The Spectral Theorem for Self-Adjoint and Unitary Operators

Michael Taylor

Contents

1. Introduction
2. Functions of a self-adjoint operator
3. Spectral theorem for bounded self-adjoint operators
4. Functions of unitary operators
5. Spectral theorem for unitary operators
6. Alternative approach
7. From Theorem 1.2 to Theorem 1.1

A. Spectral projections
B. Unbounded self-adjoint operators
C. Von Neumann's mean ergodic theorem

1. Introduction

If H is a Hilbert space, a bounded linear operator $A : H \to H$ ($A \in \mathcal{L}(H)$) has an adjoint $A^* : H \to H$ defined by

$$(Au, v) = (u, A^*v), \quad u, v \in H. \tag{1.1}$$

We say A is self-adjoint if $A = A^*$. We say $U \in \mathcal{L}(H)$ is unitary if $U^* = U^{-1}$. More generally, if $\mathcal{H}$ is another Hilbert space, we say $\Phi \in \mathcal{L}(H, \mathcal{H})$ is unitary provided Φ is one-to-one and onto, and $(\Phi u, \Phi v)_{\mathcal{H}} = (u, v)_H$, for all $u, v \in H$.

If $\dim H = n < \infty$, each self-adjoint $A \in \mathcal{L}(H)$ has the property that H has an orthonormal basis of eigenvectors of A. The same holds for each unitary $U \in \mathcal{L}(H)$. Proofs can be found in §§11–12, Chapter 2, of [T3]. Here, we aim to prove the following infinite dimensional variant of such a result, called the Spectral Theorem.

Theorem 1.1. *If $A \in \mathcal{L}(H)$ is self-adjoint, there exists a measure space $(X, \mathcal{F}, \mu)$, a unitary map $\Phi : H \to L^2(X, \mu)$, and $a \in L^\infty(X, \mu)$, such that*

$$\Phi A \Phi^{-1} f(x) = a(x) f(x), \quad \forall\, f \in L^2(X, \mu). \tag{1.2}$$

Here, a is real valued, and $\|a\|_{L^\infty} = \|A\|$.

Here is the appropriate variant for unitary operators.

Theorem 1.2. *If $U \in \mathcal{L}(H)$ is unitary, there exists a measure space $(X, \mathcal{F}, \mu)$, a unitary map $\Phi : H \to L^2(X, \mu)$, and $u \in L^\infty(X, \mu)$, such that*

$$\Phi U \Phi^{-1} f(x) = u(x) f(x), \quad \forall\, f \in L^2(X, \mu). \tag{1.3}$$

Here, $|u| = 1$ on X.

These results are proven in §1 of [T2]. The proof there makes use of the Fourier transform on the space of tempered distributions. Here we prove these results, assuming as background only §§1–5 in [T1]. We use from §5 of that appendix the holomorphic functional calculus, which we recall here.

If V is a Banach space and $T \in \mathcal{L}(V)$, we say $\zeta \in \mathbb{C}$ belongs to the resolvent set $\rho(T)$ if and only if $\zeta - T : V \to V$ is invertible. We define $\sigma(T) = \mathbb{C} \setminus \rho(T)$ to be the spectrum of T. It is shown that $\rho(T)$ is open and $\sigma(T)$ is closed and bounded, hence compact. If Ω is a neighborhood of $\sigma(T)$ with smooth boundary and f is holomorphic on a neighborhood of $\overline{\Omega}$, we set

$$f(T) = \frac{1}{2\pi i} \int_{\partial\Omega} f(\zeta)(\zeta - T)^{-1}\, d\zeta. \tag{1.4}$$

It is shown there that

$$f_k(\zeta) = \zeta^k \ (k \in \mathbb{Z}^+) \Longrightarrow f_k(T) = T^k, \tag{1.5}$$

and, if f and g are holomorphic on a neighborhood of $\overline{\Omega}$, then

$$(fg)(T) = f(T)g(T). \tag{1.6}$$

To apply these results to self-adjoint A and unitary U in $\mathcal{L}(H)$, it is useful to have the following information on their spectra.

Proposition 1.3. *If $A \in \mathcal{L}(H)$ is self-adjoint, then $\sigma(A) \subset \mathbb{R}$.*

Proof. If $\lambda = x + iy,\ y \neq 0$, then

$$\lambda - A = y(i + B), \quad B = \frac{x}{y} - \frac{1}{y}A = B^*,$$

so it suffices to show that if $B = B^*$, then $i + B : H \to H$ is invertible. First note that, given $u \in H$,

$$\begin{aligned} \|u\| = 1 \Rightarrow \|(i+B)u\| &\geq |((i+B)u, u)| \\ &= |((I - iB)u, u)| \\ &\geq \mathrm{Re}\{(u,u) - i(Bu,u)\}. \end{aligned} \tag{1.7}$$

Now

$$B = B^* \Rightarrow (Bu, u) = (u, Bu) = \overline{(Bu,u)} \Rightarrow (Bu,u) \in \mathbb{R}, \tag{1.8}$$

so the last line of (1.7) equals $\|u\|^2$, and we have

$$\|(i+B)u\| \geq \|u\|, \quad \text{when} \ \ B = B^*. \tag{1.9}$$

Hence $i + B : H \to H$ is injective and has closed range. To see that it is onto, assume $w \in H$ is orthogonal to this range, so

$$\begin{aligned} v \in H &\Longrightarrow ((i+B)v, w) = 0 \\ &\Longrightarrow (v, (i-B)w) = 0. \end{aligned} \tag{.110}$$

Hence $(i + B)w = 0$. Then (1.9), with B replaced by $-B$, gives $w = 0$. This establishes the asserted surjectivity, and proves Proposition 1.3.

From Proposition 1.3 it follows that, if $A = A^*$,

$$\sigma(A) \subset [-\|A\|, \|A\|]. \tag{1.11}$$

Here is the unitary counterpart of Proposition 1.3.

Proposition 1.4. *If $U \in \mathcal{L}(H)$ is unitary, then*

$$\sigma(U) \subset S^1 = \{\zeta \in \mathbb{C} : |\zeta| = 1\}.$$

Proof. If U is unitary, $\|U\| = 1$, so $|\zeta| > 1 \Rightarrow \zeta \in \rho(U)$. Meanwhile,

$$|\zeta| < 1 \Longrightarrow \zeta - U = -U(I - \zeta U^{-1}),$$

and $\|\zeta U^{-1}\| = |\zeta|\,\|U^*\| = |\zeta| < 1$, so again this operator is invertible.

We proceed in the rest of this note as follows. In §2 we take $I = [-\|A\| - \delta, \|A\| + \delta]$ (with $\delta > 0$ small) and show that $f \mapsto f(A)$ extends from the space $\mathcal{O}$ of functions f holomorphic on a complex neighborhood of I to $C(I)$. To do this, we show that

$$f \in \mathcal{O},\ f \ge 0 \ \text{ on } \ I \Longrightarrow f(A) \ge 0, \tag{1.12}$$

where, given $T \in \mathcal{H}$, we say

$$T \ge 0 \Longleftrightarrow T = T^* \ \text{ and } \ (Tv, v) \ge 0, \ \ \forall\, v \in H. \tag{1.13}$$

We deduce from (1.12) that

$$f \in \mathcal{O},\ |f| \le M \ \text{ on } \ I \Longrightarrow \|f(A)\| \le M. \tag{1.14}$$

From this, the extension of $f \mapsto f(A)$ to $f \in C(I)$ will follow. We will have

$$\forall\, f \in C(I), \quad f \ge 0 \Rightarrow f(A) \ge 0, \ \ |f| \le M \Rightarrow \|f(A)\| \le M. \tag{1.15}$$

We use this fact to prove Theorem 1.1 in §3.

In §4, we use arguments parallel to those in §2 to extend $f \mapsto f(U)$ from the space $\mathcal{O}$ of functions holomorphic on a complex neighborhood of S^1 to $C(S^1)$. We show that

$$f \in \mathcal{O},\ f \ge 0 \ \text{ on } \ S^1 \Longrightarrow f(U) \ge 0, \tag{1.16}$$

and

$$f \in \mathcal{O},\ |f| \le M \ \text{ on } \ S^1 \Longrightarrow \|f(U)\| \le M, \tag{1.17}$$

and from there obtain the asserted extension, and an analogue of (1.15), namely

$$\forall\, f \in C(S^1) \quad f \ge 0 \Rightarrow f(U) \ge 0, \ \ |f| \le M \Rightarrow \|f(U)\| \le M. \tag{1.18}$$

Then, in §5, we use this extension to prove Theorem 1.2.

One ingredient in the analysis in §4 is the identity

$$f(U) = \sum_{k=-\infty}^{\infty} \hat{f}(k) U^k, \quad \hat{f}(k) = \frac{1}{2\pi} \int_{S^1} f(e^{i\theta})\, d\theta, \tag{1.19}$$

for $f \in \mathcal{O}$, which follows from (1.4), with Ω an annular neighborhood of $S^1 \subset \mathbb{C}$. In §6, we take (1.19) as the definition of $f(U)$, whenever f is a function on S^1 such that $\sum |\hat{f}(k)| < \infty$, which in particular holds for all $f \in C^\infty(S^1)$. We demonstrate directly, without using any holomorphic functional calculus, that

$$f \in C^\infty(S^1),\ |f| \le M \ \text{ on } \ S^1 \Longrightarrow \|f(U)\| \le M, \tag{1.20}$$

and, from this obtain the extension $f \mapsto f(U)$ to $f \in C(S^1)$, satisfying (1.18). This again leads to a proof of Theorem 1.2, which in this setting avoids use of the holomorphic functional calculus.

In §7, we show how to go from the Spectral Theorem for unitary operators (Theorem 1.2) to the Spectral Theorem for self-adjoint operators (Theorem 1.1). This, in conjunction with §6, yields a proof of Theorem 1.1 that does not require the holomorphic functional calculus.

In Appendix A, we relate Theorems 1.1 and 1.2 to other formulations of the Spectral Theorem, involving spectral projections. In Appendix B, we define unbounded self-adjoint operators and sketch an approach to the Spectral Theorem for such operators, somewhat parallel to the derivation of Theorem 1.1 from Theorem 1.2 in §7, referring to material in [T1] and [T2] for details. In Appendix C we use Theorem 1.2 to derive von Neumann's mean ergodic theorem.

2. Functions of a self-adjoint operator

Let $A \in \mathcal{L}(H)$ be self-adjoint. As previewed in §1, we take $\delta > 0$ and set

$$I = [-\|A\| - \delta, \|A\| + \delta], \quad I_0 = [-\|A\|, \|A\|], \tag{2.1}$$

so $\sigma(A) \subset I_0 \subset I$. Let Ω be a smoothly bounded neighborhood of I_0 in $\mathbb{C}$, and assume $\overline{\Omega} \cap \mathbb{R} \subset I$. Let f be holomorphic on a neighborhoood of $\overline{\Omega}$. We have

$$f(A) = \frac{1}{2\pi i} \int\limits_{\partial\Omega} f(\zeta)(\zeta - A)^{-1}\, d\zeta. \tag{2.2}$$

Now, given $S, T \in \mathcal{L}(H)$, we have $(ST)^* = T^* S^*$, so

$$((\zeta - A)^{-1})^* = (\overline{\zeta} - A)^{-1}. \tag{2.3}$$

Hence

$$f(A)^* = -\frac{1}{2\pi i} \int\limits_{\partial\Omega} \overline{f(\zeta)}(\overline{\zeta} - A)^{-1}\, d\overline{\zeta}. \tag{2.4}$$

We can assume Ω is invariant under the reflection $\zeta \mapsto \overline{\zeta}$, which reverses orientation on $\partial\Omega$. Hence

$$f(A)^* = \frac{1}{2\pi i} \int\limits_{\partial\Omega} \overline{f(\overline{\zeta})}(\zeta - A)^{-1}\, d\zeta = f^*(A), \tag{2.5}$$

where

$$f^*(\zeta) = \overline{f(\overline{\zeta})}. \tag{2.6}$$

In particular, if f is holomorphic on a neighborhood Ω of I_0 in $\mathbb{C}$,

$$f \text{ real on } \Omega \cap \mathbb{R} \Longrightarrow f(A) \text{ self-adjoint}. \tag{2.7}$$

Here is our key positivity result.

Proposition 2.1. *If f is holomorphic on a neighborhood Ω of $I_0 \subset \mathbb{C}$, then*

$$f \geq 0 \ \text{ on } \ \Omega \cap \mathbb{R} \Longrightarrow f(A) \geq 0. \tag{2.8}$$

Proof. For $\varepsilon > 0$, set $f_\varepsilon(\zeta) = f(\zeta) + \varepsilon$. Then we have a well defined holomorphic function

$$g_\varepsilon(\zeta) = f_\varepsilon(\zeta)^{1/2}, \tag{2.9}$$

on a sufficiently small neighborhood Ω of $I_0 \subset \mathbb{C}$, with $g_\varepsilon \geq 0$ on $\Omega \cap \mathbb{R}$. Hence $g_\varepsilon(A)$ is self-adjoint, and

$$\begin{aligned} f(A) + \varepsilon = f_\varepsilon(A) &= g_\varepsilon(A)^2 \\ &\Longrightarrow f(A) + \varepsilon I \geq 0, \quad \forall\, \varepsilon > 0 \\ &\Longrightarrow f(A) \geq 0. \end{aligned} \tag{2.10}$$

This leads to the following key bound.

Proposition 2.2. *If f is holomorphic on a neighborhood Ω of $I \subset \mathbb{C}$, then*

$$|f| \le M \quad on \quad I \Longrightarrow \|f(A)\| \le M. \tag{2.11}$$

Proof. Set $g(\zeta) = f^*(\zeta)f(\zeta)$, so $g = |f|^2$ on $\Omega \cap \mathbb{R}$. Hence

$$M^2 - g \ge 0 \quad \text{on} \quad I, \tag{2.12}$$

so, by Proposition 2.1 (plus the fact that $g(A) = f(A)^* f(A)$),

$$M^2 - f(A)^* f(A) \ge 0. \tag{2.13}$$

Hence, for $v \in H$,

$$\|f(A)v\|^2 = (f^*(A)f(A)v, v) \le (M^2 v, v) = M^2 \|v\|^2, \tag{2.14}$$

and we have (2.11).

We now extend the functional calculus.

Proposition 2.3. *If $A \in \mathcal{L}(H)$ is self-adjoint, the correspondence $f \mapsto f(A)$ has a unique extension from the space $\mathcal{O}$ of functions holomorphic in a neighborhood of $I \subset \mathbb{C}$ to $C(I)$. Given $f, g \in C(I)$, we have*

$$\begin{aligned} &\|f(A)\| \le \sup_I |f|, \quad f \ge 0 \Rightarrow f(A) \ge 0, \\ &f(A)^* = \overline{f}(A), \quad (fg)(A) = f(A)g(A). \end{aligned} \tag{2.15}$$

Proof. The extension is straightforward, via Proposition 2.2 and the denseness of $\mathcal{O}$ in $C(I)$ (which follows from the Weierstrass approximation theorem). If $f \in C(I)$, $f_k \in \mathcal{O}$, $f_k \to f$ uniformly on I, then

$$\|f_j(A) - f_k(A)\| \le \sup_I |f_j - f_k|, \tag{2.16}$$

yielding convergence of $f_k(A)$ in operator norm, to a limit we denote $f(A)$. The results in (2.15) follow from their counterparts for elements of $\mathcal{O}$ by the obvious limiting arguments.

3. Spectral theorem for bounded self-adjoint operators

Given a self-adjoint $A \in \mathcal{L}(H)$ and interval I as in (2.1), take a nonzero $v \in H$ and set

$$\mu(f) = \mu_{A,v}(f) = (f(A)v, v), \quad f \in C(I). \tag{3.1}$$

We have from Proposition 2.3 that $f \geq 0 \Rightarrow f(A) \geq 0$, hence $\mu(f) \geq 0$. Thus μ gives a positive Radon measure on I:

$$(f(A)v, v) = \int_I f \, d\mu. \tag{3.2}$$

Now we define

$$W = W_{A,v} : C(I) \longrightarrow H \tag{3.3}$$

by

$$W(f) = f(A)v. \tag{3.4}$$

Note that, if also $g \in C(I)$,

$$\begin{aligned} (W(f), W(g))_H &= (f(A)v, g(A)v) \\ &= (\overline{g}(A)f(A)v, v) \\ &= \int_I f\overline{g} \, d\mu \\ &= (f, g)_{L^2(I,\mu)}. \end{aligned} \tag{3.5}$$

Consequently W has a unique continuous extension to

$$W : L^2(I, \mu) \longrightarrow H, \quad \text{an isometry.} \tag{3.6}$$

Note that the range of W in (3.6) is the closure in H of

$$\mathcal{C}(A, v) = \operatorname{Span}\{v, Av, A^2v, \dots\}. \tag{3.7}$$

We call $\overline{\mathcal{C}(A, v)} \subset H$ the cyclic subspace of H generated by A and v. If $\overline{\mathcal{C}(A, v)} = H$, we say v is a cyclic vector for A. The following is a special case of Theorem 1.1.

Proposition 3.1. *If $A \in \mathcal{L}(H)$ is self-adjoint and has a cyclic vector v, then*

$$W : L^2(I,\mu) \longrightarrow H \tag{3.8}$$

is unitary, and

$$W^{-1}AWf(x) = xf(x), \quad \forall\, f \in L^2(I,\mu). \tag{3.9}$$

Proof. The unitarity follows from (3.6)–(3.7). To get (3.9), note that

$$W(xf) = Af(A)v = AW(f), \tag{3.10}$$

the first identity by (3.4), with f replaced by xf, plus, from (2.15), the fact that $(xf)(A) = Af(A)$. The identity (3.10) holds first for $f \in C(I)$, hence, by continuity, for all $f \in L^2(I,\mu)$.

In general, we cannot say that a given self-adjoint $A \in \mathcal{L}(H)$ has a cyclic vector, but we have the following. For simplicity, we assume H is separable.

Proposition 3.2. *If H is separable and $A \in \mathcal{L}(H)$ is self-adjoint, then there exist $v_j \in H$ such that $\overline{\mathcal{C}(A,v_j)}$ are mutually orthogonal subspaces of H, with span dense in H.*

Proof. Let $\{w_j : j \in \mathbb{N}\}$ be a dense subset of H, all $w_j \neq 0$. Take $v_1 = w_1$, and construct $\overline{\mathcal{C}(A,v_1)} = H_1$, as above. Note that $A : H_1 \to H_1$.

If $H_1 = H$, we are done. If not, we proceed as follows. We claim that, whenever $H_1 \subset H$ is a linear subspace,

$$A : H_1 \to H_1 \Longrightarrow A : H_1^\perp \to H_1^\perp. \tag{3.11}$$

In fact, if $v \in H_1,\ w \in H_1^\perp$, then

$$(v, Aw) = (Av, w) = 0 \quad (\text{given } Av \in H_1,\ w \in H_1^\perp), \tag{3.12}$$

so (3.11) follows.

To continue, consider the first $j \geq 2$ such that $w_j \notin H_1$, and let v_2 denote the orthogonal projection of w_j onto $H_1^\perp$. Then set $H_2 = \overline{\mathcal{C}(A,v_2)} \subset H_1^\perp$. Clearly $H_1 \oplus H_2$ contains $\mathrm{Span}\{w_k : 1 \leq k \leq j\}$. If $H_1 \oplus H_2 = H$, we are done. If not, $A : (H_1 \oplus H_2)^\perp \to (H_1 \oplus H_2)^\perp$. Take the first $j_3 > j$ such that $w_{j_3} \notin H_1 \oplus H_2$, and let v_3 denote the orthogonal projection of w_{j_3} onto $(H_1 \oplus H_2)^\perp$. Then set $H_3 = \overline{\mathcal{C}(A,v_3)} \subset (H_1 \oplus H_2)^\perp$.

Continue. If, for some K, $H_1 \oplus \cdots \oplus H_K = H$, we are done. If not, we get a countable sequence of mutually orthogonal spaces $H_k = \overline{\mathcal{C}(A,v_k)}$, whose span contains w_j for all $j \in \mathbb{N}$, so is dense in H. Hence we have Proposition 3.2.

We can now prove Theorem 1.1, when H is separable. Write

$$H = \bigoplus_{j \geq 1} H_j, \quad H_j = \overline{\mathcal{C}(A, v_j)}, \tag{3.13}$$

and

$$W_j = W_{A,v_j} : C(I) \to H_j, \quad W_j(f) = f(A)v_j, \tag{3.14}$$

extending to unitary maps

$$W_j : L^2(I, \mu_j) \to H_j, \quad \mu_j(f) = (f(A)v_j, v_j), \tag{3.15}$$

satisfying

$$W_j^{-1} A W_j f = xf, \quad f \in L^2(I, \mu_j). \tag{3.16}$$

Then we can define the measure space (X, μ) as the *disjoint union*

$$(X, \mu) = \bigcup_{j \geq 1} (I_j, \mu_j), \quad I_j = I, \tag{3.17}$$

so

$$L^2(X, \mu) = \bigoplus_{j \geq 1} L^2(I, \mu_j), \tag{3.18}$$

and the W_j in (3.15) fit together to give the unitary map

$$W : L^2(X, \mu) \longrightarrow H, \tag{3.19}$$

satisfying

$$W^{-1} A W f = a(x) f, \quad a(x) = x \text{ on } I_j. \tag{3.20}$$

If H is not separable, one can find a suitable replacement for Proposition 3.2, using Zorn's lemma. We omit details.

4. Functions of unitary operators

Let $U \in \mathcal{L}(H)$ be unitary. We have $\sigma(U) \subset S^1 = \{\zeta \in \mathbb{C} : |\zeta| = 1\}$. If Ω is a neighborhood of S^1 in $\mathbb{C}$, containing $\{\zeta \in \mathbb{C} : 1 - \varepsilon \le |\zeta| \le 1 + \varepsilon\}$, we have, for f holomorphic on Ω,

$$\begin{aligned} f(U) &= \frac{1}{2\pi i} \int\limits_{\partial\Omega} f(\zeta)(\zeta - U)^{-1}\, d\zeta \\ &= \frac{1}{2\pi i} \Big[\int\limits_{|\zeta|=1+\varepsilon} f(\zeta)\zeta^{-1} \sum_{k=0}^{\infty} \zeta^{-k} U^k \, d\zeta \\ &\qquad - \int\limits_{|\zeta|=1-\varepsilon} f(\zeta)(-U) \sum_{\ell=0}^{\infty} \zeta^{\ell} U^{-\ell}\, d\zeta \Big]. \end{aligned} \tag{4.1}$$

We can pass to the limit $\varepsilon \searrow 0$. With $\zeta = e^{i\theta}$, we have

$$\begin{aligned} \frac{1}{2\pi i} \zeta^{-k-1}\, d\zeta &= \frac{1}{2\pi} e^{-ik\theta}\, d\theta, \\ \frac{1}{2\pi i} \zeta^{\ell}\, d\zeta &= \frac{1}{2\pi} e^{i(\ell+1)\theta}\, d\theta, \end{aligned} \tag{4.2}$$

hence

$$f(U) = \sum_{k=-\infty}^{\infty} \hat{f}(k) U^k, \tag{4.3}$$

where

$$\hat{f}(k) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) e^{-ik\theta}\, d\theta \tag{4.4}$$

are the coefficients in the Fourier series for f. It follows that $f \mapsto f(U)$ etends uniquely from the space $\mathcal{O}$ of functions holomorphic in a neighborhood of $S^1 \subset \mathbb{C}$ to larger spaces, such as $\mathcal{A}(S^1) = \{f : \sum |\hat{f}(k)| < \infty\}$, or the space $C^\infty(S^1)$:

$$f \in C^\infty(S^1) \ \text{ yields } \ f(U), \ \text{given by (4.3).} \tag{4.5}$$

The result (1.6) extends by continuity to

$$f, g \in C^\infty(S^1) \Longrightarrow (fg)(U) = f(U)g(U), \tag{4.6}$$

a result that can also be deduced directly from (4.3) and a formula for $(fg)\hat{}(k)$ in terms of $\hat{f}$ and $\hat{g}$ (cf. §6). We also have

$$g = \overline{f} \Longrightarrow \hat{g}(k) = \overline{\hat{f}(-k)}, \tag{4.7}$$

and hence

$$f(U)^* = \overline{f}(U), \quad \forall\, f \in C^\infty(S^1). \tag{4.8}$$

In particular, if $f : S^1 \to \mathbb{R}$, then $f(U)$ is self-adjoint. Here is an analogue of Proposition 2.1.

Proposition 4,1. *If $f \in C^\infty(S^1)$, then*

$$f \geq 0 \quad on \quad S^1 \Longrightarrow f(U) \geq 0. \tag{4.9}$$

Proof. For $\varepsilon > 0$, set $f_\varepsilon(\zeta) = f(\zeta) + \varepsilon,\ \zeta \in S^1$. Then we have

$$g_\varepsilon(\zeta) = f_\varepsilon(\zeta)^{1/2}, \quad g_\varepsilon \in C^\infty(S^1), \quad g_\varepsilon > 0 \ \text{ on } \ S^1. \tag{4.10}$$

Hence $g_\varepsilon(U)$ is self-adjoint, and

$$\begin{aligned} f(U) + \varepsilon = f_\varepsilon(U) &= g_\varepsilon(U)^2 \\ &\Longrightarrow f(U) + \varepsilon I \geq 0, \quad \forall\, \varepsilon > 0 \\ &\Longrightarrow f(U) \geq 0. \end{aligned} \tag{4.11}$$

This leads to the following analogue of Proposition 2.2.

Proposition 4.2. *If $f \in C^\infty(S^1)$, then*

$$|f| \leq M \quad on \quad S^1 \Longrightarrow \|f(U)\| \leq M. \tag{4.12}$$

Proof. Set $g = |f|^2 = \overline{f}\, f \in C^\infty(S^1)$. Then $M^2 - g \geq 0$ on S^1, so, by Proposition 4.1 (plus (4.6) and (4.8))

$$M^2 - f^*(U) f(U) \geq 0. \tag{4.13}$$

The desired estimate follows as in (2.14).

We then have the following analogue of Proposition 2.3.

Proposition 4.3. *If $U \in \mathcal{L}(H)$ is unitary, the correspondence $f \mapsto f(U)$ has a unique continuous extension from $C^\infty(S^1)$ to $C(S^1)$. Given $f, g \in C(S^1)$, we have*

$$\begin{gathered} \|f(U)\| \leq \sup_{S^1} |f|, \quad f \geq 0 \Rightarrow f(U) \geq 0, \\ f(U)^* = \overline{f}(U), \quad (fg)(U) = f(U) g(U). \end{gathered} \tag{4.14}$$

Proof. Straightforward variant of the proof of Proposition 2.3.

5. Spectral theorem for unitary operators

Given a unitary operator $U \in \mathcal{L}(H)$, take a nonzero $v \in H$ and set

$$\mu(f) = \mu_{U,v}(f) = (f(U)v, v), \quad f \in C(S^1). \tag{5.1}$$

We have from Proposition 4.3 that $f \geq 0 \Rightarrow f(U) \geq 0$, hence $\mu(f) \geq 0$. Thus, as in §3, μ gives a Radon measure on S^1:

$$(f(U)v, v) = \int_{S^1} f \, d\mu. \tag{5.2}$$

Now we can define

$$W = W_{U,v} : C(S^1) \longrightarrow H \tag{5.3}$$

by

$$W(f) = f(U)v. \tag{5.4}$$

Parallel to (3.5), if also $g \in C(S^1)$,

$$\begin{aligned} (W(f), W(g))_H &= (f(U)v, g(U)v) \\ &= (\overline{g}(U)f(U)v, v) \\ &= \int_{S^1} f\overline{g} \, d\mu \\ &= (f, g)_{L^2(S^1,\mu)}. \end{aligned} \tag{5.5}$$

Consequently, W has a unique continuous extension to

$$W : L^2(S^1, \mu) \longrightarrow H, \quad \text{an isometry.} \tag{5.6}$$

The range of W in (5.6) is the closure in H of

$$\mathcal{C}(U, v) = \operatorname{Span}\{U^k v : k \in \mathbb{Z}\}. \tag{5.7}$$

We call $\overline{\mathcal{C}(U, v)}$ the cyclic subspace of H generated by U and v. If $\overline{\mathcal{C}(U, v)} = H$, we say v is a cyclic vector for U. The following is parallel to Proposition 3.1.

Proposition 5.1. *If $U \in \mathcal{L}(H)$ is unitary and has a cyclic vector v, then*

$$W : L^2(S^1, \mu) \longrightarrow H \tag{5.8}$$

is unitary, and

$$W^{-1}UWf(\zeta) = \zeta f(\zeta), \quad \forall\, f \in L^2(S^1, \mu). \tag{5.9}$$

(Note that in (3.9), x ranges over $I \subset \mathbb{R}$, while in (5.9), ζ ranges over $S^1 \subset \mathbb{C}$.)

Proof. The unitarity of W in (5.8) follows from (5.6)–(5.7). To get (5.9), note that

$$W(\zeta f) = Uf(U) = UW(f), \tag{5.10}$$

the first identity via (5.4), with $f(\zeta)$ replaced by $\zeta f(\zeta)$, plus, from (4.14), $(\zeta f)(U) = Uf(U)$.

Passing from Proposition 5.1 to Theorem 1.2 is done via a straightforward analogue of Proposition 3.2. We omit the details.

6. Alternative approach

Here we provide an alternative approach to the functional calculus $f \mapsto f(U)$ when $U \in \mathcal{L}(H)$ is unitary. As a definition, we take

$$f(U) = \sum_{k=-\infty}^{\infty} \hat{f}(k) U^k, \quad \hat{f}(k) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) e^{-ik\theta}\, d\theta, \tag{6.1}$$

whenever $\sum |\hat{f}(k)| < \infty$ (we say $f \in \mathcal{A}(S^1)$). In fact, to begin we typically take $f \in C^\infty(S^1)$. Our goal is to extend $f \mapsto f(U)$ to $f \in C(S^1)$. Note that (6.1) coincides with (4.3)–(4.4). The difference is that here we make no use of the holomorphic functional calculus, used in §§1–4. Thus we derive basic identities from scratch, from the definition (6.1).

To begin,

$$e_\ell(\zeta) = \zeta^\ell \Rightarrow \hat{e}_\ell(k) = \delta_{k\ell} \Rightarrow e_\ell(U) = U^\ell, \quad \forall\, \ell \in \mathbb{Z}. \tag{6.2}$$

Note that $\varphi(\zeta) = \overline{f(\zeta)} \Rightarrow \overline{\hat{f}(k)} = \hat{\varphi}(-k)$, so, since $U^* = U^{-1}$,

$$f(U)^* = \sum_{k=-\infty}^{\infty} \overline{\hat{f}(k)} U^{-k} = \overline{f}(U). \tag{6.3}$$

To proceed with an analysis of $(fg)(U)$ (given $f, g \in \mathcal{A}(S^1)$), first note that

$$g_\ell(\zeta) = \zeta^\ell g(\zeta) = e_\ell(\zeta) g(\zeta) \Rightarrow \hat{g}_\ell(k) = \hat{g}(k - \ell), \tag{6.4}$$

so

$$g_\ell(U) = \sum \hat{g}_\ell(k) U^k = \sum \hat{g}(k) U^{k+\ell} = U^\ell g(U). \tag{6.5}$$

Now

$$(fg)(\zeta) = \sum_\ell \hat{f}(\ell) \zeta^\ell g(\zeta) = \sum_\ell \hat{f}(\ell) g_\ell(\zeta), \tag{6.6}$$

so

$$(fg)(U) = \sum_\ell \hat{f}(\ell) g_\ell(U) = \sum_\ell \hat{f}(\ell) U^\ell g(U) = f(U) g(U), \tag{6.7}$$

given $f, g \in \mathcal{A}(S^1)$ (especially, $f, g \in C^\infty(S^1)$).

If $f : S^1 \to \mathbb{R}$ is smooth, (6.3) implies $f(U)$ is self-adjoint. The following is key to extending $f \mapsto f(U)$ to all continuous f.

Proposition 6.1. *If $f \in C^\infty(S^1)$ and $f : S^1 \to [0,\infty)$, then $f(U) \geq 0$, i.e.,*

$$(f(U)v, v) \geq 0, \quad \forall\, v \in H. \tag{6.10}$$

Proof. For $\varepsilon > 0$, set $f_\varepsilon(\zeta) = f(\zeta) + \varepsilon$, so there exists smooth $g_\varepsilon(\zeta) = f_\varepsilon(\zeta)^{1/2},\ g_\varepsilon : S^1 \to (0,\infty)$. Hence

$$f(U) + \varepsilon = f_\varepsilon(U) = g_\varepsilon(U)^2, \quad \text{and } g_\varepsilon(U) \text{ is self-adjoint.} \tag{6.11}$$

Hence

$$((f(U)+\varepsilon)v, v) = (g_\varepsilon(U)^2 v, v) = \|g_\varepsilon(U)v\|^2 \geq 0, \tag{6.12}$$

for all $\varepsilon > 0$, which yields (6.10).

Remark. Given $f_\varepsilon : S^1 \to (0,\infty)$ and $g_\varepsilon = f_\varepsilon^{1/2}$, it is clear that

$$f_\varepsilon \in C^\infty(S^1) \Longrightarrow g_\varepsilon \in C^\infty(S^1). \tag{6.13}$$

The result

$$f_\varepsilon \in \mathcal{A}(S^1) \Longrightarrow g_\varepsilon \in \mathcal{A}(S^1) \tag{6.14}$$

is also true, but harder to prove.

From here we obtain the following crucial norm estimate. The proof is like that of Proposition 4.2, but to make this section self-contained, we include it here.

Proposition 6.2. *If $f \in C^\infty(S^1)$,*

$$\sup_{S^1} |f| = M \Longrightarrow \|f(U)\| \leq M. \tag{6.15}$$

Proof. We have

$$g = M^2 - |f|^2 \geq 0, \quad g \in C^\infty(S^1), \tag{6.16}$$

so Proposition 6.1 implies $g(U) \geq 0$. Hence, for all $v \in H$,

$$\begin{aligned} \|f(U)v\|^2 &= (|f|^2(U)v, v) \\ &= (M^2 v, v) - (g(U)v, v) \\ &\leq M^2 \|v\|^2, \end{aligned} \tag{6.17}$$

giving (6.15).

Corollary 6.3. *Given $U \in \mathcal{L}(H)$ unitary, the map $f \mapsto f(U)$ has a unique continuous extension from $\{f \in C^\infty(S^1)\}$ to $\{f \in C(S^1)\}$. We have*

$$f, g \in C(S^1) \Longrightarrow (fg)(U) = f(U)g(U), \quad \overline{f}(U) = f(U)^*, \tag{6.18}$$

and

$$f \in C(S^1), \ |f| \le M \Longrightarrow \|f(U)\| \le M. \tag{6.19}$$

From here, the arguments of §5 give the Spectral Theorem for unitary operators. We will not repeat these arguments here, just the result.

Theorem 6.4. *If $U \in \mathcal{L}(H)$ is unitary, there exists a measure space $(X, \mathcal{F}, \mu)$, a unitary map*

$$\Phi : H \longrightarrow L^2(X, \mu), \tag{6.20}$$

and $u \in L^\infty(X, \mu)$ such that $|u| = 1$, μ-a.e. on X, and, for all $v \in H$,

$$\Phi U v(x) = u(x) \Phi v(x). \tag{6.21}$$

Furthermore, for all $f \in C(S^1)$,

$$\Phi f(U) v(x) = f(u(x)) \Phi v(x). \tag{6.22}$$

Remark. The proof of the Spectral Theorem given here is a variant of a familiar approach using Fourier series, but other such approaches typically use nontrivial results from harmonic analysis, such as the Bochner-Herglotz theorem (cf. [P], Appendix). By comparison, the proof here is fairly elementary.

7. From Theorem 1.2 to Theorem 1.1

In §§1–5, we established the Spectral Theorems 1.1 and 1.2, for bounded self-adjoint $A \in \mathcal{L}(H)$ and unitary $U \in \mathcal{L}(H)$, making use of the holomorphic functional calculus. In §6, we gave an alternative proof of the Spectral Theorem 1.2 for unitary U, using Fourier series, but not holomorphic functional calculus. Here, we show how to obtain Theorem 1.1 from Theorem 1.2, also not using the holomorphic functional calculus.

So take a self-adjoint $A \in \mathcal{L}(H)$. As seen in Proposition 1.3, $A = A^* \Rightarrow \sigma(A) \subset \mathbb{R}$, so $A \pm i$ are invertible:

$$(A+i)^{-1},\ (A-i)^{-1} \in \mathcal{L}(H). \tag{7.1}$$

We define $U \in \mathcal{L}(H)$ to be the "Cayley transform":

$$U = (A+i)(A-i)^{-1}. \tag{7.2}$$

Note that $(A-i)(A-i)^{-1} = I = (A-i)^{-1}(A-i)$ implies $(A-i)^{-1}$ commutes with A, hence with $A+i$, so also

$$U = (A-i)^{-1}(A+i). \tag{7.3}$$

We claim that U is unitary. To see this, note that $(TS)^* = S^*T^*$ yields $(T^*)^{-1} = (T^{-1})^*$ if $T \in \mathcal{L}(H)$ is invertible. Hence (7.2) yields

$$\begin{aligned} U^* &= ((A-i)^{-1})^*(A+i)^* \\ &= (A+i)^{-1}(A-i), \end{aligned} \tag{7.4}$$

and comparison with (7.3) gives

$$U^*U = UU^* = I. \tag{7.5}$$

To invert the correspondence (7.2) (or (7.3)), note that $(A+i) = U(A-i) = (A-i)U$, hence

$$AU - A = i(U+1), \quad \text{or} \quad A(U-1) = i(U+1). \tag{7.6}$$

Lemma 7.1. *If $A \in \mathcal{L}(H)$ is self-adjoint and U is given by (7.2)–(7.3), then*

$$1 \in \rho(U). \tag{7.7}$$

Proof. We can rewrite (7.2) as

$$(7.8)\qquad \begin{aligned} U &= (A+i)(A+i)(A+i)^{-1}(A-i)^{-1} \\ &= (A+i)^2(A^2+1)^{-1}, \end{aligned}$$

since

$$(7.9)\qquad (A+i)^{-1}(A-i)^{-1} = ((A-i)(A+i))^{-1} = (A^2+1)^{-1}.$$

Hence

$$(7.10)\qquad U = (A^2-1)(A^2+1)^{-1} + 2iA(A^2+1)^{-1}.$$

The operators $A^2-1,\ A^2+1$, and A all commute, and so do A^2-1 and A with $(A^2+1)^{-1}$. Hence the operators

$$(7.11)\qquad (A^2-1)(A^2+1)^{-1} \ \text{ and } \ A(A^2+1)^{-1} \ \text{ are self-adjoint,}$$

and, for all $v \in H$,

$$(7.12)\qquad \begin{aligned} \operatorname{Re}(Uv,v) &= ((A^2-1)(A^2+1)^{-1}v,v) \\ &= \|v\|^2 - 2((A^2+1)^{-1}v,v), \end{aligned}$$

so

$$(7.13)\qquad \begin{aligned} \operatorname{Re}((I-U)v,v) &= 2((A^2+1)^{-1}v,v) \\ &= 2(w,(A^2+1)w) \\ &= 2\|w\|^2 + 2\|Aw\|^2 \\ &\geq 2\|w\|^2, \end{aligned}$$

where $w = (A^2+1)^{-1}v$. Now

$$(7.14)\qquad \|v\| = \|(A^2+1)w\| \leq C\|w\|,$$

with $C = \|A^2+1\| < \infty$, so (7.13) gives

$$(7.15)\qquad \operatorname{Re}((I-U)v,v) \geq \frac{2}{C^2}\|v\|^2,$$

which via Cauchy's inequality yields

$$(7.16)\qquad \|(I-U)v\| \geq \frac{2}{C^2}\|v\|.$$

Hence

$$I - U : H \longrightarrow H \quad \text{is injective, with closed range.} \tag{7.17}$$

Now, if $w \in H$ is orthogonal to the range of $I - U$, then, for all $v \in H$,

$$\begin{aligned} 0 = ((I-U)v, w) &= (v, (I - U^{-1})w) \\ &= -(v, U^{-1}(I-U)w) \\ &= -(Uv, (I-U)w). \end{aligned} \tag{7.18}$$

Since $U : H \to H$ is invertible, this implies that, for all $\tilde{v} \in H$,

$$(\tilde{v}, (I-U)w) = 0, \quad \text{hence} \quad (I-U)w = 0, \tag{7.19}$$

and (7.17) then implies $w = 0$. This proves Lemma 7.1.

Having Lemma 7.1, we deduce from (7.6) that

$$A = i(U+1)(U-1)^{-1}. \tag{7.20}$$

Note that $(U-1)^{-1}$ commutes with $U-1$, hence with U, hence with $U+1$, so also

$$A = i(U-1)^{-1}(U+1). \tag{7.21}$$

As a check, note that (7.20) implies

$$\begin{aligned} A^* &= -i((U-1)^{-1})^*(U+1)^* \\ &= -i(U^{-1}-1)^{-1}(U^{-1}+1), \end{aligned} \tag{7.22}$$

and $(U^{-1}-1)^{-1} = (1-U)^{-1}U$, so

$$A^* = i(U-1)^{-1}(U+1) = A, \tag{7.23}$$

as it should.

We now apply Theorem 1.2 (or equivalently, Theorem 6.4) to U. We have a measure space $(X, \mathcal{F}, \mu)$, a unitary map

$$\Phi : H \longrightarrow L^2(X, \mu), \tag{7.24}$$

and $u \in L^\infty(X, \mu)$ ($|u| = 1$, μ-a.e. on X), such that, for all $v \in H$,

$$\Phi U v(x) = u(x) \Phi v(x). \tag{7.25}$$

The result (7.7) implies $1 \in \rho(M_u)$, so there exists $\delta > 0$ such that

$$|u(x) - 1| \geq \delta, \quad \text{for } \mu\text{-a.e. } x \in X. \tag{7.26}$$

Then, as suggested by (7.20), we can set

$$a(x) = i\, \frac{u(x)+1}{u(x)-1}, \tag{7.27}$$

obtaining $a \in L^\infty(X, \mu)$. Calculations parallel to (7.22)–(7.23) give

$$\overline{a(x)} = a(x), \quad \text{for } \mu\text{-a.e. } x \in X, \tag{7.28}$$

and we get, for all $v \in H$,

$$\Phi A v(x) = a(x) \Phi v(x), \tag{7.29}$$

proving Theorem 1.1.

A. Spectral projections

The Spectral Theorem 1.1 says that each self-adjoint $A \in \mathcal{L}(H)$ is unitarily equivalent to a multiplication operator on a space $L^2(X,\mu)$:

$$A = \Phi^{-1} M_a \Phi, \quad \Phi : H \to L^2(X,\mu) \text{ unitary}, \tag{A.1}$$

with

$$M_a f(x) = a(x) f(x), \quad f \in L^2(X,\mu). \tag{A.2}$$

We obtain spectral projections as follows. Let $\Sigma \subset \mathbb{R}$ be a Borel set. Then $a^{-1}(\Sigma) \subset X$ is μ-measurable; take

$$\begin{aligned} \chi_\Sigma(x) = 1 \quad &\text{if} \ \ a(x) \in \Sigma, \\ 0 \quad &\text{if} \ \ a(x) \notin \Sigma, \end{aligned} \tag{A.3}$$

and set

$$Q_\Sigma f(x) = \chi_\Sigma(x) f(x), \quad f \in L^2(X,\mu). \tag{A.4}$$

Then each Q_Σ is an orthogonal projection on $L^2(X,\mu)$. We have

$$Q_\Sigma Q_\Gamma = Q_{\Sigma\cap\Gamma}, \tag{A.5}$$

and the assignment $\Sigma \mapsto Q_\Sigma$ is strongly countably additive, i.e.,

$$\Sigma = \bigcup_{j\ge 1} \Sigma_j \ \text{(disjoint union)}, \ f \in L^2(X,\mu) \Rightarrow Q_\Sigma f = \lim_{N\to\infty} \sum_{j=1}^{N} Q_{\Sigma_j} f, \tag{A.6}$$

the limit taken in the L^2-norm. This follows from the Lebesgue dominated convergence theorem.

If we set

$$Q_\lambda = Q_{(-\infty,\lambda]}, \tag{A.7}$$

then, for all $f \in L^2(X,\mu)$, we have the Stieltjes integral representations

$$f = \int dQ_\lambda \, f, \quad af = \int \lambda \, dQ_\lambda \, f. \tag{A.8}$$

Now we define a projection valued measure P_Σ on H by

$$P_\Sigma = \Phi^{-1} Q_\Sigma \Phi. \tag{A.9}$$

This is strongly countably additive, we have

$$P_\Sigma P_\Gamma = P_{\Sigma\cap\Gamma}, \tag{A.10}$$

and, if we set $P_\lambda = P_{(-\infty,\lambda]}$, then for all $v \in H$,

$$v = \int dP_\lambda \, v, \quad Av = \int \lambda \, dP_\lambda \, v. \tag{A.11}$$

This last formula is often given as one version of the Spectral Theorem.

B. Unbounded self-adjoint operators

Here, we give a sketchy discussion of results that can be found in more detail in §8 of [T1] and §1 of [T2]. Rather than a bounded operator $A : H \to H$, we consider a linear transformation defined on a dense linear subspace $\mathcal{D}(A)$ of A,

$$A : \mathcal{D}(A) \longrightarrow H. \tag{B.1}$$

We assume A is closed, i.e., its graph

$$\mathcal{G}_A = \{(v, Av) \in H \oplus H : v \in \mathcal{D}(A)\}$$

is a closed linear subspace of $H \oplus H$. We say A is symmetric if

$$(Av, w) = (v, Aw), \quad \forall\, v, w \in \mathcal{D}(A). \tag{B.2}$$

We say A is self-adjoint if, furthermore, whenever $v \in H$ has the property that there exists $C < \infty$ such that

$$|(v, Aw)| \le C\|w\|, \quad \forall\, w \in \mathcal{D}(A),$$

then in fact $v \in \mathcal{D}(A)$ and (B.2) holds.

It is the case (cf. [T1], §8), that a closed operator A is self-adjoint if and only if it is symmetric and

$$A + i,\ A - i : \mathcal{D}(A) \longrightarrow H \ \text{ are both bijective.} \tag{B.3}$$

In such a case,

$$U = (A + i)(A - i)^{-1} \tag{B.4}$$

belongs to $\mathcal{L}(H)$ and in fact is unitary. Hence the Spectral Theorem 1.2 applies to U. We have a unitary map

$$\Phi : H \longrightarrow L^2(X, \mu) \tag{B.5}$$

and $u \in L^\infty(X, \mu),\ |u| = 1, \mu$-a.e., such that

$$\Phi U \Phi^{-1} f(x) = u(x) f(x), \quad \forall\, f \in L^2(X, \mu). \tag{B.6}$$

We can go from there to a Spectral Theorem for the unbounded self-adjoint operator A, by a process somewhat parallel to §7, though with some differences in detail. We have

$$\Phi A \Phi^{-1} = \widetilde{A}, \tag{B.7}$$

with

$$\widetilde{A}f(x) = a(x)f(x), \quad a(x) = i\,\frac{u(x)+1}{u(x)-1}. \tag{B.8}$$

In this setting, a might not belong to $L^\infty(X,\mu)$, but in place of (7.26) we do have

$$u(x) - 1 \neq 0, \quad \text{for } \mu\text{-a.e. } x \in X. \tag{B.9}$$

We also have

$$\begin{aligned} &\Phi : \mathcal{D}(A) \longrightarrow \mathcal{D}(\widetilde{A}) \ \ \text{bijective}, \\ &\mathcal{D}(\widetilde{A}) = \{f \in L^2(X,\mu) : af \in L^2(X,\mu)\}. \end{aligned} \tag{B.10}$$

For more details, see [T2], §1.

C. Von Neumann's mean ergodic theorem

Let U be a unitary operator on a Hilbert space H, and consider

$$A_N = \frac{1}{N} \sum_{k=0}^{N-1} U^k, \tag{C.1}$$

i.e., $A_N = \varphi_N(U)$, where

$$\varphi_N(\zeta) = \frac{1}{N} \sum_{k=0}^{N-1} \zeta^k, \quad \zeta \in S^1. \tag{C.2}$$

Note that

$$\varphi_N(1) \equiv 1, \quad \varphi_N(\zeta) = \frac{1}{N} \frac{1-\zeta^N}{1-\zeta}, \quad \forall\, \zeta \in S^1 \setminus \{1\}. \tag{C.3}$$

Hence

$$|\varphi_N(\zeta)| \le 1, \quad \lim_{N\to\infty} \varphi_N(\zeta) = 0, \quad \forall\, \zeta \ne 1. \tag{C.4}$$

If $\Phi : H \to L^2(X,\mu)$ is the unitary operator given in Theorem 1.2, so (1.3) holds, then, for all $f \in L^2(X,\mu)$,

$$\begin{aligned} \Phi A_N \Phi^{-1} f(x) &= \varphi_N(u(x)) f(x) \\ &\to \chi(x) f(x), \quad \text{as } \ N \to \infty, \end{aligned} \tag{C.5}$$

in L^2-norm, where

$$\begin{aligned} \chi(x) = \ & 1 \quad \text{if } \ u(x) = 1, \\ & 0 \quad \text{if } \ u(x) \ne 1. \end{aligned} \tag{C.6}$$

Convergence in $L^2(X,\mu)$ follows from the Lebesgue dominated convergence theorem. Now $Qf = \chi f$ defines Q as the orthogonal projection of $L^2(X,\mu)$ onto $\operatorname{Ker}(I - M_u)$. It follows that, for each $v \in H$,

$$A_N v \longrightarrow Pv \ \text{ in } \ H\text{-norm}, \tag{C.7}$$

where

$$P = \Phi^{-1} Q \Phi \ \text{ is the orthogonal projection of } H \text{ onto } \operatorname{Ker}\,(I - U). \tag{C.8}$$

We formally state the result just derived.

Proposition C.1. *If $U : H \to H$ is unitary, then, for each $v \in H$, we have convergence in H-norm,*

$$\text{(C.9)} \quad \frac{1}{N} \sum_{k=0}^{N-1} U^k v \longrightarrow Pv, \quad P = \textit{ orthogonal projection of } H \textit{ onto } \operatorname{Ker}(I - U).$$

There are other proofs of this, which do not use the Spectral Theorem, and which moreover extend the scope to include operators $U : H \to H$ that are isometries, but are not necessarily invertible, hence not unitary. For such results, and applications to ergodic theory, see [P], Chapter 1, or [T4], Chapter 14. The proof given above is closer to von Neumann's original proof.

References

[P] W. Parry, Topics in Ergodic Theory, 2nd ed., Cambridge Univ. Press, 2004.

[T1] M. Taylor, Outline of functional analysis, Appendix A in *Partial Differential Equations, Vol. 1*, Springer-Verlag, New York, 1996 (2nd ed., 2011).

[T2] M. Taylor, Spectral theory, Chapter 8 in *Partial Differential Equations, Vol. 2*, Springer-Verlag, New York 1996 (2nd ed., 2011).

[T3] M. Taylor, Introduction to Differential Equations, American Mathematical Society, Providence, RI, 2011.

[T4] M. Taylor, Measure Theory and Integration, American Mathematical Society, Providence, RI, 2006.