Bruce K. Driver

Functional Analysis Tools with Examples

January 6, 2020 *File:241Functional˙2020s.tex*

Contents

Part I Background

1 Vector Valued Integration Theory . . . 3
1.1 Pettis Integral . . . 3
1.2 Riemann Integrals of Continuous Integrands . . . 4
1.3 Bochner Integral (integrands with separable range) . . . 7
1.3.1 Proof of Theorem 1.12 . . . 7
1.4 Strong Bochner Integrals . . . 11
1.5 Weak integrals for Hilbert Spaces . . . 13

Part II Banach Algebra Spectral Theory

2 Banach Algebras and Linear ODE . . . 21
2.1 Basic Definitions, Examples, and Properties . . . 21
2.2 Calculus in Banach Algebras . . . 24
2.3 General Linear ODE in $\mathcal{A}$ . . . 26
2.4 Logarithms . . . 30
2.5 C^*-Banach algebras . . . 31
2.6 Exercises . . . 34

References . . . 37

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 [3–6] and especially [11]. Other references are **Pettis Integral** (See Craig Evans PDE book?) also see

$$http://en.wikipedia.org/wiki/Pettis\_integral$$

and

$$http://www.math.umn.edu/\tilde{}garrett/m/fun/Notes/07\_vv\_integrals.pdf$$

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 1.12 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

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

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. $\sup\left\{\left\|\frac{f_n(t+h)-f_n(t)}{h}\right\| : (t,h) \in J\times\mathbb{R}\ \ni\ t+h \in J \text{ and } h \ne 0\right\} \le a_n$.
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 $t \in J$,
$$\dot{F}(t) = \sum_{n=1}^{\infty} \dot{f}_n(t).$$

Note: if X is a complex Banach space, J is an open subset of $\mathbb{C}$, and $f_n : J \to X$ are analytic functions (see Definition 1.9 below) such that Eq. (1.10) holds, then the results of the exercise continue to hold provided $\dot{f}_n(t)$ is interpreted as the complex derivative of f_n.

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*

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

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 $D(z_0,\rho)$ 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

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

■

Corollary 1.11 (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

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

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

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

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

1.3 Bochner Integral (integrands with separable range)

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

Theorem 1.12. *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*

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

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 1.12.

1.3.1 Proof of Theorem 1.12

Proposition 1.13. *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 1.14. *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 1.15. *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 1.14 of the appendix which asserts that $\sigma(X^*) = \mathcal{B}_X$ when X is separable. ■

Corollary 1.16. *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 1.15. ■

Corollary 1.17. *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 1.18. *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 1.13 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 1.19. *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 1.20. *Suppose $a_n \in X$ and $\|a_{n+1} - a_n\| \le \varepsilon_n$ and $\sum_{n=1}^\infty \varepsilon_n < \infty$. Then $\lim_{n\to\infty} a_n = a \in X$ exists and $\|a - a_n\| \le \delta_n := \sum_{k=n}^\infty \varepsilon_k$.*

Proof. Let $m > n$ then

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

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

Lemma 1.21. *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 1.20, $F(\omega) := \lim_{j\to\infty} G_j(\omega)$ 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 1.20 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, $G_N \xrightarrow{\mu} F$ as $N \to \infty$.

With this in hand, it is straightforward to show $F_n \xrightarrow{\mu} F$. 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 $G_j \xrightarrow{\mu} F$. ■

Theorem 1.22. *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 1.21 there exists a subsequence $\{G_j\}$ of $\{F_n\}$ such that $G_j \to F$ a.e. By Fatou's Lemma,

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

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 1.23 (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 1.24. *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 $\left\|F_\varepsilon^N\right\| \leq \|F_\varepsilon\| \leq 2\|F\|$ for all N. Therefore another application of the dominated convergence theorem implies, $\lim_{N\to\infty} \left\|F_\varepsilon^N - F_\varepsilon\right\|_{L^p(\mu;X)} = 0$. Thus any $F \in L^p(\mu; X)$ may be arbitrarily well approximated by one of the $F_\varepsilon^N \in \mathcal{S}(\mu; X)$ with ε sufficiently small and N sufficiently large. ■

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

Lemma 1.25. *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 1.24 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 1.26. *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 1.27. *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 1.28 (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 1.29 (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 1.26. 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

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

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

Proof. The existence of a continuous linear map $\bar{I} : L^1(\Omega, \mathcal{F}, \mu; X) \to X$ such that $\bar{I}|_{\mathcal{S}(\mu;X)} = I$ and Eq. (1.18) holds follows from Propositions 1.27 and 1.24 and the bounded linear transformation Theorem 1.28. If $\varphi \in X^*$ and $F \in L^1(\Omega, \mathcal{F}, \mu; X)$, choose $F_n \in \mathcal{S}(\mu; X)$ such that $F_n \to F$ in $L^1(\Omega, \mathcal{F}, \mu; X)$ as $n \to \infty$. Then $\bar{I}(F) = \lim_{n\to\infty} I(F_n)$ and hence by Eq. (1.17),

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

This proves Eq. (1.19) since

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

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

Example 1.30. 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.

Remark 1.31. 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, $X_0 := \overline{\text{span}(\text{essran}_\mu(F))}$. 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 1.32 (DCT). *If $\{u_n\} \subset L^1(\mu; X)$ is such that $u(\omega) = \lim_{n\to\infty} u_n(\omega)$ exists for μ-a.e. x and there exists $g \in L^1(\mu)$ such that $\|u_n\|_X \le g$ a.e. for all n, then $u \in L^1(\mu; X)$ and $\lim_{n\to\infty} \|u - u_n\|_1 = 0$ and in particular,*

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

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

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

■

1.4 Strong Bochner Integrals

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

Definition 1.33. *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 1.34. *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 1.35. *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 1.24 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 1.16. ■

Corollary 1.36. *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 1.35. ■

Definition 1.37. *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 1.38. *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 1.36 that $C \otimes D$ is measurable on the product space. We further have

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

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

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

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

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

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

Since,

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

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

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

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

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

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

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

Show

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

is differentiable and

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

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 $\sigma-$finite measures spaces.

Definition 1.39. *A function $\psi:X\to H$ is said to be **weakly measurable** if $X\ni x\to\langle h,\psi\left(x\right)\rangle\in\mathbb{F}$ is $\mathcal{M}-$measurable for all $h\in H$.*

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

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

Definition 1.40. *A function $\psi:X\to H$ is **weakly-integrable** if ψ is weakly measurable and*

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

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

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

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

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

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

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

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

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

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

Theorem 1.41. *There is a unique linear map,*

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

such that

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

Moreover this map satisfies;

1. $$\left\|\int_X\psi\left(x\right)d\mu\left(x\right)\right\|_H\le\|\psi\|_{L^1(\mu:H)}.$$
2. *If $B\in L\left(H,K\right)$ is a bounded linear operator from H to K, then*

$$B\int_X\psi\left(x\right)d\mu\left(x\right)=\int_X B\psi\left(x\right)d\mu\left(x\right).$$

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 1.42. *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 1.43. *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 1.44. *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 1.45 (Weak Integrals) *We denote the $\bar{C}$ in Theorem 1.44 by either $\mu(C)$ or $\int_X C(x)\,d\mu(x)$.*

Theorem 1.46. *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,

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

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

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

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

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

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

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

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

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

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

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

■

Exercise 1.6. Let us continue to use the notation in Theorem 1.46. If $B \in B(H)$ is a linear operator such that $[C(x), B] = 0$ for μ – a.e. x, show $[\mu(C), B] = 0$.

Part II

Banach Algebra Spectral Theory

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 [8, 9].For general Banach and C^*-algebra stuff have a look at [7, 13]. Also see the lecture notes in [10, 12]. Putnam's file looked quite good. For a very detailed statements see [2, 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!]

Examples 2.3 *1. 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.*

2. 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.

3. If $(\Omega,\mathcal{B},\mu)$ is a measure space then $L^\infty(\mu) := L^\infty(\Omega,\mathcal{B},\mu:\mathbb{C})$ is a commutative complex Banach algebra with identity. In this case $\|f\| = \|f\|_{L^\infty(\mu)}$ is the essential supremum of $|f|$ defined by

$$\|f\|_{L^\infty(\mu)} = \inf\{M > 0 : |f| \le M \ \mu\text{-a.e.}\}.$$

4. 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 $\dim X > 1$.

*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*

$$\delta_0(n) := \begin{cases} 1 \text{ if } n = 0 \\ 0 \text{ if } n \neq 0 \end{cases}.$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

and

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

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

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

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

Define

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

and

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

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

Proof. Straightforward. ■

Remark 2.7. If $\mathcal{A}$ is a C^*-algebra as in Definition 2.48 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

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

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 $\mathcal{A}$.

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

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

As $(x + h) = x\left(1 + x^{-1}h\right)$, it follows that $x + h$ is invertible and

$$(x+h)^{-1} = \left(1 + x^{-1}h\right)^{-1} x^{-1}.$$

Taking norms of this equation then gives the estimate in Eq. (2.6). ■

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 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 2.1 below.]*

Proof. Suppose that $x \in \mathcal{A}_{inv}$ and $h \in \mathcal{A}$ is sufficiently small so that $\left\|x^{-1}h\right\| \le \left\|x^{-1}\right\| \|h\| < 1$. Then $x+h$ is invertible by Corollary 2.15 and we find the identity,

$$(x+h)^{-1} - x^{-1} = (x+h)^{-1}(x-(x+h))x^{-1} = -(x+h)^{-1}hx^{-1}. \tag{2.8}$$

From Eq. (2.8) and Corollary 2.15 it follows that

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

■

2.2 Calculus in Banach Algebras

Exercise 2.1. 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:** rewrite Eq. (2.8) as

$$(x+h)^{-1} = x^{-1} - (x+h)^{-1}hx^{-1} \tag{2.9}$$

and then iterate the identity once.

Exercise 2.2. Suppose that $a \in \mathcal{A}$ and $t \in \mathbb{R}$ (or $\mathbb{C}$ if $\mathcal{A}$ is a complex Banach algebra). Show directly that:

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

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

Proof. In the proof to follows we will use $e^{ta}b = be^{ta}$ for all $t \in \mathbb{R}$. [Proof is left to the reader.] Let $f(t) := e^{-ta}e^{t(a+b)}$, then by the product rule,

$$\dot{f}(t) = -e^{-ta}ae^{t(a+b)} + e^{-ta}(a+b)e^{t(a+b)} = e^{-ta}be^{t(a+b)} = be^{-ta}e^{t(a+b)} = bf(t).$$

Therefore, $\frac{d}{dt}\left[e^{-tb}f(t)\right] = 0$ and hence $e^{-tb}f(t) = e^{-0b}f(0) = 1$. Altogether we have shown,

$$e^{-tb}e^{-ta}e^{t(a+b)} = e^{-tb}f(t) = 1.$$

Taking $t = \pm 1$ and $b = 0$ in this identity shows $e^{-a}e^{a} = 1 = e^{a}e^{-a}$, i.e. $(e^a)^{-1} = e^{-a}$. Knowing this fact it then follows from the previously displayed equation that $e^{t(a+b)} = e^{ta}e^{tb}$ which at $t = 1$ gives, $e^a e^b = e^{a+b}$. Interchanging the roles of a and b then completes the proof. ■

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

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

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

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

Moreover,

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

We also have the following converse to this corollary whose proof is outlined in Exercise 2.14 below.

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

1. *(Semi-group property.) $T_0 = 1 \in \mathcal{A}$ and $T_t T_s = T_{t+s}$ for all $s, t \ge 0$.*
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 2.3. 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 2.4. 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

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

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

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 $f(t) := e^{ta}be^{-ta}$, then

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

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,

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

■

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,

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

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.

$$e^{a+b} = e$$

■

Exercise 2.5. 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

$$\begin{aligned}\left\|\dot{f}_n\right\| &\le \frac{1}{(n-1)!}\|a\|^{n-1}\|\dot{a}\|, \\ \|f_n'\| &\le \frac{1}{(n-1)!}\|a\|^{n-1}\|a'\|, \\ \left\|\ddot{f}_n\right\| &\le \frac{1}{(n-2)!}\|a\|^{n-2}\|\dot{a}\|^2 + \frac{1}{(n-1)!}\|a\|^{n-1}\|\ddot{a}\| \\ \left\|\dot{f}_n'\right\| &\le \frac{1}{(n-2)!}\|a\|^{n-2}\|\dot{a}\|\|a'\| + \frac{1}{(n-1)!}\|a\|^{n-1}\|\dot{a}'\| \\ \|f_n''\| &\le \frac{1}{(n-2)!}\|a\|^{n-2}\|a'\|^2 + \frac{1}{(n-1)!}\|a\|^{n-1}\|a''\|\end{aligned}$$

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

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

Proof. The function, $u(s,t) := e^{t(a+sb)}$ is C^2 by Exercise 2.5 and therefore we find,

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

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 2.24. *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 2.23,

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

and therefore,

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

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 2.23,

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

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 2.25 *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 2.28 that more generally we have

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

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

Lemma 2.26. *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 2.27. *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;

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

■

Corollary 2.28. *For all $n \in \mathbb{N}$,*

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

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

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

Proof. This follows by the simple estimate along with Lemma 2.26 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 2.29. *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 2.30 *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 2.31 (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 2.32 (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 2.33 (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*

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

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. ■

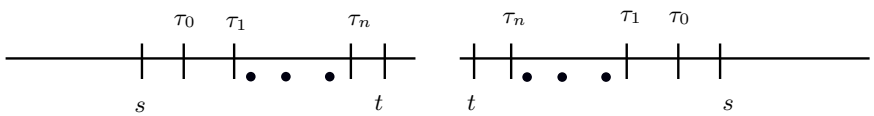

Fig. 2.1. This figures shows how to find the new limits of integration when $t > s$ and $t < s$ respectively.

Lemma 2.34. *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

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

Proof. Using Lemma 2.33 shows,

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

Similarly,

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

■

Proposition 2.35 (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,

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

and hence

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

which specializes to Eq. (2.25) when $\varphi(t) = \mathbf{1}$.Differentiating Eq. (2.25) on s then gives Eq. (2.26). Another proof of Eq. (2.26) may be given using Proposition 2.32 to conclude that $u(t,s) = u(s,t)^{-1}$ and then differentiating this equation shows

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

■

Theorem 2.36 (Du hamell'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

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

where by summing Eq. (2.24),

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

and the proof is complete.

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

$$\frac{d}{dt}\int_s^t u^A(t,\sigma)\,h(\sigma)\,d\sigma$$
$$= \frac{d}{d\varepsilon}\Big|_0\int_s^{t+\varepsilon} u^A(t,\sigma)\,h(\sigma)\,d\sigma + \frac{d}{d\varepsilon}\Big|_0\int_s^t u^A(t+\varepsilon,\sigma)\,h(\sigma)\,d\sigma$$
$$= u^A(t,t)\,h(t) + \int_s^t \frac{d}{dt}u^A(t,\sigma)\,h(\sigma)\,d\sigma$$
$$= h(t) + \int_s^t A(t)\,u^A(t,\sigma)\,h(\sigma)\,d\sigma$$
$$= h(t) + A(t)\int_s^t u^A(t,\sigma)\,h(\sigma)\,d\sigma.$$

Thus it follows that

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

■

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

Theorem 2.37. *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 2.1 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 2.38 (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 2.39. *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 2.40. *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 2.41. *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 2.42. *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 $C \geq 0$.

Proof. For $0 \leq t \leq 1$ 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

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

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

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

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

2.5 C^*-Banach 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 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 2.43. *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 $1 \in \mathcal{B}$, then automatically we have $1^* = 1$. Indeed, applying the involution to the identity, $1^* = 1 \cdot 1^*$ gives

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

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

Definition 2.44. *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 2.45. 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 2.46 (C^* - algebras). *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.

Remark 2.47. Conditions 1) and 2) in Definition 2.46 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

$$\|a\|^2 \leq \|a^*\| \|a\| \implies \|a\| \leq \|a^*\|.$$

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

It is fact the case that seemingly weaker condition 1. by itself implies condition 3., see Theorem 16.1 on page 45 of [2].[1]

Definition 2.48. *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 $a \in \mathcal{B}$.*

Example 2.49. 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.

Example 2.50. If $(\Omega, \mathcal{B}, \mu)$ is a measure space then $L^\infty(\mu) := L^\infty(\Omega, \mathcal{B}, \mu : \mathbb{C})$ is a commutative complex C^*-algebra with identity. Again we let $f^*(\omega) = \overline{f(\omega)}$. The C^*-condition is

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

Example 2.51. Let H be a Hilbert space and $\mathcal{B}$ be a $*-$ closed and operator norm-closed sub-algebra of $B(H)$. Then $(\mathcal{B}, *)$ is a C^*–algebra. The key point is to show $\|A^*A\| = \|A\|^2$ for all $A \in \mathcal{B}$ which we now prove in case you have forgotten this fact.

Proof. 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 the isometry condition. Similarly,

$$\begin{aligned}\|A\|^2 &= \sup_{h\in H:\|h\|=1} \|Ah\|^2 = \sup_{h\in H:\|h\|=1} |\langle Ah|Ah\rangle| \\ &= \sup_{h\in H:\|h\|=1} |\langle h|A^*Ah\rangle| \le \sup_{h\in H:\|h\|=1} \|A^*Ah\| = \|A^*A\| \qquad (2.38)\end{aligned}$$

[1] See https://en.wikipedia.org/wiki/C*-algebra for information about B* algebras being the same as a C* algebra.

and therefore,

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

Alternate proof. Using the Rayleigh quotient in Theorem **??** below, 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 $*$-closed 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 every abstract C^*-algebra may be "represented" by a "concrete" (i.e. sub-algebra of $B(H)$) C^*-algebra – see Conway for details of how to do this by looking up the GNS construction.

Example 2.53. 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 $\left(ei^A\right)^* = e^{-iA^*} = e^{-iA} = \left(e^{iA}\right)^{-1}$, which shows e^{iA} is unitary. This result is generalized in the following proposition.

Proposition 2.54. *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*

$$\dot{u}(t) = A(t)\,u(t) \text{ with } u(0) = 1 \qquad (2.39)$$

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 2.32 it follows that $u(t)^{-1} = u(0,t)$ and from Proposition 2.35 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. (2.39) shows

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

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

Lemma 2.55. *If $\mathcal{B}$ is a 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. ■

Definition 2.56. *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 2.57. *If H is a complex Hilbert space and $\mathcal{B}$ is a C^*-sub-algebra of $B(H)$, then $\mathcal{B}$ is symmetric.*

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

$$\|(I + A^*A)h\| \ge \|h\| .$$

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 invertible and moreover, $\left\|(I + A^*A)^{-1}\right\| \le 1$. ■

Example 2.58. Referring to Example 2.45 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

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

See Exercise **??** for more on this example.

Remark 2.59. According to Wikipedia[2], condition 1. and condition 3 of Definition **??** are equivalent but the implication 1. $\implies$ 3. is quite non-trivial. Historically condition 1. is called the C^*-condition on a norm and condition 3. is called the B^* – condition on a norm. Moreover, Wikipedia claims the term C^*-algebra was coined by I. Segal by which he mean a **C**losed $*$ – subalgebra of $B(H)$ where H is a Hilbert space as in Lemma 2.57. 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 [1, 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.]

As a consequence of Lemma 2.57 and assuming Remark 2.59, every C^*-algebra is symmetric. We will explicitly prove this fact for commutative C^*-algebras below in Lemma **??** and in general later without reference to Remark 2.59.

Definition 2.60. *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 $A \to A^*$.]*

Example 2.61. If $T \in \mathcal{B}$, then $C^*(\{T\}) = C^*(\{T, T^*\})$ is the norm closure of non-commutative polynomials in T, and T^* without constant term. If $1 \in \mathcal{B}$, then $C^*(\{T, 1\})$ is the norm closure of non-commutative polynomials in T and T^* with constant terms.

Before ending this section let us describe the C^*-subalgebras which are generated by commuting normal elements of $\mathcal{B}$. First we will need the following result.

Theorem 2.62 (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 $[N^*, B] = 0$.*

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

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

Then $u(0) = B$ and

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

[2] See https://en.wikipedia.org/wiki/C*-algebra for information about B* algebras being the same as a C* algebra.

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

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

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

Using what we have just proved and the normality assumptions[3] on N and M we have for any $w \in \mathbb{C}$ that

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

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

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

and hence by Example 2.53 and Lemma 2.55,

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

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 1.11) we conclude $f(z) = f(0) = B$ for all $z \in \mathbb{C}$, i.e.

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

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

Corollary 2.63. *Suppose that $\mathcal{B}$ be a C^*-algebra with identity and $\mathbf{T} := \{T_j\}_{j=1}^n \subset \mathcal{B}$ are commuting normal operators, then $\mathbf{T} \cup \mathbf{T}^* := \left\{T_j, T_j^*\right\}_{j=1}^n$ is a list of pairwise commuting operators and $C^*(\mathbf{T}, 1)$ is the norm closure of all elements of $\mathcal{B}$ of the form $p(\mathbf{T}, \mathbf{T}^*)$ where $p(z_1, \ldots, z_n, w_1, \ldots, w_n)$ is a polynomial in $2n$-variables. Moreover, $C^*(\mathbf{T}, 1)$ is a commutative C^*-subalgebra of $\mathcal{B}$.*

Remark 2.64. For the fun of it, here are two elementary proofs of Theorem 2.62 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 $E_\lambda^M := \operatorname{Nul}(M - \lambda I)$ 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 $N^*B = BM^*$.

2.6 Exercises

Exercise 2.6. To each $A \in \mathcal{A}$, we may define $L_A, R_A : \mathcal{A} \to \mathcal{A}$ by

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

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 2.7. 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

$$\dot{V}(t) = A(t)V(t) \text{ with } V(0) = I \tag{2.40}$$

and

$$\dot{U}(t) = -U(t)A(t) \text{ with } U(0) = I. \tag{2.41}$$

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 2.6 may be useful here.) Then use the uniqueness of solutions to linear O.D.E.s

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

Exercise 2.8. Suppose that $A \in \mathcal{A}$ and $v \in X$ is an eigenvector of A with eigenvalue λ, i.e. that $Av = \lambda v$. Show $e^{tA}v = e^{t\lambda}v$. Also show that if $X = \mathbb{R}^n$ and A is a diagonalizable $n \times n$ matrix with

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

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

Exercise 2.9. Suppose that $A, B \in \mathcal{A}$ let $ad_A B = [A, B] := AB - BA$. Show $e^{tA}Be^{-tA} = e^{tad_A}(B)$. In particular, if $[A, B] = 0$ then $e^{tA}Be^{-tA} = B$ for all $t \in \mathbb{R}$.

Exercise 2.10. Suppose that $A, B \in \mathcal{A}$ and $[A, B] := AB - BA = 0$. Show that $e^{(A+B)} = e^A e^B$.

Exercise 2.11. 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 2.12. 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 2.13. 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 2.14 (L. Gårding's trick I.). Prove Theorem 2.19, i.e. suppose that $T_t \in \mathcal{A}$ for $t \geq 0$ satisfies;

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

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

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

Exercise 2.15 (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 (2.40) 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:

$$\dot{W}(t) = A(t)\,W(t) + H(t) \;\text{ with } W(0) = W_0 \tag{2.42}$$

is given by

$$W(t) = V(t)\,W_0 + V(t)\int_0^t V(\tau)^{-1}\,H(\tau)\;d\tau. \tag{2.43}$$

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

References

1. John B. Conway, *A course in functional analysis*, second ed., Graduate Texts in Mathematics, vol. 96, Springer-Verlag, New York, 1990. MR 1070713
2. Robert Doran, *Characterizations of c* algebras: the gelfand naimark theorems*, CRC Press, 05 2019.
3. Robert F. Geitz, *Pettis integration*, Proc. Amer. Math. Soc. **82** (1981), no. 1, 81–86. MR 603606
4. ———, *Geometry and the Pettis integral*, Trans. Amer. Math. Soc. **269** (1982), no. 2, 535–548. MR 637707
5. 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
6. Robert Frederick Geitz, *THE PETTIS INTEGRAL*, ProQuest LLC, Ann Arbor, MI, 1980, Thesis (Ph.D.)–University of Illinois at Urbana-Champaign. MR 2630764
7. Laurent W. Marcoux, *An introduction to operator algebras*, University of Waterloo preprint book, 2016.
8. 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
9. ———, *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
10. Ian F. Putnam, *Lecture notes on c-star-algebras*, http://www.math.uvic.ca/faculty/putnam/ln/C*-algebras.pdf, 2019.
11. Michel Talagrand, *Pettis integral and measure theory*, Mem. Amer. Math. Soc. **51** (1984), no. 307, ix+224. MR 756174
12. Dana P. Williams, *A (very) short course on c-star algebras*, https://math.dartmouth.edu/ dana/bookspapers/cstar.pdf, 2019.
13. Ke He Zhu, *An introduction to operator algebras*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1993. MR 1257406