$, then $z\to f(z)\frac{1}{\zeta-z}\frac{1}{\zeta-T}$ is analytic inside of C_k and therefore

$$\oint_{C_k}dz f(z)\,\frac{1}{\zeta-z}\frac{1}{z-T}=0.$$

From these last two identities and Fubini's theorem it follows that $A_{k,l}=0$ if $k\neq l$.

Now suppose that $k=l$ so that

$$A_{k,k}=\oint_{C_k}dz\oint_{\tilde{C}_k}d\zeta f(z)\,g(\zeta)\,\frac{1}{\zeta-z}\left(\frac{1}{z-T}-\frac{1}{\zeta-T}\right).$$

For $z\in C_k$, $\zeta\to g(\zeta)\frac{1}{\zeta-z}\frac{1}{z-T}$ is still analytic for ζ inside $\tilde{C}_l$ and therefore Eq. (4.1) holds for $l=k$ and

$$A_{k,k}=\oint_{\tilde{C}_k}d\zeta\frac{g(\zeta)}{\zeta-T}\oint_{C_k}dz\frac{f(z)}{z-\zeta}=2\pi i\oint_{\tilde{C}_k}d\zeta\frac{g(\zeta)\,f(\zeta)}{\zeta-T}.$$

Thus we have shown

$$\begin{aligned}f(T)\,g(T)&=\left(\frac{1}{2\pi i}\right)^2\sum_{k=1}^{n}A_{k,k}\\&=\frac{1}{2\pi i}\sum_{k=1}^{n}\oint_{\tilde{C}_k}d\zeta\frac{g(\zeta)\,f(\zeta)}{\zeta-T}=(f\cdot g)(T)\end{aligned}$$

which shows that $T\to f(T)$ is an algebra homomorphism

If $f(z)$ is an entire function of z we may replace the contour C by $z=\rho e^{i\theta}$ for any ρ sufficiently large so that $\sigma(T)\subset D(0,\rho)$. Since $|z|=\rho$ with $\rho>\|T\|$ we have

$$\frac{1}{z-T}=\frac{1}{z}\frac{1}{1-T/z}=\frac{1}{z}\sum_{n=0}^{\infty}\frac{1}{z^n}T^n=\sum_{n=0}^{\infty}\frac{1}{z^{n+1}}T^n,$$

it now follows that

$$\begin{aligned}f(T)&=\frac{1}{2\pi i}\oint_C\frac{f(z)}{z-T}dz=\frac{1}{2\pi i}\oint_{|z|=\rho}\frac{f(z)}{z-T}dz\\&=\frac{1}{2\pi i}\oint_{|z|=\rho}\sum_{m=0}^{\infty}\frac{f(z)}{z^{m+1}}T^m dz\\&=\sum_{m=0}^{\infty}\left(\frac{1}{2\pi i}\oint_{|z|=\rho}\frac{f(z)}{z^{m+1}}dz\right)T^m=\sum_{m=0}^{\infty}\frac{f^{(m)}(0)}{m!}T^m.\end{aligned}$$

■

Theorem 4.2 (Spectral Mapping Theorem). *Keeping the same notation as above,* $f(\sigma(T))=\sigma(f(T))$.

Proof. Suppose that $\mu\in\sigma(T)$ 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 $f(T)-f(\mu)=(T-\mu)\,g(T)$ and so if $f(\mu)\notin\sigma(f(T))$ then $f(T)-f(\mu)$ is invertible and therefore $T-\mu$ would be invertible contradicting $\mu\in\sigma(T)$. Thus we have shown $f(\sigma(T))\subset\sigma(f(T))$. Conversely if $\alpha\notin f(\sigma(T))$ then $g(z):=\frac{1}{f(z)-\alpha}$ is holomorphic on a neighborhood of $\sigma(T)$. Since $(f(z)-\alpha)\,g(z)=1$ it follows that $(f(T)-\alpha)\,g(T)=I$ and therefore $\alpha\notin\sigma(f(T))$ and we have shown $[f(\sigma(T))]^c\subset[\sigma(f(T))]^c$, i.e. $\sigma(f(T))\subset\sigma(f(T))$. ■

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* **??** *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

$$\left\|a^{2^n}\right\|=\|a\|^{2^n}. \tag{4.2}$$

We will prove Eq. (4.2) 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. (4.2) 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. (4.2) 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 $i \le j$, 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 $\sigma(a) \subset \mathbb{R}$. [See Lemma* **??** *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[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 $r(x) \le \|x\|$ for all $x \in \mathcal{B}$ along with the C^*-identity shows,

$$(\lambda+1)^2 \le [r(\lambda - ia)]^2 \le \|\lambda - ia\|^2$$

wherein

$$\|\lambda - ia\|^2 \overset{C^*\text{-cond}}{=} \left\|(\lambda - ia)^*(\lambda - ia)\right\| = \|(\lambda + ia)(\lambda - ia)\|$$
$$= \left\|\lambda^2 + a^2\right\| \le \lambda^2 + \left\|a^2\right\| \overset{C^*\text{-cond}}{=} \lambda^2 + \|a\|^2.$$

Combining the last two displayed equation leads to the nonsensical inequality, $2\lambda + 1 \le \|a\|^2$ for all $\lambda \in \mathbb{R}$, and we have arrived at the desired contradiction and hence $i \notin \sigma(a)$.

[1] More directly,

$$\lambda + 1 - (\lambda - ia) = 1 + ia = i(a - i)$$

is not invertible by assumption and hence $\lambda + 1 \in (\lambda - ia)$.

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 $\sigma(a) \subset \mathbb{R}$. ■

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 $\varphi_a(x) = a$ ore equivalently, $\varphi_a(p) = p(a)$ for all $p \in \mathbb{C}[x]$. [We usually write $\varphi_a(f)$ as $f(a)$.]*

Proof. By the classical Stone-Weiersrtass theorem, $\left\{p|_{\sigma(a)} : p \in \mathbb{C}[x]\right\}$ 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 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 4.8 (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 4.9. *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 4.10. *Let H be a separable Hilbert space and $\mathcal{A}$ be a unital $*$ – sub-algebra 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 4.1. Show that Theorem 4.10 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 $\{(v_\alpha, H_\alpha)\}_{\alpha \in I}$ where H_α is a closed $\mathcal{A}$ in variant subspace of H, $v_\alpha \neq 0$ is an $\mathcal{A}$-cyclic vector, and $H = \oplus_{\alpha \in I} H_\alpha$ (orthogonal direct sum).

Before leaving this topic let us explore the meaning of cyclic vectors by looking at the finite dimensional case.

Proposition 4.11. *Let T be a $n \times n$-diagonal matrix, $T = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ for some $\lambda_i \in \mathbb{C}$ and set $\sigma(T) := \{\lambda_1, \ldots, \lambda_n\}$. If $u \in \mathbb{C}^n$ is expressed as*

$$u = \sum_{\lambda \in \sigma(T)} e_\lambda \tag{4.3}$$

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. (4.3) 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

Proposition 4.12. *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. Then for each $v \in H \setminus \{0\}$, there exists a unique finite radon measure, μ_v, on $(Y, \mathcal{B}_Y)$ such that*

$$\langle \varphi(f)v, v\rangle = \int_Y f d\mu_v \ \forall\ f \in C(Y). \tag{4.4}$$

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 finite radon measure, μ_v, on $(Y, \mathcal{B}_Y)$ such that

$$\langle \varphi(f) v, v\rangle = \Lambda(f) = \int_Y f d\mu_v \ \forall\ f \in C(Y).$$

■

Proposition 4.13. *Continue the notation and assumptions in Proposition 4.12 and for each $v \in H \setminus \{0\}$, let*

$$H_v := \overline{\mathcal{B}v}^H \subset H. \tag{4.5}$$

Then there exists a unique unitary isomorphism, $U_v : L^2(\mu_v) \to H_v$ which is uniquely determined by requiring

$$U_v f = \varphi(f) v \in H_v \text{ for all } f \in C(Y). \tag{4.6}$$

Moreover, this unitary map satisfies,

$$U_v^* \varphi(f)|_{H_v} U_v = M_f \text{ on } L^2(\mu_v) \quad \forall\ f \in C(Y). \tag{4.7}$$

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 $f, g \in C(Y)$,

$$U_v^* \varphi(f) U_v g = U_v^* \varphi(f) \varphi(g) v = U_v^* \varphi(fg) v = fg = M_f g. \tag{4.8}$$

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. (4.8) 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. (4.7). ■

Theorem 4.14. *Continue the notation and assumptions in Proposition 4.12. Then there exist $N \in \mathbb{N} \cup \{\infty\}$, a probability measure μ measure on $\Omega := \Lambda_N \times Y \subset \Lambda_N \times \mathbb{R}$ equipped with the product σ – algebra (here $\Lambda_N = \{1, 2, \ldots, N\} \cap \mathbb{N}$), and a unitary map $U : L^2(\mu) \to H$ such that*

$$U^* \varphi(f) U = M_{f \circ \pi} \text{ on } L^2(\mu)$$

where $\pi : \Omega \to \mathbb{C}$ is defined by $\pi(j, w) = w$ for all $j \in \Lambda_N$ and $w \in Y$.

Proof. By Theorem 4.10, 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 $\mathcal{A}$. 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 4.12 and let $\Omega := \Lambda \times Y$ which we equip with the product $\sigma -$ algebra, $\mathcal{F}$, and the probability measure μ defined as follows. Every $G \in \mathcal{F}$ may be written (see Remark 4.15 below) may be uniquely written as

$$G = \sum_{i \in \Lambda} \{i\} \times G_i \text{ for some } \{G_i\}_{i \in \Lambda} \subset \mathcal{B}_Y$$

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,

$$Ug = \sum_{i \in \Lambda} U_{v_i} g(i, \cdot) \oplus_{i \in \Lambda} H_i = H,$$

where U_{v_i} is the unitary map in Proposition 4.13. 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

$$\begin{aligned} U M_{f \circ \pi} g &= \sum_{i \in \Lambda} U_{v_i} [f \circ \pi(i, \cdot) g(i, \cdot)] = \sum_{i \in \Lambda} U_{v_i} [f g(i, \cdot)] \\ &= \sum_{i \in \Lambda} U_{v_i} [M_f g(i, \cdot)] = \sum_{i \in \Lambda} \varphi(f) U_{v_i} g(i, \cdot) = \varphi(f) U g. \end{aligned}$$

This completes the proof. ■

Remark 4.15. The product σ-algebra on $\Lambda \times Y$ is given by the collection of sets

$$\mathcal{F} := \left\{ \sum_{j\in\Lambda} (\{j\} \times G_j) : \ \{G_j\}_{j=1}^{\infty} \subset \mathcal{B}_Y \right\}.$$

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,

$$\begin{aligned}\mu\left(\sum_{n\in\mathbb{N}} G(n)\right) &= \sum_{i\in\Lambda} \mu_i \left(\sum_{n\in\mathbb{N}} G_i(n)\right) = \sum_{i\in\Lambda}\sum_{n\in\mathbb{N}} \mu_i(G_i(n)) \\ &= \sum_{n\in\mathbb{N}}\sum_{i\in\Lambda} \mu_i(G_i(n)) = \sum_{n\in\mathbb{N}} \mu(G(n)).\end{aligned}$$

Corollary 4.16 (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 $A = UM_aU^*$.*

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 4.14 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 4.14. ∎

4.5 Measurable Functional Calculus

For general measurable functional calculus, see Theorem **??**.

Theorem 4.17 (Measurable Functional Calculus). *Let H be a separable Hilbert space and A be a self-adjoint element of $B(H)$. Then there exists a unique map $\psi_A : \mathcal{B}^\infty(\sigma(A)) \to B(H)$ 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 $f, g \in \mathcal{B}^\infty(\sigma(A))$.*
2. *$\|\psi_A(f)\|_{op} \le \|f\|_\infty$ for all $f \in \mathcal{B}^\infty(\sigma(A))$.*
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 $f_n \in \mathcal{B}^\infty(\sigma(A))$ 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 $Ah = \lambda h$ for some $h \in H$ and $\lambda \in \mathbb{R}$, then $\psi_A(f)h = f(\lambda)h$.*
7. *If $B \in B(H)$ and $[B, A] = 0$, then $[B, \psi_A(f)] = 0$ for all $f \in \mathcal{B}^\infty(\sigma(A))$.*

Proof. Uniqueness. Suppose that $\psi : \mathcal{B}^\infty(\sigma(A)) \to B(H)$ is another map satisfying (1) – (4). Let

$$\mathbb{H} := \{f \in \mathcal{B}^\infty(\sigma(A), \mathbb{C}) : \psi(f) = \psi_A(f)\}.$$

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 $\mathbb{H} = \mathcal{B}^\infty(\sigma(A))$, i.e. $\psi = \psi_A$.

Existence. Let $U : L^2(\Omega, \mu) \to H$ be as in Corollary 4.16 and then define

$$\psi_A(f) := UM_{f\circ a}U^* \ \forall \ f \in \mathcal{B}^\infty(\sigma(A)).$$

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. 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 $\{g \ne 0\}$. Thus it follows that $f \circ a = f(\lambda) \ \mu$ – a.e. on $\{g \ne 0\}$ and this implies $M_{f\circ a} g = f(\lambda) g$ which then implies,

$$\psi_A(f)h = \psi_A(f)Ug = UM_{f\circ a}g = Uf(\lambda)\,g = f(\lambda)\,h.$$

7. Let

$$\mathbb{H} := \{f \in \mathcal{B}^\infty(\sigma(A), \mathbb{C}) : [B, \psi_A(f)] = 0\}$$

which is vector space closed under conjugation[2] 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 follows by an application of the multiplicative system Theorem A.9 applied using the multiplicative system,

$$\mathbb{M} = \left\{p|_{\sigma(A)} : p \in \mathbb{C}[x]\right\}.$$

■

[2] Again we use Theorem 2.68 and the fact that $\psi_A(f)$ is normal for all f.

Part III

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. A. Ancona and J. C. Taylor, *Some remarks on Widder's theorem and uniqueness of isolated singularities for parabolic equations*, Partial differential equations with minimal smoothness and applications (Chicago, IL, 1990), IMA Vol. Math. Appl., vol. 42, Springer, New York, 1992, pp. 15–23. MR MR1155849
3. T. Ando and F. Hiai, *Inequality between powers of positive semidefinite matrices*, Linear Algebra Appl. **208/209** (1994), 65–71. MR 1287340 (95k:15023)
4. 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
5. J. Brüning and M. Lesch, *Hilbert complexes*, J. Funct. Anal. **108** (1992), no. 1, 88–132. MR MR1174159 (93k:58208)
6. 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)
7. Paul R. Chernoff, *Essential self-adjointness of powers of generators of hyperbolic equations*, J. Functional Analysis **12** (1973), 401–414. MR 0369890 (51 #6119)
8. Henry B. Cohen, *Sequential denseness and the Eberlein-Šmulian theorem*, Math. Ann. **172** (1967), 209–210. MR MR0212549 (35 #3420)
9. John B. Conway, *A course in functional analysis*, second ed., Graduate Texts in Mathematics, vol. 96, Springer-Verlag, New York, 1990. MR 1070713
10. Claude Dellacherie, *Capacités et processus stochastiques*, Springer-Verlag, Berlin, 1972, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 67. MR MR0448504 (56 #6810)
11. Robert Doran, *Characterizations of c* algebras: the gelfand naimark theorems*, CRC Press, 05 2019.
12. 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)
13. Gerald B. Folland, *Harmonic analysis in phase space*, Annals of Mathematics Studies, vol. 122, Princeton University Press, Princeton, NJ, 1989. MR 983366 (92k:22017)
14. 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)
15. Masatoshi Fukushima, Yōichi Ōshima, and Masayoshi Takeda, *Dirichlet forms and symmetric Markov processes*, de Gruyter Studies in Mathematics, vol. 19, Walter de Gruyter & Co., Berlin, 1994. MR MR1303354 (96f:60126)
16. 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)
17. ———, *The conservation property of the heat equation on Riemannian manifolds.*, Comm. Pure Appl. Math. **12** (1959), 1–11. MR MR0102097 (21 #892)
18. Robert F. Geitz, *Pettis integration*, Proc. Amer. Math. Soc. **82** (1981), no. 1, 81–86. MR 603606
19. ———, *Geometry and the Pettis integral*, Trans. Amer. Math. Soc. **269** (1982), no. 2, 535–548. MR 637707

20. 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
21. Robert Frederick Geitz, *THE PETTIS INTEGRAL*, ProQuest LLC, Ann Arbor, MI, 1980, Thesis (Ph.D.)–University of Illinois at Urbana-Champaign. MR 2630764
22. A. A. Grigor′yan, *Stochastically complete manifolds*, Dokl. Akad. Nauk SSSR **290** (1986), no. 3, 534–537. MR MR860324 (88a:58209)
23. Mikhael Gromov, *Carnot-Carathéodory spaces seen from within*, Sub-Riemannian geometry, Progr. Math., vol. 144, Birkhäuser, Basel, 1996, pp. 79–323. MR MR1421823 (2000f:53034)
24. Brian C. Hall, *Phase space bounds for quantum mechanics on a compact Lie group*, Comm. Math. Phys. **184** (1997), no. 1, 233–250.
25. ———, *Quantum mechanics in phase space*, Perspectives on quantization (South Hadley, MA, 1996), Amer. Math. Soc., Providence, RI, 1998, pp. 47–62. MR 99e:22015
26. ———, *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
27. Erhard Heinz, *Beiträge zur Störungstheorie der Spektralzerlegung*, Math. Ann. **123** (1951), 415–438. MR 0044747 (13,471f)
28. H. Hess, R. Schrader, and D. A. Uhlenbrock, *Domination of semigroups and generalization of Kato's inequality*, Duke Math. J. **44** (1977), no. 4, 893–904. MR MR0458243 (56 #16446)
29. Joe Howard, *Mackey compactness in Banach spaces*, Proc. Amer. Math. Soc. **37** (1973), 108–110. MR MR0312212 (47 #774)
30. Svante Janson, *Gaussian Hilbert spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997. MR MR1474726 (99f:60082)
31. David Jerison, *The Poincaré inequality for vector fields satisfying Hörmander's condition*, Duke Math. J. **53** (1986), no. 2, 503–523. MR MR850547 (87i:35027)
32. David Jerison and Antonio Sánchez-Calle, *Subelliptic, second order differential operators*, Complex analysis, III (College Park, Md., 1985–86), Lecture Notes in Math., vol. 1277, Springer, Berlin, 1987, pp. 46–77. MR MR922334 (89b:35021)
33. Tosio Kato, *Notes on some inequalities for linear operators*, Math. Ann. **125** (1952), 208–212. MR 0053390 (14,766e)
34. ———, *Schrödinger operators with singular potentials*, Proceedings of the International Symposium on Partial Differential Equations and the Geometry of Normed Linear Spaces (Jerusalem, 1972), vol. 13, 1972, pp. 135–148 (1973). MR MR0333833 (48 #12155)
35. Xue-Mei Li, *Stochastic differential equations on noncompact manifolds: moment stability and its topological consequences*, Probab. Theory Related Fields **100** (1994), no. 4, 417–428. MR MR1305780 (95k:58176)
36. ———, *Strong p-completeness of stochastic differential equations and the existence of smooth flows on noncompact manifolds*, Probab. Theory Related Fields **100** (1994), no. 4, 485–511. MR MR1305784 (96b:58119)
37. Karl Löwner, *Über monotone Matrixfunktionen*, Math. Z. **38** (1934), no. 1, 177–216. MR 1545446
38. Laurent W. Marcoux, *An introduction to operator algebras*, University of Waterloo preprint book, 2016.
39. Richard Montgomery, *A tour of subriemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, vol. 91, American Mathematical Society, Providence, RI, 2002. MR MR1867362 (2002m:53045)
40. Alexander Nagel, Elias M. Stein, and Stephen Wainger, *Balls and metrics defined by vector fields. I. Basic properties*, Acta Math. **155** (1985), no. 1-2, 103–147. MR MR793239 (86k:46049)
41. 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
42. ———, *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
43. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983. MR 710486
44. 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
45. Gert K. Pedersen, *Some operator monotone functions*, Proc. Amer. Math. Soc. **36** (1972), 309–310. MR 0306957 (46 #6078)
46. A. Prasad, *An easy proof of the Stone-von Neumann-Mackey theorem*, ArXiv e-prints (2009), 1–10.
47. Ian F. Putnam, *Lecture notes on c-star-algebras*, http://www.math.uvic.ca/faculty/putnam/ln/C*-algebras.pdf, 2019.
48. 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
49. ———, *Methods of modern mathematical physics. III*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1979, Scattering theory. MR 80m:81085
50. ———, *Methods of modern mathematical physics. I*, second ed., Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980, Functional analysis. MR 85e:46002
51. ———, *Methods of modern mathematical physics. I*, second ed., Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980, Functional analysis. MR MR751959 (85e:46002)
52. Franz Rellich, *Halbbeschränkte gewöhnliche Differentialoperatoren zweiter Ordnung*, Math. Ann. **122** (1951), 343–368. MR 0043316 (13,240g)
53. Derek W. Robinson, *Elliptic operators and Lie groups*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1991, Oxford Science Publications. MR MR1144020 (92m:58133)
54. Walter Roelcke, *Über den Laplace-Operator auf Riemannschen Mannigfaltigkeiten mit diskontinuierlichen Gruppen*, Math. Nachr. **21** (1960), 131–149. MR MR0151927 (27 #1908)

55. 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)
56. 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
57. ———, *Real and complex analysis*, third ed., McGraw-Hill Book Co., New York, 1987. MR 924157
58. ———, *Functional analysis*, second ed., International Series in Pure and Applied Mathematics, McGraw-Hill, Inc., New York, 1991. MR 1157815
59. L. Saloff-Coste, *Parabolic Harnack inequality for divergence-form second-order differential operators*, Potential Anal. **4** (1995), no. 4, 429–467, Potential theory and degenerate partial differential operators (Parma). MR MR1354894 (96m:35031)
60. Laurent Saloff-Coste, *Aspects of Sobolev-type inequalities*, London Mathematical Society Lecture Note Series, vol. 289, Cambridge University Press, Cambridge, 2002. MR MR1872526 (2003c:46048)
61. 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
62. Konrad Schmüdgen, *Unbounded self-adjoint operators on Hilbert space*, Graduate Texts in Mathematics, vol. 265, Springer, Dordrecht, 2012. MR 2953553
63. B. Simon, *Convergence in trace ideals*, Proc. Amer. Math. Soc. **83** (1981), no. 1, 39–43. MR 82h:47042
64. Barry Simon, *Distributions and their Hermite expansions*, J. Mathematical Phys. **12** (1971), 140–148. MR 0278067 (43 #3799)
65. ———, *An abstract Kato's inequality for generators of positivity preserving semigroups*, Indiana Univ. Math. J. **26** (1977), no. 6, 1067–1073. MR MR0461209 (57 #1194)
66. ———, *Kato's inequality and the comparison of semigroups*, J. Funct. Anal. **32** (1979), no. 1, 97–101. MR MR533221 (80e:47036)
67. ———, *Trace ideals and their applications*, London Mathematical Society Lecture Note Series, vol. 35, Cambridge University Press, Cambridge, 1979. MR 80k:47048
68. 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
69. Robert S. Strichartz, *Analysis of the Laplacian on the complete Riemannian manifold*, J. Funct. Anal. **52** (1983), no. 1, 48–79. MR 84m:58138
70. K. T. Sturm, *Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality*, J. Math. Pures Appl. (9) **75** (1996), no. 3, 273–297. MR MR1387522 (97k:31010)
71. Karl-Theodor Sturm, *Analysis on local Dirichlet spaces. I. Recurrence, conservativeness and L^p-Liouville properties*, J. Reine Angew. Math. **456** (1994), 173–196. MR MR1301456 (95i:31003)
72. ———, *On the geometry defined by Dirichlet forms*, Seminar on Stochastic Analysis, Random Fields and Applications (Ascona, 1993), Progr. Probab., vol. 36, Birkhäuser, Basel, 1995, pp. 231–242. MR MR1360279 (96j:31009)
73. ———, *Is a diffusion process determined by its intrinsic metric?*, Chaos Solitons Fractals **8** (1997), no. 11, 1855–1860. MR MR1477263 (98h:58203)
74. Michel Talagrand, *Pettis integral and measure theory*, Mem. Amer. Math. Soc. **51** (1984), no. 307, ix+224. MR 756174
75. J. v. Neumann, *Die Eindeutigkeit der Schrödingerschen Operatoren*, Math. Ann. **104** (1931), no. 1, 570–578. MR 1512685
76. 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
77. N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon, *Analysis and geometry on groups*, Cambridge Tracts in Mathematics, vol. 100, Cambridge University Press, Cambridge, 1992. MR 95f:43008
78. Robert Whitley, *An elementary proof of the Eberlein-Šmulian theorem*, Math. Ann. **172** (1967), 116–118. MR MR0212548 (35 #3419)
79. Dana P. Williams, *A (very) short course on c-star algebras*, https://math.dartmouth.edu/ dana/bookspapers/cstar.pdf, 2019.
80. 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 [10, p. 14]. The style of proof given here may be found in Janson [30, 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 $\sigma(\mathbb{M}) \subset \mathcal{B}$.

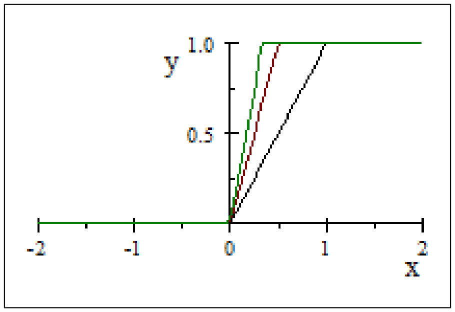

Fig. A.1. Plots of φ_1, φ_2 and φ_3 which are continuous functions used to approximate, $x \to 1_{x\ge0}$.

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 $\mathbb{M}_0 = \operatorname{span}_{\mathbb{C}}(\mathbb{M} \cup \{1\})$ 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 $\operatorname{Re} f = \frac{1}{2}(f + \bar{f})$ and $\operatorname{Im} f = \frac{1}{2i}(f - \bar{f})$ 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 $f = \operatorname{Re} f + i \operatorname{Im} f \in \mathbb{H}$ since $\operatorname{Re} f$ and $\operatorname{Im} f$ 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

$$f_n(x) := \begin{cases} 0 & \text{on } (-\infty, a] \cup [b + \frac{1}{n}, \infty) \\ n(x-a) & \text{if} \quad a \le x \le a + \frac{1}{n} \\ 1 & \text{if} \quad a + \frac{1}{n} \le x \le b \\ 1 - n(b-x) & \text{if} \quad b \le x \le b + \frac{1}{n} \end{cases}.$$

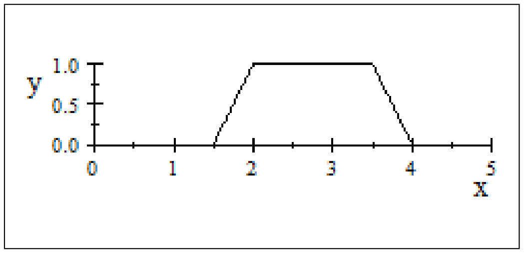

Fig. A.2. Here is a plot of $f_2(x)$ when $a = 1.5$ and $b = 3.5$.

■

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. $\mathcal{M}_2 := \sigma\left(x \to f_1(x_1) \dots f_d(x_d) : f_i \in C_c(\mathbb{R})\right)$
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. $\mathbb{M} := C_c(\mathbb{R}^d)$, *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.
$$\begin{aligned}\|A\| &= \sup_{\|x\|=1}\|Ax\| = \sup_{\|x\|=1}\sup_{\|f\|=1}|f(Ax)| \\ &= \sup_{\|f\|=1}\sup_{\|x\|=1}\left|A^\dagger f(x)\right| = \sup_{\|f\|=1}\left\|A^\dagger f\right\| = \left\|A^\dagger\right\|.\end{aligned}$$
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
$$f \in X^* \overset{\hat{}}{\to} \hat{f} \in X^{***} \overset{J^\dagger}{\to} J^\dagger\hat{f} \in X^*$$
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 $X^* \overset{\hat{}}{\to} X^{***}$ 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
$$\begin{aligned}\|\pi(x)\| &= \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{\left|\sum_{i=1}^n \alpha_i \bar{f}_i(\pi(x))\right|}{\left\|\sum_{i=1}^n \alpha_i \bar{f}_i\right\|} = \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{|\sum_{i=1}^n \alpha_i \lambda(f_i)|}{\|\sum_{i=1}^n \alpha_i f_i\|} \\ &= \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{|\lambda(\sum_{i=1}^n \alpha_i f_i)|}{\|\sum_{i=1}^n \alpha_i f_i\|} \\ &\le \|\lambda\| \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{\|\sum_{i=1}^n \alpha_i f_i\|}{\|\sum_{i=1}^n \alpha_i f_i\|} = 1.\end{aligned}$$
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 Operator Ordering and the Lattice of Orthogonal Projections

See Definition **??** below and related material for operator ordering basics.

Definition A.26. *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 the collection of orthogonal projection on H. If $\mathcal{P}$ is an family of orthogonal projections on H then we 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 A.27. 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 complements then shows $\operatorname{Ran}(P) \subset \operatorname{Ran}(Q)$, i.e. $P \le Q$ as in Definition A.26.

Lemma A.28. *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 **??** which states; If $\mathcal{A}$ is a $*$ subalgebra of $L(H)$, 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 A.29. *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

$$A \operatorname{Ran}(P) \subset \operatorname{Ran}(P) \text{ and } A^* \operatorname{Ran}(P) \subset \operatorname{Ran}(P) \ \forall\ P \in \mathcal{P}. \tag{A.9}$$

By Eq. (A.9),

$$A\left[\cap_{P\in\mathcal{P}} \operatorname{Ran}(P)\right] \subset \left[\cap_{P\in\mathcal{P}} \operatorname{Ran}(P)\right] \text{ and}$$
$$A^*\left[\cap_{P\in\mathcal{P}} \operatorname{Ran}(P)\right] \subset \left[\cap_{P\in\mathcal{P}} \operatorname{Ran}(P)\right]$$

and therefore both A and A^* both preserve $\operatorname{Ran}(P_{\inf})$, i.e.

$$AP_{\inf} = P_{\inf}AP_{\inf} \text{ and } A^*P_{\inf} = P_{\inf}A^*P_{\inf}.$$

Taking adjoints of these equations also shows,

$$P_{\inf}A^* = P_{\inf}A^*P_{\inf} \text{ and } P_{\inf}A = P_{\inf}AP_{\inf}$$

and therefore $[A, P_{\inf}] = 0$.

Similarly by Eq. (A.9) 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 A.30. *Let $\mathcal{P}$ be a family of orthogonal projections on a Hilbert space H.*

1. *If there exists and orthogonal projection such 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 A.27 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 A.27 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 $Q = P_{\inf}$. ■

A.5 Rayleigh Quotient

Theorem A.31 (Rayleigh quotient). *If H is a Hilbert space and $T \in B(H)$ is a bounded self-adjoint operator, then*

$$M := \sup_{f\neq 0} \frac{|\langle Tf, f\rangle|}{\|f\|^2} = \|T\| \ \left(= \sup_{f\neq 0} \frac{\|Tf\|}{\|f\|}\right).$$

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. (**??**) with $Q(f,g) = \langle Tf, g\rangle$ and Eq. (**??**) with $Q(f,g) = \langle f, g\rangle$ along with the Cauchy-Schwarz inequality implies,

$$\begin{aligned} 4\operatorname{Re}\langle Tf, g\rangle &= \langle T(f+g), (f+g)\rangle - \langle T(f-g), (f-g)\rangle \\ &\le M\left[\|f+g\|^2 + \|f-g\|^2\right] = 2M\left[\|f\|^2 + \|g\|^2\right]. \end{aligned}$$

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

$$4\,|\langle Tf, g\rangle| \le 2M\left[\|f\|^2 + \|g\|^2\right]$$

and therefore,

$$\|T\| = \sup_{\|f\|=\|g\|=1} |\langle f, Tg\rangle| \le M$$

and since it is clear $M \le \|T\|$ we have shown $M = \|T\|$.

If $f \in H\setminus\{0\}$ and $\|T\| = |\langle Tf, f\rangle| / \|f\|^2$ then, using Schwarz's inequality,

$$\|T\| = \frac{|\langle Tf, f\rangle|}{\|f\|^2} \le \frac{\|Tf\|}{\|f\|} \le \|T\|. \tag{A.10}$$

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.10) 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,

$$|\langle Tf, f\rangle| = \left|\int_\Omega g\,|f|^2\, d\mu\right| \le \|g\|_{L^\infty(\mu)} \int_\Omega |f|^2\, d\mu = \|g\|_{L^\infty(\mu)}\, \|f\|^2_{L^2(\mu)}.$$

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 $\varepsilon g 1_A \ge m 1_A$. For this f it follows that

$$|\langle Tf, f\rangle| = \int_A \varepsilon g\, d\mu \ge m\cdot\mu(A) = m\,\|f\|^2_{L^2(\mu)}.$$

Combining these last two assertions shows

$$m \le \sup_{\|f\|\neq 0} \frac{|\langle Tf, f\rangle|}{\|f\|^2} \le \|T\|_{op}$$

which completes this proof as $m < \|T\|_{op}$ was arbitrary. ■