UNIVERSITÀ DI PISA

Dipartimento di Matematica
Corso di Laurea Triennale in Matematica

Anno Accademico 2023 - 2024

Tesi di Laurea

A Non-commutative Extension of the Sewing Lemma with Applications to Quantum Mechanics

Relatore
Prof. Dario Trevisan

Candidato
Samuele Biscaro

Table of Contents

Introduction

The Sewing Lemma is a powerful tool in the study of rough paths and differential equations, widely known for its role in constructing generalized integrals in cases where classical techniques are insufficient due to low regularity.

While the results presented in this thesis are well established in the literature, they do not appear to have been applied in the specific context of quantum computing before. This thesis explores such an application, focusing on the extension of the Sewing Lemma to the non-commutative setting, which is essential for dealing with problems in quantum mechanics, where matrix-valued and operator-valued functions dominate.

In particular, the non-commutative extension developed here offers new tools for addressing irregular Hamiltonians. While much of the existing theory assumes smooth or at least differentiable operators, our approach allows for a less regular setting, broadening the scope of applicable models in quantum mechanics.

A significant part of this thesis is dedicated to outlining the technical aspects of this extension and discussing its potential for advancing (growing) fields such as quantum computing.

Although quantum machine learning will not be treated here, it is worth mentioning that this rapidly growing field will likely benefit from the mathematical techniques introduced in this work, especially as quantum algorithms increasingly deal with low-regularity operators.

CHAPTER 1

The Additive Sewing Lemma

The *Sewing Lemma* [Gub04, FdLP06] was introduced as a powerful analytical tool to study integration when dealing with functions of low regularity . It allows for the unique definition of integrals of the form

$$I_t = \int_0^t X_s \, dY_s,$$

in cases where both X and Y are not necessarily smooth. For example, in the so-called *Young regime*, where X and Y have Hölder regularities α and β, respectively, with $\alpha + \beta > 1$, a classical result by Young [You36] and Kondurar [Kon37] establishes that a well-defined integration theory exists. The Sewing Lemma generalizes this result, proving the existence and uniqueness of a function $I : [0, T] \to \mathbb{R}$ such that

$$|I_b - I_a - Y_a(X_b - X_a)| \leq \mathbf{c} \cdot |b - a|^{\alpha+\beta}.$$

This provides a rigorous foundation for integration in low-regularity settings, enabling its use in various applications.

1.1 Control Functions

Before introducing the Sewing Lemma, it is necessary to define a more general notion of modulus of continuity.

Definition 1.1.1. We say that a real valued function $V(t)$ defined on $[0, T]$ is a *control function* if it is non decreasing, $V(0) = 0$ and

$$\sum_{n \geq 1} V(1/n) < +\infty. \tag{1.1.1}$$

Remark 1.1.2. The final condition is equivalent to

$$\overline{V}(t) = \sum_{k\geq 0} 2^k \cdot V(2^{-k} \cdot t) < +\infty, \tag{1.1.2}$$

which is a more convenient form, as we shall see.

Also, we have

$$\overline{V}(t) = V(t) + \cdots + 2^n V(2^n \cdot t) + 2^{n+1}\overline{V}(2^{-n-1} \cdot t)$$

from which follows that

$$\lim_{n\to+\infty} 2^n \cdot \overline{V}(2^{-n} \cdot t) = 0 \tag{1.1.3}$$

Example 1.1.3. A common control function is $V(t) = t^\alpha$ with $\alpha > 1$, in fact in this case:

$$\begin{aligned}\overline{V}(t) &= \sum_{k\geq 0} 2^k \cdot V(2^{-k} \cdot t) = \sum_{k\geq 0} 2^k \cdot t^\alpha \cdot 2^{-\alpha k} \\ &= t^\alpha \cdot \sum_{k\geq 0} 2^{(1-\alpha)k} < +\infty\end{aligned}$$

since $1 - \alpha < 0$.

Example 1.1.4. We can also consider the function defined on $[0, 1]$

$$V(t) = \frac{t}{(\ln(t^{-1}))^\alpha},$$

with $\alpha > 1$. Obviously $V(0) = 0$ and by taking the derivative we get

$$\alpha \ln(t^{-1})^{-\alpha-1} + \ln(t^{-1})^{-\alpha} > 0 \quad \forall t \in [0, 1],$$

so V is also increasing. Lastly,

$$\begin{aligned}\overline{V}(t) &= \sum_{k\geq 0} 2^k \cdot V(2^{-k} \cdot t) = \sum_{k\geq 0} 2^k \cdot \frac{2^{-k} \cdot t}{(\ln(2^k \cdot t^{-1}))^\alpha} \\ &= t \cdot \sum_{k\geq 0} \frac{1}{(k \ln 2 + \ln(t^{-1}))^\alpha} < +\infty.\end{aligned}$$

1.2 The Additive Sewing Lemma

Theorem 1.2.1. Consider a continuous function $\mu(a, b)$ with real values defined for $0 \leq a \leq b \leq T$, satisfying the relation

$$|\mu(a, b) - \mu(a, c) - \mu(c, b)| \leq V(b - a)$$

for every $c \in [a, b]$, where V is a control function. Then, there exists a unique function $\phi(t)$ on $[0, T]$, up to an additive constant, such that

$$|\phi(b) - \phi(a) - \mu(a, b)| \leq \overline{V}(b - a).$$

We will prove this result in the course of this section.

First, let $\mu_0(a,b) = \mu(a,c) + \mu(c,b)$ for $c = (a+b)/2$, and define recursively $\mu_{n+1}(a,b) = \mu_n(a,c) + \mu_n(c,b)$.

Lemma 1.2.2. Let μ_n as before, then for $n \geq 0$, we obtain

$$|\mu_n(a,b) - \mu_{n+1}(a,b)| \leq 2^n \cdot V(2^{-n} \cdot |b-a|).$$

Proof. We show this by induction. For $n = 0$, this is exactly the assumed relation. For $n > 0$, we have:

$$\begin{aligned}
|\mu_n(a,b) - \mu_{n+1}(a,b)| &= |\mu_{n-1}(a,c) + \mu_{n-1}(c,b) - \mu_n(a,c) - \mu_n(c,b)| \\
&\leq 2^{n-1} \cdot V(2^{-n+1} \cdot |c-a|) + 2^{n-1} \cdot V(2^{-n+1} \cdot |b-c|) \\
&= 2^{n-1} \cdot V\left(2^{-n+1} \cdot \frac{|b-a|}{2}\right) + 2^{n-1} \cdot V\left(2^{-n+1} \cdot \frac{|b-a|}{2}\right) \\
&= 2^n \cdot V(2^{-n} \cdot |b-a|).
\end{aligned}$$

□

From this, it follows that

Proposition 1.2.3. The sequence μ_n converges uniformly to a limit u.

Proof. Because of Lemma 1.2.2, the series

$$\sum_{n \geq 0} |\mu_n(a,b) - \mu_{n+1}(a,b)| \leq \overline{V}(b-a)$$

converges, so

$$\mu_n(a,b) = \sum_{k=0}^{n-1} \mu_{k+1}(a,b) - \mu_k(a,b)$$

converges to a limit $u(a,b)$. Since V is non-decreasing, we have uniform convergence because for every $0 \leq a \leq b \leq T$:

$$\begin{aligned}
|\mu_n(a,b) - u(a,b)| &\leq \sum_{k \geq n} |\mu_k(a,b) - \mu_{k+1}(a,b)| \\
&\leq \sum_{k \geq n} 2^k \cdot V(2^{-k} \cdot (b-a)) \\
&\leq \sum_{k \geq n} 2^k \cdot V(2^{-k} \cdot T),
\end{aligned}$$

which tends to zero since the sum is convergent.

□

Note that, since μ is continuous, so are all the μ_n, then the limit u is also continuous because we have uniform convergence.

Also, for $c = (a+b)/2$, $\mu_{n+1}(a,b) = \mu_n(a,c) + \mu_n(c,b)$, which implies

$$u(a,b) = u(a,c) + u(c,b).$$

We say that u is *midpoint-additive*.

Also, we have that

$$\begin{aligned} |u(a,b) - \mu(a,b)| &\leq \sum_{n\geq 0} |\mu_n(a,b) - \mu_{n+1}(a,b)| \\ &= \overline{V}(b-a). \end{aligned}$$

This function u is actually the only function satisfying this properties, indeed we have

Proposition 1.2.4. Let $v(a,b)$ be a midpoint-additive function such that for every $0 \leq a \leq b \leq T$:

$$|v(a,b) - \mu(a,b)| \leq \mathbf{c} \cdot \overline{V}(b-a),$$

then $v = u$.

Where, here and in the following, $\mathbf{c}$ will denote a costant.

Proof. Let v be such a function, then we obtain for a constant K that

$$|v(a,b) - u(a,b)| \leq K \cdot \overline{V}(b-a).$$

By induction, assuming

$$|v(a,b) - u(a,b)| \leq 2^{n-1} K \cdot \overline{V}(2^{-n+1} \cdot (b-a))$$

we have

$$\begin{aligned} |v(a,b) - u(a,b)| &= |v(a,c) + v(c,b) - u(a,c) - u(c,b)| \\ &\leq 2^{n-1} K \cdot \overline{V}\left(2^{-n+1}(c-a)\right) \\ &\quad + 2^{n-1} K \cdot \overline{V}(2^{-n+1}(b-c)) \\ &\leq 2^n K \cdot \overline{V}(2^{-n}(b-a)), \end{aligned}$$

which vanishes as $n \to \infty$, as shown in (1.1.3), so that $v = u$.

□

We will now show that the function u is *additive*, meaning that for every $a \leq c \leq b$:

$$u(a,c) + u(c,b) = u(a,b).$$

To do so, let k be an integer with $k \geq 3$, and define the function

$$w(a,b) = \sum_{i=0}^{k-1} u(t_i, t_{i+1}),$$

where $t_i = a + i \cdot \frac{b-a}{k}$.

Proposition 1.2.5. The function w is midpoint-additive and satisfies

$$|w(a,b) - u(a,b)| \leq \mathbf{c} \cdot \overline{V}(b-a). \tag{1.2.1}$$

Proof. To prove midpoint-additivity, set $s_i = a + i \cdot \frac{b-a}{2k}$, for $c = \frac{a+b}{2}$ we have

$$\begin{aligned} w(a,c) + w(c,b) &= \sum_{i=0}^{k-1} u(s_i, s_{i+1}) + \sum_{i=k}^{2k-1} u(s_i, s_{i+1}) \\ &= \sum_{i=0}^{2k-1} u(s_i, s_{i+1}) = \sum_{i=0}^{k-1} u(t_i, t_{i+1}) \\ &= w(a,b). \end{aligned}$$

In the last step, we used the fact that u is midpoint-additive. Thus, w is also midpoint-additive.

For proving (1.2.1), we have

$$\begin{aligned} |w(a,b) - \mu(a,b)| &= |\sum_{i=0}^{k-1} u(t_i, t_{i+1}) - \mu(a,b)| \\ &\leq |\sum_{i=0}^{k-1} u(t_i, t_{i+1}) - \sum_{i=0}^{k-1} \mu(t_i, t_{i+1})| \\ &\quad + |\sum_{i=0}^{k-1} \mu(t_i, t_{i+1}) - \mu(a,b)|. \end{aligned}$$

For the first term, we get

$$|\sum_{i=0}^{k-1} u(t_i, t_{i+1}) - \sum_{i=0}^{k-1} \mu(t_i, t_{i+1})| \leq k \cdot \overline{V}(\frac{b-a}{i}) \leq \mathbf{c}_k \cdot \overline{V}(b-a).$$

While for the second one, we proceed by induction. We know that for $c \in [a,b]$ it is true that $|\mu(a,b) - \mu(a,c) - \mu(c,b)| \leq V(b-a)$, suppose that for every a, b:

$$|\mu(a,b) - \sum_{i=0}^{k-2} \mu(t_i, t_{i+1}) - \mu(t_k, b)| \leq \mathbf{c}_{k-1} \cdot \overline{V}(b-a),$$

then

$$\begin{aligned}|\mu(a,b) - \sum_{i=0}^{k-1} \mu(t_i, t_{i+1})| &= |\mu(a,b) - \sum_{i=0}^{k-2} \mu(t_i, t_{i+1}) - \mu(t_{k-1}, b)| \\ &\leq |\mu(a,b) - \mu(a, t_{k-1}) - \mu(t_{k-1}, b)| \\ &\quad + |\mu(a, t_{k-1}) - \sum_{i=0}^{k-2} \mu(t_i, t_{i+1})| \\ &\leq \mathbf{c} \cdot \overline{V}(b-a) + \mathbf{c}_{k-1} \cdot \overline{V}(t_{k-1} - a) \leq \mathbf{c}_k \cdot \overline{V}(b-a).\end{aligned}$$

Hence,

$$|w(a,b) - \mu(a,b)| \leq \mathbf{c}_k \cdot \overline{V}(b-a).$$

□

This implies that $w = u$, meaning that u is actually *rationally additive*. Since u is continuous, it is additive over the entire interval. Therefore, we can define

$$\phi(t) = u(0,t).$$

This shows that $\phi(t)$ is the desired function, uniquely determined up to an additive constant, and this concludes the proof of Theorem 1.2.1.

Note that in this proof we also shown that

Corollary 1.2.6. Let μ as in Theorem 1.2.1, then there exist a unique additive function $u(a,b)$ such that

$$|u(a,b) - \mu(a,b)| \leq \mathbf{c} \cdot \overline{V}(b-a)$$

for every $a \leq b$.

The same proof works in general for a function μ with values in a Banach space X, in particular for a matrix-valued function.

1.3 The Young Integral

To appreciate the importance of this result we will derive some important conclusions from *Young's integration theory*. To do so, we will need the following proposition about *Riemann sums*.

Proposition 1.3.1. Let $\sigma = \{t_i\}$ be a finite subdivision of $[a,b]$. Define $\delta = \sup_i |t_{i+1} - t_i|$. Then

$$\lim_{\delta \to 0} \sum_i \mu(t_i, t_{i+1}) = \phi(b) - \phi(a),$$

where μ and ϕ are as above.

Proof. We have

$$\phi(b) - \phi(a) - \sum_i \mu(t_i, t_{i+1}) = \sum_i [\phi(t_{i+1}) - \phi(t_i) - \mu(t_i, t_{i+1})].$$

Using the inequality

$$|\phi(t_{i+1}) - \phi(t_i) - \mu(t_i, t_{i+1})| \leq \overline{V}(t_{i+1} - t_i),$$

we get

$$\left|\phi(b) - \phi(a) - \sum_i \mu(t_i, t_{i+1})\right| \leq \sum_i \overline{V}(t_{i+1} - t_i).$$

Since $\overline{V}(\delta)/\delta \leq \epsilon$ as $\delta \to 0$, we can bound the sum:

$$\sum_i \overline{V}(t_{i+1} - t_i) \leq \epsilon \sum_i (t_{i+1} - t_i) = \epsilon(b - a).$$

Therefore,

$$\left|\phi(b) - \phi(a) - \sum_i \mu(t_i, t_{i+1})\right| \leq \epsilon(b - a).$$

As $\delta \to 0$, also $\epsilon \to 0$, which implies

$$\lim_{\delta \to 0} \sum_i \mu(t_i, t_{i+1}) = \phi(b) - \phi(a).$$

□

We are now ready to define the *Young Integral*.

Let x and y be two real valued α-Hölder continuous functions on $[0, T]$, with $\alpha > 1/2$. We can define

$$\mu(a, b) = x_a(y_b - y_a),$$

so that

$$\mu(a, b) - \mu(a, c) - \mu(c, b) = -(x_c - x_a)(y_b - y_c),$$

which leads to

$$|\mu(a, b) - \mu(a, c) - \mu(c, b)| \leq \|x\|_\alpha \|y\|_\alpha |b - a|^{2\alpha},$$

where $\|x\|_\alpha$ denotes the norm in the space $\mathcal{C}^\alpha$ (see Section A.2).
The function $V(t) = t^{2\alpha}$, since $2\alpha > 1$, is indeed a control function, as in example 1.1.3. This means that the additive sewing lemma applies, giving us a function ϕ. We can then define the integral as

$$\int_a^b x_t \, dy_t = \phi(b) - \phi(a),$$

so that

$$\int_a^b x_t \, dy_t = \lim_{\delta \to 0} \sum_i x_{t_i}(y_{t_{i+1}} - y_{t_i})$$

where t_i is a partition of $[a, b]$ and $\delta = \sup(t_{i+1} - t_i)$. This integral is known as the Young integral.

Remark 1.3.2. It is also possible to take $x \in \mathcal{C}^\alpha$ and $y \in \mathcal{C}^\beta$ as long as $\alpha + \beta > 1$, but this would unnecessarily complicate the notation, since we will only consider the case $\alpha = \beta$ in this thesis.

Proposition 1.3.3. Given $w, x, y, z \in \mathcal{C}^\alpha([0, T], \mathbb{R})$ and $\gamma, \delta \in \mathbb{R}$, applying the Sewing Lemma to

$$\mu(a, b) = \delta w_a(x_b - x_a) + \gamma y_a(z_a - z_b)$$

yields

$$\delta \int_a^b w_t \, dx_t + \gamma \int_a^b y_t \, dz_t. \tag{1.3.1}$$

To simplify the notation, we will often use the following shorthand:

$$\int_a^b (\delta w_t \, dx_t + \gamma y_t \, dz_t).$$

Proof. By the definition of the Young integral, we know that

$$\left| \int_a^b w_t \, dx_t - w_a(x_b - x_a) \right| \leq \mathbf{c} \cdot |b - a|^{2\alpha},$$

and similarly,

$$\left| \int_a^b y_t \, dz_t - y_a(z_b - z_a) \right| \leq \mathbf{c} \cdot |b - a|^{2\alpha}.$$

Combining these inequalities, we get

$$\begin{aligned} \left| \delta \int_a^b w_t \, dx_t + \gamma \int_a^b y_t \, dz_t - (\delta w_a(x_b - x_a) + \gamma y_a(z_b - z_a)) \right| \leq \\ \leq \delta \left| \int_a^b w_t \, dx_t - w_a(x_b - x_a) \right| \\ + \gamma \left| \int_a^b y_t \, dz_t - y_a(z_b - z_a) \right| \\ \leq \mathbf{c} \cdot |b - a|^{2\alpha}. \end{aligned}$$

Since the expression in (1.3.1) is additive, by the Sewing Lemma it is the unique solution constructed from the increments $\mu(a, b)$. □

An obvious application of proposition 1.3.3 is that the functions

$$x \mapsto \int_a^b x_t \, dy_t \quad \text{and} \quad y \mapsto \int_a^b x_t \, dy_t$$

are linear.

Note that even for matrix-valued functions

$$A \in \mathcal{C}^\alpha([0,T], \mathbb{R}^{d\times d}) \quad \text{and} \quad B \in \mathcal{C}^\alpha([0,T], \mathbb{R}^{d\times d}),$$

we can consider the increment $\mu(a,b) = A_a(B_b - B_a)$, and the Sewing Lemma applies as before. This allows us to define the integral

$$\int_a^b A_t \, dB_t. \tag{1.3.2}$$

The integral we have just defined is quite natural. In fact, we have the following result:

Proposition 1.3.4. Given A and B as above, for all $i, j \in \{1, \dots, d\}$, the components of the matrix-valued integral satisfy

$$\left(\int_a^b A_t \, dB_t\right)^{ij} = \int_a^b \sum_{k=1}^d A_t^{ik} \, dB_t^{kj}.$$

Proof. Let M be the matrix that, for some $a < b$, has entries

$$M^{ij} = \int_a^b (A_t \, dB_t)^{ij},$$

where the right hand side is to be intended as the function that we get applying the Sewing Lemma to

$$(A_a(B_b - B_a))^{ij}.$$

Because of Proposition 1.3.3 it is clear that

$$M^{ij} = \int_a^b (A_t \, dB_t)^{ij} = \int_a^b \sum_{k=1}^d A_t^{ik} \, dB_t^{kj}.$$

We only need to show that

$$M = \int_a^b A_t \, dB_t.$$

This is another simple application of the Sewing Lemma, because for every i, j we know that

$$\left| \int_a^b (A_t \, dB_t)^{ij} - (A_a(B_b - B_a))^{ij} \right| \leq \mathbf{c} \cdot |b-a|^{2\alpha},$$

then

$$|M - A_a(B_b - B_a)| = \sum_{i,j} \left| \int_a^b (A_t \, dB_t)^{ij} - (A_a(B_b - B_a))^{ij} \right| \\ \leq \mathbf{c} \cdot |b - a|^{2\alpha},$$

and since M is additive we obtain exactly the above equality. □

Remark 1.3.5. Of course, we do not need to restrict ourselves to square matrices; the only requirement is that A and B are compatible for multiplication. This more general setting introduces no additional complexity beyond a slight increase in notational burden.

Now that we are working with matrices, they may not commute. In fact, by applying the Sewing Lemma to the function

$$\mu(a, b) = (B_b - B_a)A_a,$$

we generally obtain a function that is different from (1.3.2). To denote this expression, we will use the notation

$$\int_a^b dB_t \, A_t.$$

We can obtain a result analogous to Proposition 1.3.4 for this integral as well, and the proof follows the same reasoning.

1.4 Properties of the Young Integral

We will now derive some important properties of the Young integral, which will be useful in the following sections.

Proposition 1.4.1. Let $A, B \in \mathcal{C}^\alpha([0, T], \mathbb{R}^{d \times d})$. Then the functions

$$I_t = \int_0^t A_s \, dB_s \quad \text{and} \quad J_t = \int_0^t dB_s \, A_s$$

are α-Hölder continuous.

Proof. We need to consider, for $a < b$,

$$|I_b - I_a| \leq |I_b - I_a - A_a(B_b - B_a)| + |A_a(B_b - B_a|).$$

By the definition of the Young integral. For the first term, we have:

$$|I_b - I_a - A_a(B_b - B_a)| \leq \mathbf{c} \cdot |b - a|^{2\alpha}.$$

For the second term, using the continuity of the function A, which implies it is bounded, we obtain:

$$\begin{aligned}|A_a(B_b - B_a)| &\leq |A_a| \cdot |B_b - B_a| \\ &\leq \sup_{0\leq s\leq T} |A_s| \cdot |B_b - B_a| \\ &\leq \mathbf{c} \cdot |b-a|^{\alpha}.\end{aligned}$$

Combining the above results, we find:

$$\begin{aligned}|I_b - I_a| &\leq \mathbf{c} \cdot (|b-a|^{2\alpha} + |b-a|^{\alpha}) \\ &\leq \mathbf{c} \cdot (T^{\alpha}|b-a|^{\alpha} + |b-a|^{\alpha}) \\ &\leq \mathbf{c} \cdot |b-a|^{\alpha}.\end{aligned}$$

In an analogous way, we can derive the same result for J. □

The next property is also known as the *transitivity* of the Young integral.

Proposition 1.4.2. Let $A, B, C \in \mathcal{C}^{\alpha}([0,T], \mathbb{R}^{d\times d})$, and

$$I_t = \int_0^t B_s \, dC_s.$$

Then

$$\int_0^t A_s \, dI_s = \int_0^t A_s B_s \, dC_s.$$

Proof. We just need to prove that

$$\left| \int_a^b A_s \, dI_s - A_a B_a (C_b - C_a) \right| \leq \mathbf{c} \cdot |b-a|^{2\alpha}$$

By definition

$$\left| \int_a^b A_s \, dI_s - A_a (I_b - I_a) \right| \leq \mathbf{c} \cdot |b-a|^{2\alpha},$$

and

$$\begin{aligned}|A_a(I_b - I_a) - A_a B_a (C_b - C_a)| &\leq |A_a| \cdot |I_b - I_a - B_a(C_b - C_a)| \\ &\leq \mathbf{c} \cdot |b-a|^{2\alpha},\end{aligned}$$

where we used that A is bounded and the definition of I_t. Putting all togheter, we get exacty what we wanted. Then, since

$$\int_a^b A_s \, dI_s$$

is additive, by the Sewing Lemma it is indeed equal to

$$\int_0^t A_s B_s \, dC_s.$$

□

Also, in a similar fashion, we can get

Proposition 1.4.3. Let $A, B, C \in \mathcal{C}^\alpha([0,T], \mathbb{R}^{d\times d})$, and

$$I_t = \int_0^t dA_s \, B_s.$$

Then

$$\int_0^t dI_s \, C_s = \int_0^t dA_s \, B_s C_s.$$

The last property we will prove is what is called the *Integration by Parts formula*, and it states that

Theorem 1.4.4. Let $A, B \in \mathcal{C}^\alpha([0,T], \mathbb{R}^{d\times d})$, then

$$A_b B_b = A_a B_a + \int_a^b dA_s \, B_s + \int_a^b A_s \, dB_s$$

Proof. We need to show that

$$A_b B_b - A_a B_a = \int_a^b dA_s \, B_s + \int_a^b A_s \, dB_s.$$

The function on the left-hand side is additive, so by Proposition 1.3.3 it suffices to show that

$$|A_b B_b - A_a B_a - ((A_b - A_a)B_a - A_a(B_b - B_a))| \leq \mathbf{c} \cdot |b-a|^{2\alpha}.$$

The left-hand side simplifies to $(A_b - A_a)(B_b - B_a)$, which completes the proof.

□

CHAPTER 2

The Multiplicative Sewing Lemma

The *Multiplicative Sewing Lemma* [FdLPM08] is a non-commutative extension of the classical Sewing Lemma, designed to handle integration in settings where the objects involved do not commute.

In analogy to the classical case the Multiplicative Sewing Lemma constructs a *product integral*, allowing us to define expressions of the form

$$U_t = \prod_0^t e^{dA_s},$$

where A is typically a matrix or operator valued function of low regularity.

This lemma extends the applicability of the classical Sewing Lemma to non-commutative settings, offering a robust framework to study the evolution of non-commutative processes, including applications in areas like quantum mechanics, as we shall see.

2.1 Strong Control Functions

Here we need a strong notion of control function.

Definition 2.1.1. We say that a function $V(t)$ defined on $[0, T]$ is a *strong control function* if it is a control function (see Definition 1.1.1) and there exists a $\theta > 2$ such that for every t

$$\overline{V}(t) = \sum_{k \geq 0} \theta^k \cdot V(2^{-k} \cdot t) < \infty.$$

Note that also in this case

$$\overline{V}(t) = V(t) + \cdots + \theta^n \cdot V \cdot (t \cdot 2^{-n}) + \theta^{n+1} \cdot \overline{V}(2^{-n-1} \cdot t),$$

which means that

$$\lim_{n\to+\infty} \theta^n \cdot \overline{V}(2^{-n}\cdot t) = 0.$$

Example 2.1.2. The function $V(t) = t^\alpha$, with $\alpha > 1$, is a strong control function.

We already saw in Example 1.1.3 that it is a control function. We can then choose $0 < \beta < \alpha - 1$, and by setting $\theta = 2^{1+\beta} > 2$, we get:

$$\begin{aligned}\overline{V}(t) &= \sum_{k\geq 0} \theta^k \cdot V(2^{-k}\cdot t) = \sum_{k\geq 0} 2^{(1+\beta)k}\cdot 2^{-\alpha k}\cdot t^\alpha \\ &= t^\alpha \cdot \sum_{k\geq 0} 2^{(1-\alpha+\beta)k} < +\infty\end{aligned}$$

since $1 - \alpha + \beta < 0$.

2.2 The Multiplicative Sewing Lemma

We consider an *associative monoid* $\mathcal{M}$ with a unit element I, that is a *complete metric space* under a distance d. Let also

$$|\cdot| : \mathcal{M} \to \mathbb{R}$$

be a Lipschitz function such that $|I| = 1$.

We also assume that d that satisfies the following conditions for every $x, y, z \in \mathcal{M}$:

$$d(xz, yz) \leq |z|\, d(x, y), \quad d(zx, zy) \leq |z|\, d(x, y). \tag{2.2.1}$$

Let $\mu(a, b)$ be an $\mathcal{M}$-valued function defined for $0 \leq a \leq b \leq T$. We assume that μ is continuous, that $\mu(a, a) = I$ for every a, and that for every $a \leq c \leq b$ we have

$$d(\mu(a, b), \mu(a, c)\mu(c, b)) \leq V(b - a) \tag{2.2.2}$$

We say that an $\mathcal{M}$-valued function $u(a, b)$ is *multiplicative* if for every $a \leq c \leq b$ we have $u(a, b) = u(a, c)u(c, b)$.

The Multiplicative Sewing Lemma states:

Theorem 2.2.1. Let μ as above satisfying (2.2.2), then there exists a unique multiplicative function u such that

$$d(\mu(a, b), u(a, b)) \leq \mathbf{c} \cdot \overline{V}(b - a)$$

for every $a \leq b$.

We will prove the theorem by breaking it down into the following lemmas.
We start by setting $\mu_0 = \mu$. By induction, we define

$$\mu_{n+1}(a,b) = \mu_n(a,c)\mu_n(c,b) \quad \text{where} \quad c = \frac{a+b}{2}.$$

Next, we introduce the functions

$$h_n(t) = \sup_{b-a\leq t} |\mu_n(a,b)|, \quad \text{and} \quad U_n(t) = \sup_{b-a\leq t} d(\mu_{n+1}(a,b), \mu_n(a,b)).$$

Lemma 2.2.2. The functions h_n and U_n are continuous and non decreasing, with $h_n(0) = 1$ and $U_n(0) = 0$.

Proof. We start by noting that $h_n(t)$ and $U_n(t)$ are defined as suprema over a bounded interval, ensuring that they are non decreasing.

To show continuity, note that since $\mu_n(a,b)$ is continuous in its arguments and the supremum of continuous functions over a compact set is continuous, $h_n(t)$ is continuous, an analogous argument shows that $U_n(t)$ is continuous.

For $h_n(0)$, since $\mu_{n+1}(a,a) = \mu_n(a,a) = I$ for every $n \in \mathbb{N}$ and for every $a \in \mathcal{M}$, we have

$$\sup_{a\in\mathcal{M}} |\mu_n(a,a)| = |I| = 1,$$

thus $h_n(0) = 1$. Similarly, for $U_n(0)$ we get

$$d(\mu_{n+1}(a,a), \mu_n(a,a)) = 0,$$

so $U_n(0) = 0$. □

Lemma 2.2.3. For the functions h_n and U_n the following inequalities hold:

1. $h_{n+1}(t) \leq h_n(t) + \kappa U_n(t)$, where κ is the Lipschitz constant of the map $z \mapsto |z|$.
2. $U_{n+1}(t) \leq [h_n(t/2) + h_{n+1}(t/2)]U_n(t/2)$.

Proof. We will prove the two inequalities individually.

1. By definition of Lipschitzianity of $|\cdot|$ we have

$$||\mu_{n+1}(a,b)| - |\mu_n(a,b)|| \leq \kappa d(\mu_{n+1}(a,b), \mu_n(a,b))$$

If $|\mu_{n+1}(a,b)| \geq |\mu_n(a,b)|$ then

$$|\mu_{n+1}(a,b)| \leq |\mu_n(a,b)| + \kappa d(\mu_{n+1}(a,b), \mu_n(a,b)),$$

otherwise

$$|\mu_{n+1}(a,b)| \leq |\mu_n(a,b)| \leq |\mu_n(a,b)| + \kappa d(\mu_{n+1}(a,b), \mu_n(a,b)).$$

Taking the supremum on both sides we get the desired inequality. Note also that iterating we get:

$$h_{n+1}(t) \leq h_0(t) + \kappa U_0(t) + \cdots + \kappa U_n(t).$$

2. We have, by expanding the definitions and using the triangular inequality that

$$\begin{aligned} d(\mu_{n+1}(a,b), \mu_n(a,b)) &= d(\mu_n(a,c)\mu_n(c,b), \mu_{n-1}(a,c)\mu_{n-1}(c,b)) \\ &\leq d(\mu_n(a,c)\mu_n(c,b), \mu_{n-1}(a,c)\mu_n(c,b)) \\ &\quad + d(\mu_{n-1}(a,c)\mu_n(c,b), \mu_{n-1}(a,c)\mu_{n-1}(c,b)). \end{aligned}$$

Using (2.2.1) the first term is bounded by

$$|\mu_n(c,b)| \cdot d(\mu_n(a,c), \mu_{n-1}(a,c)),$$

and also the second one by

$$|\mu_{n-1}(a,c)| \cdot d(\mu_n(c,b), \mu_{n-1}(c,b)).$$

Taking the supremum on both sides we get exactly

$$U_{n+1}(t) \leq [h_n(t/2) + h_{n+1}(t/2)]U_n(t/2).$$

□

Lemma 2.2.4. The sequence h_n is bounded and the series $\sum_{n\geq 0} U_n$ converges uniformly on $[0,T]$.

Proof. Since $\theta > 2$ we may take $\tau > 0$ such that $h_0(\tau) + \kappa\overline{V}(\tau) \leq \theta/2$.

Assume that for $t \leq \tau$ and $i \leq n$ the following inequality holds:

$$h_i(t) \leq \theta/2,$$

$$U_i(t) \leq \theta^i \cdot V(t/2^i).$$

Then we have using the last lemma

$$h_{n+1}(t) \leq h_0(t) + \kappa U_0(t) + \cdots + \kappa U_n(t) \leq h_0(t) + \kappa\overline{V}(t) \leq \theta/2$$

and

$$\begin{aligned} U_{n+1}(t) &\leq [h_n(t/2) + h_{n+1}(t/2)]U_n(t/2) \\ &\leq \theta \cdot U_n(t/2) \leq \theta^{n+1} \cdot V(t/2^{n+1}) \end{aligned}$$

for $t \leq \tau$ and every n by induction.

Thus, for $t \leq \tau$, the series

$$\sum_{n\geq 0} U_n(t) \leq \sum_{n\geq 0} \theta^n V(t \cdot 2^{-n}) = \overline{V}(t) < \infty$$

converges, indicating that the sequence $h_n(\tau)$ is bounded.

Using the second inequality from the last lemma, also the series

$$\sum_{n\geq 0} U_n(2\tau)$$

converges, meaning the sequence $h_n(2\tau)$ is also bounded. By proceeding step-by-step, we get that the sequence h_n is bounded, and the series $\sum_{n\geq 0} U_n$ converges uniformly on $[0, T]$.

□

Consequently, since $\mathcal{M}$ is complete, the sequence $\mu_n(a, b)$ converges uniformly to a continuous function $u(a, b)$.

This function $u(a, b)$ has the property that $u(a, b) = u(a, c)u(c, b)$ for any midpoint $c = (a + b)/2$, as before we say that u is *midpoint-multiplicative*.

Finally, we have the inequality

$$d(u(a, b), \mu(a, b)) \leq \mathbf{c} \cdot \overline{V}(b - a). \tag{2.2.3}$$

As in the additive sewing lemma, we will first establish the uniqueness of this function, and once uniqueness has been demonstrated, we will then proceed to show that the function is also multiplicative.

Lemma 2.2.5. Let v be a continuous function that is midpoint multiplicative and such that for every $0 \leq a \leq b \leq T$

$$d(v(a, b), \mu(a, b)) \leq \mathbf{c} \cdot \overline{V}(b - a)$$

then $v = u$.

Proof. Let v be such a function. Define $K(t)$ as follows:

$$K(t) = \sup_{b-a\leq t} \max\left(|u(a, b)|, |v(a, b)|\right)$$

Let $\tau > 0$ be such that $K(\tau) \leq \theta/2$. Given that

$$d(u(a, b), v(a, b)) \leq k\overline{V}(b - a)$$

for some constant k, we then have:

$$\begin{aligned} d(u(a,b),v(a,b)) &= d(u(a,c)u(c,b),v(a,c)v(c,b)) \\ &\leq d(u(a,c)u(c,b),u(a,c)v(c,b)) \\ &\quad + d(u(a,c)v(c,b),v(a,c)v(c,b)) \\ &\leq |u(a,c)|\cdot d(u(c,b),v(c,b)) \\ &\quad + |v(c,b)|\cdot d(u(a,c),v(a,c)) \\ &\leq 2K(t/2)k\overline{V}(t/2) \leq k\theta\overline{V}(t/2) \end{aligned}$$

for $b-a \leq t \leq \tau$. If we assume

$$d(u(a,b),v(a,b)) \leq k\theta^{n-1}\overline{V}(t\cdot 2^{-n+1})$$

for $b-a \leq t \leq \tau$, then, as before:

$$\begin{aligned} d(u(a,b),v(a,b)) &\leq |u(a,c)|\cdot d(u(c,b),v(c,b)) \\ &\quad + |v(c,b)|\cdot d(u(a,c),v(a,c)) \\ &\leq 2K(t/2)k\theta^{n-1}\overline{V}(2^{-n+1}\cdot t/2) \\ &\leq k\theta^{n}\overline{V}(t\cdot 2^{-n}) \end{aligned}$$

for every n by induction.

Using the fact that $\lim_n \theta^n\cdot\overline{V}(2^{-n}\cdot t)=0$, we get that $u(a,b)=v(a,b)$ for all $b-a\leq\tau$.

Finally, this equality can be extended to any $b-a$ using the midpoint-multiplicativity property.

□

We argue now as in the additive case, defining for $k\geq 3$ the function

$$w(a,b)=\prod_{i=0}^{k-1} u(t_i,t_{i+1}),$$

where $t_i = a+i\cdot\frac{b-a}{k}$. In a manner similar to the additive case, it follows that

Lemma 2.2.6. The function w is midpoint-multiplicative and satisfies

$$d(w(a,b),\mu(a,b)) \leq \mathbf{c}_k\overline{V}(b-a). \tag{2.2.4}$$

Proof. The fact that w is midpoint multiplicative follows exactly as in the additive case. We will show the second statement by induction on k. Note that for $k=2$ the relation is exactly the (2.2.3).

Assume now that (2.2.4) holds for $k-1$, then

$$\begin{aligned} d(w(a,b),\mu(a,b)) &\leq d(w(a,b),\mu(a,t_{k-1})\mu(t_{k-1},b)) \\ &\quad + d(\mu(a,t_{k-1})\mu(t_{k-1},b),\mu(a,b)). \end{aligned}$$

Which, expanding the definition and using again the triangular inequality, is less or equal than

$$\begin{aligned} & d(\prod_{i=0}^{k-2} u(t_i, t_{i+1})\ u(t_{k-1}, b), \prod_{i=0}^{k-2} u(t_i, t_{i+1})\ \mu(t_{k-1}, b)) \\ & \quad + d(\prod_{i=0}^{k-2} u(t_i, t_{i+1})\ \mu(t_{k-1}, b), \mu(a, t_{k-1})\mu(t_{k-1}, b)) \\ & \quad + d(\mu(a, t_{k-1})\mu(t_{k-1}, b), \mu(a, b)). \end{aligned}$$

The first term of the sum is, using (2.2.1),

$$\begin{aligned} |\prod_{i=0}^{k-2} u(t_i, t_{i+1})|\ d(u(t_{k-1}, b), \mu(t_{k-1}, b)) &\leq \mathbf{c} \cdot \overline{V}(b - t_{k-1}) \\ &\leq \mathbf{c} \cdot \overline{V}(b - a), \end{aligned}$$

while the second one, using also the induction hypothesis is

$$\begin{aligned} |\mu(t_{k-1}, b)|\ d(\prod_{i=0}^{k-2} u(t_i, t_{i+1}), \mu(a, t_{k-1})) &\leq \mathbf{c}_{k-1} \cdot \overline{V}(t_{k-1} - a) \\ &\leq \mathbf{c}_{k-1} \cdot \overline{V}(b - a). \end{aligned}$$

Lastly, we can bound the last term using (2.2.2) and get finally

$$d(w(a, b), \mu(a, b)) \leq \mathbf{c}_k \cdot \overline{V}(b - a).$$

□

Therefore, it follows that $w = u$ meaning that u is, in fact, rationally multiplicative, and, since it is continous, it is also multiplicative.
This concludes the proof of the theorem.

Corollary 2.2.7. Note that if ν is a function with the same properties as μ, and if it satisfies $d(\nu(a, b), \mu(a, b)) \leq \mathbf{c} \cdot \overline{V}(b - a)$ for every $a \leq b$, then ν will define the same multiplicative function u as μ.

2.3 The Integral Product

As in the additive case, we have a result similar to a *"Riemann product"*.

Proposition 2.3.1. Let $\sigma = \{t_i\}$ be a finite subdivision of $[a, b]$. Define $\delta = \sup_i |t_{i+1} - t_i|$. Then

$$\lim_{\delta \to 0} \prod_i \mu(t_i, t_{i+1}) = u(a, b).$$

Proof. We have

$$d(u(a,b), \prod_i \mu(t_i, t_{i+1})) = d(\prod_i u(t_i, t_{i+1}), \prod_i \mu(t_i, t_{i+1})).$$

Using the inequality

$$|u(t_i, t_{i+1}) - \mu(t_i, t_{i+1})| \leq \mathbf{c} \cdot \overline{V}(t_{i+1} - t_i),$$

and using the fact that for $w, x, y, z \in \mathcal{M}$

$$\begin{aligned} d(wx, yz) &\leq d(wx, wz) + d(wz, yz) \\ &\leq |w| d(x, z) + |z| d(w, y) \leq \mathbf{c} \cdot (d(x, z) + d(w, y)), \end{aligned}$$

we get

$$d(u(a,b), \prod_i \mu(t_i, t_{i+1})) \leq \mathbf{c} \cdot \sum_i \overline{V}(t_{i+1} - t_i).$$

Since $\overline{V}(\delta)/\delta \leq \epsilon$ as $\delta \to 0$, we can bound the sum:

$$\sum_i \overline{V}(t_{i+1} - t_i) \leq \epsilon \sum_i (t_{i+1} - t_i) = \epsilon(b - a).$$

Therefore,

$$d(u(a,b), \prod_i \mu(t_i, t_{i+1})) \leq \mathbf{c} \cdot \epsilon(b - a).$$

As $\delta \to 0$, also $\epsilon \to 0$, which implies

$$\lim_{\delta \to 0} \sum_i \mu(t_i, t_{i+1}) = u(a, b).$$

□

Let $t \to A_t$ a $\mathcal{C}^\alpha$ function with values in a Banach Algebra $\mathcal{A}$ with a unit element I. Put $A_{ab} = A_b - A_a$ and

$$\mu(a, b) = I + A_{ab}.$$

Obviously $\mu(a, a) = I$ and we get for $a < c < b$

$$\begin{aligned} \mu(a,b) - \mu(a,c)\mu(c,b) &= I + A_{ab} - (I + A_{ac})(I + A_{cb}) \\ &= A_{ab} - A_{ac} - A_{cb} - A_{ac}A_{cb} \\ &= -A_{ac}A_{cb}. \end{aligned}$$

Then

$$d(\mu(a,b), \mu(a,c)\mu(c,b)) = \| - A_{ac}A_{cb}\| \leq \|A\|_\alpha^2 \cdot (b - a)^{2\alpha}.$$

Therefore, since $V(t) = t^{2\alpha}$ is a strong control function, the multiplicative sewing lemma applies, giving us a function u such that

$$d(u(a,b), \mu(a,b)) \leq \mathbf{c} \cdot (b-a)^{2\alpha}.$$

Because of the the last proposition, a good notation for this function is

$$u(a,b) = \prod_a^b (I + dA_s).$$

We can also take $\nu(a,b) = e^{A_{ab}}$, it satisfies $\nu(a,a) = I$ and

$$\begin{aligned}
e^{A_{ab}} - e^{A_{ac}} e^{A_{cb}} &= e^{A_b - A_a} - e^{A_c - A_a} e^{A_b - A_c} \\
&= \sum_{k\geq 0} \frac{1}{k!}(A_b - A_a)^k \\
&\quad - \left(\sum_{k\geq 0} \frac{1}{k!}(A_c - A_a)^k \right) \cdot \left(\sum_{k\geq 0} \frac{1}{k!}(A_b - A_c)^k \right) \\
&= \sum_{k\geq 2} \frac{1}{k!}(A_b - A_a)^k \\
&\quad - \left(\sum_{k\geq 2} \frac{1}{k!}(A_c - A_a)^k \right) \cdot \left(\sum_{k\geq 2} \frac{1}{k!}(A_b - A_c)^k \right) \\
&\quad - (A_c - A_a)(A_b - A_c).
\end{aligned}$$

We can bound this expression using the fact that

$$\begin{aligned}
\| \sum_{k\geq 2} \frac{1}{k!}(A_b - A_a)^k \| &\leq \sum_{k\geq 2} \frac{1}{k!} \|A\|_\alpha^k (b-a)^{\alpha k} \\
&= \sum_{k\geq 2} \frac{1}{k!} \|A\|_\alpha^k (b-a)^{\alpha(k-2)} (b-a)^{2\alpha} \\
&= \sum_{k\geq 2} \frac{1}{k!} \|A\|_\alpha^k \cdot T^{\alpha(k-2)} \ (b-a)^{2\alpha} \\
&\leq \mathbf{c} \cdot (b-a)^{2\alpha}
\end{aligned}$$

so that

$$d(\nu(a,b), \nu(a,c)\nu(c,b)) \leq \mathbf{c} \cdot (b-a)^{2\alpha}.$$

Also, in a similar way

$$\begin{aligned}
e^{A_{ab}} - (I + A_{ab}) &= \sum_{k\geq 0} \frac{1}{k!}(A_b - A_a)^k - I - A_b + A_a \\
&= \sum_{k\geq 2} \frac{1}{k!}(A_b - A_a)^k,
\end{aligned}$$

so we get that

$$d(\nu(a,b),\mu(a,b)) \leq \mathbf{c}\cdot(b-a)^{2\alpha}.$$

Because of Corollary 2.2.7, we get the same u by using ν instead of μ. Therefore, we can also write $u(a,b)$ as:

$$u(a,b)=\prod_a^b(I+dA_t)=\prod_a^b e^{dA_t}.$$

Theorem 2.3.2. Let $U_t=u(0,t)$. Then this is a solution to the ODE:

$$U_t = I+\int_0^t U_s\,dA_s \tag{2.3.1}$$

where the integral is taken in the Young sense.

Proof. Since $U_0=u(0,0)=I$, we only need to show that U_t is the function that we get using the additive sewing lemma on $\xi(a,b)=U_aA_{ab}$. This means we only have to verify that

$$\|U_b-U_a-U_aA_{ab}\|\leq \mathbf{c}\cdot(b-a)^{2\alpha}.$$

The left hand side is worth

$$\begin{aligned}\|U_b-U_a-U_aA_{ab}\| &= \|u(0,b)-u(0,a)-u(0,a)A_{ab}\|\\ &= \|u(0,a)\cdot(u(a,b)-I-A_{ab})\|\\ &\leq \|u(0,a)\|\cdot\|u(a,b)-\mu(a,b)\|\\ &\leq \mathbf{c}\cdot(b-a)^{2\alpha}\end{aligned}$$

so we are done.

□

We will study better this equation in Chapter 4.

2.4 A Trotter Type Formula

Let $A,B\in\mathcal{C}^\alpha([0,T],\mathcal{A})$ as before, and define now

$$\mu(a,b)=(I+A_{ab})(I+B_{ab}).$$

This function satisfies, for $a\leq c\leq b$:

$$\begin{aligned}d(\mu(a,b),\mu(a,c)\mu(c,b)) &= \|\mu(a,b)-\mu(a,c)\mu(c,b)\|\\ &= \|(I+A_{ab})(I+B_{ab})\\ &\quad -(I+A_{ac})(I+B_{ac})(I+A_{cb})(I+B_{cb})\|,\end{aligned}$$

and it is easy to verify that

$$d(\mu(a,b),\mu(a,c)\mu(c,b)) \leq \mathbf{c}\cdot|b-a|^{2\alpha}.$$

As before, we can apply the Multiplicative Sewing Lemma and obtain the function

$$u(a,b)=\prod_a^b (I+dA_t)(I+dB_t)=\prod_a^b e^{dA_t}e^{dB_t}.$$

By Proposition 2.3.1 we can write

$$u(a,b)=\lim_{n\to+\infty}\prod_{i=0}^{2^n-1} e^{A_{t_it_{i+1}}}e^{B_{t_it_{i+1}}},$$

for

$$t_i=a+i\cdot\frac{b-a}{2^n}.$$

Defining $C=A+B$ we can observe that the function

$$\nu(a,b)=I+C_{ab}$$

is such that

$$d(\nu(a,b),\mu(a,b))\leq \mathbf{c}\cdot|b-a|^{2\alpha}$$

so that they define the same u. By applying Theorem 2.3.2 we then have that

$$u(0,t)=I+\int_0^t u(0,s)\,dC_s. \tag{2.4.1}$$

In particular, we can fix $A,B\in\mathcal{A}$ and apply what we said to the functions $t\mapsto tA$ and $t\mapsto tB$, which are obviously α-Hölder continuous whith $\alpha=1$. We get a function $u(a,b)$, and by (2.4.1)

$$u(0,t)=e^{t(A+B)}.$$

With $t=1$ we get the classical Lie-Trotter formula:

$$\begin{aligned}e^{A+B}=u(0,1)&=\lim_{n\to+\infty}\prod_{i=0}^{2^n-1} e^{A_{t_it_{i+1}}}e^{B_{t_it_{i+1}}}\\&=\lim_{n\to+\infty}\prod_{i=0}^{2^n-1} e^{A/2^n}e^{B/2^n}\\&=\lim_{n\to+\infty}\left(e^{A/2^n}e^{B/2^n}\right)^{2^n}.\end{aligned}$$

CHAPTER 3

An Introduction to Quantum Computing

Quantum computing studies how information can be stored and processed using systems governed by quantum mechanical laws. Unlike classical computing, which relies on *bits*, quantum computing uses *qubits*, exploiting quantum phenomena such as *superposition* and *entanglement* to perform computations more efficiently for certain problems. This field has significant implications not only for computational speed and cryptography but also for our fundamental understanding of reality.

Quantum computing is of particular interest in this work because it requires manipulating operators on Hilbert spaces, which are central objects in the mathematical formulation of quantum mechanics. In particular, our study of the multiplicative sewing lemma provides a framework to generalize the behavior of quantum operations governed by more irregular Hamiltonians.

Throughout this chapter, we will follow the presentation of quantum computing from the book *Mathematics of Quantum Computing* [Sch19], which offers a rigorous mathematical foundation for the concepts discussed.

3.1 Basic Notions of Quantum Mechanics

Quantum mechanics is a theory that predicts the statistical behavior of microscopic objects (such as electrons, protons, atoms) and often has implications for macroscopic phenomena. Measurements on these objects yield real-number outcomes, and repeated measurements on identically prepared systems reveal that the values are distributed around a *mean*, following a relative *frequency distribution*. The mathematical formula-

tion of quantum mechanics will be presented through four *foundational postulates*, and we will only focus on the so called *"pure states"*. For the notation and mathematical objects used throughout this chapter, we refer the reader to Appendix A.

Postulate 1 (Observable and Pure States)**.** An *observable*, i.e., a physically measurable quantity of a quantum system, is represented by a self-adjoint operator on a Hilbert space $\mathbb{H}$. If the preparation of a statistical ensemble is such that, for any observable represented by its self-adjoint operator A, the mean value of the observable can be calculated using a vector $|\psi\rangle \in \mathbb{H}$ satisfying $\|\psi\| = 1$ as:

$$\langle A\rangle_\psi \coloneqq \langle\psi|A\psi\rangle. \tag{3.1.1}$$

The preparation of the system is then said to be described by a *pure state*, represented by the vector $|\psi\rangle \in \mathbb{H}$. This vector is called the *state vector* or simply the *state*, and $\langle A\rangle_\psi$ is referred to as the (quantum mechanical) *expectation value* of the observable A in the pure state $|\psi\rangle$.

The space $\mathbb{H}$ is known as the *Hilbert space* of the quantum system, and often we will refer to a quantum system by its Hilbert space, for example, if a system S is described by $\mathbb{H}^S$, we will simply say "the system $\mathbb{H}^S$".

We will also require that

$$\langle \mathbf{1}\rangle_\psi = \|\psi\|^2 = 1,$$

for any state vector $|\psi\rangle$, because the operator $\mathbf{1}$ can be interpreted as the observable *"is there anything present?"*.

Using the diagonal representation of a self-adjoint operator A in terms of its eigenbasis, the expectation value of the observable A in a state ψ is given by

$$\langle A\rangle_\psi = \langle\psi|A|\psi\rangle = \sum_j \lambda_j |\langle e_j|\psi\rangle|^2,$$

where $\{\lambda_j\}$ are the eigenvalues of A and $\{e_j\}$ are its corresponding eigenvectors. The quantities $|\langle e_j|\psi\rangle|^2$ represent the probabilities of measuring the eigenvalue λ_j.

In finite-dimensional systems, the eigenvalues $\{\lambda_j\}$ of the self-adjoint operator A correspond to the possible measurement outcomes for the observable. If the spectrum is non-degenerate, the probabilities of obtaining each λ_j are precisely $|\langle e_j|\psi\rangle|^2$. This idea can be extended to infinite-dimensional systems, where the spectrum may include continuous parts, but in this context, we focus only on finite-dimensional systems. This concept is formalized in the following postulate.

Postulate 2. In a quantum system with Hilbert space $\mathbb{H}$, the possible measurement values of an observable are given by the spectrum $\sigma(A)$ of

the operator $A \in B_{\text{sa}}(\mathbb{H})$ associated with the observable. The probability $P_\psi(\lambda)$ that a measurement of the observable yields the eigenvalue λ of A for a quantum system in the pure state $|\psi\rangle \in \mathbb{H}$ is given by:

$$P_\psi(\lambda) = \|P_\lambda|\psi\rangle\|^2, \tag{3.1.2}$$

where P_λ is the projection onto the eigenspace $\text{Eig}(A, \lambda)$ of λ.

That (3.1.2) indeed defines a probability measure on the spectrum of A requires, in the general case, a technically demanding proof, so we will not show this.

As a consequence of (3.1.1), for any observable A and any complex number of the form $e^{i\alpha} \in \mathbb{C}$ with $\alpha \in \mathbb{R}$, we have

$$\langle A\rangle_{e^{i\alpha}\psi} = \langle e^{i\alpha}\psi|A|e^{i\alpha}\psi\rangle = \langle\psi|A|\psi\rangle = \langle A\rangle_\psi,$$

which means that the expectation values of the observable A are the same in the state $e^{i\alpha}|\psi\rangle$ as in the state $|\psi\rangle$.

Furthermore, since

$$|\langle e^{i\alpha}\psi|e_j\rangle|^2 = |\langle\psi|e_j\rangle|^2,$$

the measurement probabilities in the two states are also identical. This shows that the states $e^{i\alpha}|\psi\rangle$ and $|\psi\rangle$ are *physically indistinguishable*, meaning they represent the same quantum state. This leads us to the following definition

Definition 3.1.1. For every $|\psi\rangle \in \mathbb{H}$ with $\|\psi\| = 1$, the set

$$S_\psi := \{e^{i\alpha}|\psi\rangle \mid \alpha \in \mathbb{R}\}$$

is called a *ray* in $\mathbb{H}$ with $|\psi\rangle$ as a representative.

Every element of a ray S_ψ describes the same physical state, and the phase $\alpha \in \mathbb{R}$ can be chosen arbitrarily. More precisely, pure quantum states are described by a representative $|\psi\rangle$ of a ray S_ψ in the Hilbert space $\mathbb{H}$. In practice, only the symbol $|\psi\rangle$ of a representative of the ray is used, with the understanding that $|\psi\rangle$ and $e^{i\alpha}|\psi\rangle$ represent the same physical state.

Conversely, every unit vector in a Hilbert space $\mathbb{H}$ corresponds to a physical state, meaning that it captures the statistical properties of a quantum system. If $|\phi\rangle, |\psi\rangle \in \mathbb{H}$ are quantum states, then any linear combination $a|\phi\rangle + b|\psi\rangle$ with $a, b \in \mathbb{C}$ and $\|a|\phi\rangle + b|\psi\rangle\| = 1$ is also a valid quantum state. This is known as the quantum *superposition principle*: any normalized linear combination of quantum states is itself a state and is, in principle, physically realizable.

However, it is important to note that while the *global phase* of a state is physically irrelevant, the *relative phases* between terms in a linear combination are not. This will be made more clear by the following example.

Example 3.1.2. Let $|\phi\rangle$ and $|\psi\rangle \in \mathbb{H}$ be two *orthogonal states*, meaning $\langle\phi|\psi\rangle = 0$. Then the states $\frac{1}{\sqrt{2}}(|\phi\rangle + |\psi\rangle)$ and $\frac{1}{\sqrt{2}}(|\phi\rangle + e^{i\alpha}|\psi\rangle)$ are both normalized state vectors, but they correspond to different physical situations. In fact, for any observable A we have:

$$\begin{aligned}
\langle A\rangle_{\frac{|\phi\rangle+|\psi\rangle}{\sqrt{2}}} &= \frac{1}{2}\langle\phi+\psi|A\left(|\phi\rangle+|\psi\rangle\right)\rangle \\
&= \frac{1}{2}\left(\langle\phi|A\phi\rangle + \langle\psi|A\psi\rangle + \langle\phi|A\psi\rangle + \langle\psi|A\phi\rangle\right) \\
&= \frac{1}{2}\left(\langle\phi|A\phi\rangle + \langle\psi|A\psi\rangle + \langle\phi|A\psi\rangle + \langle A\psi|\phi\rangle\right) \\
&= \frac{1}{2}\left(\langle A\rangle_\phi + \langle A\rangle_\psi\right) + \mathrm{Re}(\langle\phi|A\psi\rangle).
\end{aligned}$$

Where the term $\mathrm{Re}(\langle\phi|A\psi\rangle)$ is often called the *interference term*, and this will be the term that is different in the case of $\frac{1}{\sqrt{2}}(|\phi\rangle + e^{i\alpha}|\psi\rangle)$. In fact, doing the same calcultaions, one gets

$$\langle A\rangle_{\frac{|\phi\rangle+e^{i\alpha}|\psi\rangle}{\sqrt{2}}} = \frac{1}{2}\left(\langle A\rangle_\phi + \langle A\rangle_\psi\right) + \mathrm{Re}(e^{i\alpha}\langle\phi|A\psi\rangle).$$

If a quantum system is prepared in the state $|\psi\rangle$, we can determine the likelihood of measuring it in the state $|\phi\rangle$ using the following proposition.

Proposition 3.1.3. Let the states of a quantum system be represented by rays in a Hilbert space $\mathbb{H}$. If the system is prepared in the state $|\psi\rangle \in \mathbb{H}$, the probability of observing it in the state $|\phi\rangle \in \mathbb{H}$ is given by

$$P\begin{pmatrix}\text{System prepared in state } |\psi\rangle \\ \text{is observed in state } |\phi\rangle\end{pmatrix} = |\langle\phi|\psi\rangle|^2.$$

Proof. Let $|\psi\rangle, |\phi\rangle \in \mathbb{H}$ with $\|\psi\| = \|\phi\| = 1$. The observable measured when determining if the system is in the state $|\phi\rangle$ is the orthogonal projection $P_\phi = |\phi\rangle\langle\phi|$ onto that state. This observable has eigenvalues 0 and 1. The eigenvalue $\lambda = 1$ is non-degenerate, and its eigenspace is spanned by $|\phi\rangle$. Therefore, the projection onto the eigenspace for eigenvalue $\lambda = 1$ is also given by P_ϕ. Thus, equation (3.1.2) from Postulate 2 becomes:

$$\begin{aligned}
P_\psi(\lambda = 1) = \|P_1|\psi\rangle\|^2 = \|P_\phi|\psi\rangle\|^2 = \||\phi\rangle\langle\phi|\psi\rangle\|^2 = \\
= \langle\phi|\psi\rangle^2\|\phi\|^2 = 1 = |\langle\phi|\psi\rangle|^2.
\end{aligned}$$

□

How widely are the measurement results distributed around their expectation value? This question is addressed by the concept of *uncertainty* or *standard deviation*, defined similarly to the corresponding notions in standard probability theory.

Definition 3.1.4. The *uncertainty* of an observable A in the state $|\psi\rangle$ is defined as

$$\Delta_\psi(A) \coloneqq \sqrt{\langle\psi|(A - \langle A\rangle_\psi \mathbf{1})^2\psi\rangle} = \sqrt{\langle(A - \langle A\rangle_\psi)^2\rangle_\psi}.$$

If the uncertainty vanishes, i.e., $\Delta_\psi(A) = 0$, we say that the value of the observable A in the state $|\psi\rangle$ is *sharp*. A sharp value of an observable A in a state $|\psi\rangle$ means that all measurements of A on systems in the state $|\psi\rangle$ will always yield the same result. This occurs if and only if $|\psi\rangle$ is an eigenvector of A, as stated in the following proposition.

Proposition 3.1.5. For any observable A and state $|\psi\rangle$, the following equivalence holds:

$$\Delta_\psi(A) = 0 \iff A|\psi\rangle = \langle A\rangle_\psi|\psi\rangle.$$

Proof. Since the observable A is self-adjoint it follows that $\langle A\rangle_\psi \in \mathbb{R}$. Consequently, we have:

$$(A - \langle A\rangle_\psi \mathbf{1})^* = A - \langle A\rangle_\psi \mathbf{1}.$$

From this, we can derive:

$$\begin{aligned}(\Delta_\psi(A))^2 &= \langle\psi|(A - \langle A\rangle_\psi \mathbf{1})^2\psi\rangle \\ &= \langle(A - \langle A\rangle_\psi \mathbf{1})\psi|(A - \langle A\rangle_\psi \mathbf{1})\psi\rangle.\end{aligned}$$

This leads to the expression:

$$\Delta_\psi(A) = 0 \iff A|\psi\rangle = \langle A\rangle_\psi|\psi\rangle.$$

In conclusion, the value of the observable A is sharp if and only if $|\psi\rangle$ is an eigenvector of A with eigenvalue $\langle A\rangle_\psi$.

□

A state that is an *eigenvector* of an operator associated with an observable is referred to as an *eigenstate* of that operator or observable.

A preparation in an eigenstate of A guarantees that all measurements of A conducted in that state will consistently yield the corresponding eigenvalue. Conversely, if the uncertainty of A vanishes for a given preparation, this indicates that the preparation is described by an eigenstate of A.

Definition 3.1.6. Two observables A and B are said to be *compatible* if the associated operators commute, i.e., if $[A, B] = 0$. Conversely, if $[A, B] \neq 0$, they are referred to as *incompatible*.

A result from linear algebra indicates that if A and B are self-adjoint and commute, $[A, B] = 0$, then there exists an orthonormal basis $\{|e_j\rangle\}$ in which both A and B are diagonal. Specifically, we can express the operators as:

$$A = \sum_j a_j |e_j\rangle\langle e_j| \quad \text{and} \quad B = \sum_j b_j |e_j\rangle\langle e_j|.$$

In a state $|e_k\rangle$, the system is then in an eigenstate of both A and B. Consequently, measurements of the compatible observables A and B in this state yield sharp results (the values a_k and b_k) for both observables, exhibiting no uncertainty.

On the other hand, the uncertainties of incompatible observables are subject to a lower bound, as demonstrated by the following proposition.

Proposition 3.1.7. For any observables $A, B \in \mathcal{B}_{sa}(\mathbb{H})$ and state $|\psi\rangle \in \mathbb{H}$, the following uncertainty relation holds:

$$\Delta_\psi(A)\Delta_\psi(B) \geq \left| \left\langle \frac{1}{2i}[A, B] \right\rangle_\psi \right|, \tag{3.1.3}$$

Proof. The relation is a consequence of the following estimates:

$$\begin{aligned}
\Delta_\psi(A)^2\Delta_\psi(B)^2 &= \|\langle A - \langle A\rangle_\psi \mathbf{1}\rangle_\psi\|^2 \, \|\langle B - \langle B\rangle_\psi \mathbf{1}\rangle_\psi\|^2 \\
&\geq |\langle (A - \langle A\rangle_\psi \mathbf{1})\psi | (B - \langle B\rangle_\psi \mathbf{1})\psi\rangle|^2 \\
&\geq \left(\operatorname{Im}\left(\langle (A - \langle A\rangle_\psi \mathbf{1})\psi | (B - \langle B\rangle_\psi \mathbf{1})\psi\rangle\right)\right)^2 \\
&= \Bigg(\frac{1}{2i}\langle (A - \langle A\rangle_\psi \mathbf{1})\psi | (B - \langle B\rangle_\psi \mathbf{1})\psi\rangle \\
&\quad - \frac{1}{2i}\overline{\langle (A - \langle A\rangle_\psi \mathbf{1})\psi | (B - \langle B\rangle_\psi \mathbf{1})\psi\rangle}\Bigg)^2 \\
&= \Bigg(\frac{1}{2i}\langle (A - \langle A\rangle_\psi \mathbf{1})\psi | (B - \langle B\rangle_\psi \mathbf{1})\psi\rangle \\
&\quad - \frac{1}{2i}\langle (B - \langle B\rangle_\psi \mathbf{1})\psi | (A - \langle A\rangle_\psi \mathbf{1})\psi\rangle\Bigg)^2 \\
&= \Bigg(\frac{1}{2i}\langle \psi | (A - \langle A\rangle_\psi \mathbf{1})(B - \langle B\rangle_\psi \mathbf{1})\psi\rangle \\
&\quad - \frac{1}{2i}\langle \psi | (B - \langle B\rangle_\psi \mathbf{1})(A - \langle A\rangle_\psi \mathbf{1})\psi\rangle\Bigg)^2 \\
&= \left(\left\langle \frac{1}{2i}[A - \langle A\rangle_\psi \mathbf{1}, B - \langle B\rangle_\psi \mathbf{1}] \right\rangle_\psi\right)^2 \\
&= \left(\left\langle \frac{1}{2i}[A, B] \right\rangle_\psi\right)^2 .
\end{aligned}$$

□

Example 3.1.8. The *Heisemberg uncertainty relation* can be viewed as a specific instance of equation (3.1.3). In this context, we consider $\mathbb{H} = L^2(\mathbb{R}^3)$, where A represents one of the position operators Q_j and B corresponds to one of the momentum operators P_j across the three spatial dimensions $j \in \{1, 2, 3\}$.

For these operators, their actions on states $|\psi\rangle$ in the Hilbert space $H = L^2(\mathbb{R}^3)$ are defined as follows[1]:

$$\begin{aligned} Q_j|\psi\rangle(x) &= x_j\psi(x), \\ P_j|\psi\rangle(x) &= -i\frac{\partial}{\partial x_j}\psi(x). \end{aligned}$$

From these definitions, we can derive the commutation relation:

$$\begin{aligned} [Q_j, P_k]|\psi\rangle(x) &= -ix_j\frac{\partial}{\partial x_k}\psi(x) - \left(-i\frac{\partial}{\partial x_k}(x_j\psi(x))\right) \\ &= i\delta_{jk}|\psi\rangle(x). \end{aligned}$$

Thus, we have $[Q_j, P_k] = i\delta_{jk}$. As a result, the uncertainty relation in this scenario is expressed as:

$$\Delta_\psi(Q_j)\Delta_\psi(P_k) \geq \frac{1}{2}\delta_{jk}.$$

A measurement of an observable $A = \sum_j \lambda_j|e_j\rangle\langle e_j|$ on an object in the state $|\psi\rangle = \sum_j |e_j\rangle\langle e_j|\psi\rangle$ yields an eigenvalue $\lambda_k \in \sigma(A)$. After this measurement, if no external interaction occurs, any subsequent measurement of A will always yield λ_k.

The object is now in a state where A has the sharp value λ_k, described by the eigenvector $|e_k\rangle$. Thus, the measurement *"projects"* the object from $|\psi\rangle$ into the eigenstate $|e_k\rangle$ with probability $|\langle e_k|\psi\rangle|^2$. This is known as the *Projection Postulate*.

Postulate 3. If a measurement of the observable A on a quantum mechanical system in the pure state $|\psi\rangle \in \mathbb{H}$ yields the eigenvalue λ, then the measurement causes a *state transition*. Specifically, the state $|\psi\rangle$ before the measurement transitions to the new state

$$\frac{P_\lambda|\psi\rangle}{\|P_\lambda|\psi\rangle\|},$$

where P_λ is the projection onto the eigenspace corresponding to λ. This new state represents the system immediately after the measurement.

[1] As always in this thesis, here the system of units with $\hbar = 1$ is used, since otherwise one would have for the momentum operators $P_j = -i\frac{\partial}{\partial x_j}$

Historically, the state $|\psi\rangle$ of a quantum mechanical system has been referred to as the *wave function*. For this reason, the Projection Postulate is also known as the *collapse of the wave function*.

A state can also evolve without measurement. The time evolution of a state, when no measurement is performed, is governed by a unitary operator, which is the solution to an operator initial value problem, as stated in the next postulete.

Postulate 4. In a quantum system with Hilbert space $\mathbb{H}$, every change of a pure state over time that is not caused by a measurement is described by the *time evolution operator* $U(t, t_0) \in \mathcal{U}(\mathbb{H})$.

Let $|\psi(t_0)\rangle$ be the state at time t_0 and $|\psi(t)\rangle$ be the state at time t. The time-evolved state $|\psi(t)\rangle$ originating from the initial state $|\psi(t_0)\rangle$ is given by:

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle.$$

The time evolution operator $U(t, t_0)$ is the solution of the initial value problem[2]:

$$\begin{aligned} i\frac{d}{dt}U(t, t_0) &= H(t)U(t, t_0), \\ U(t_0, t_0) &= I, \end{aligned} \tag{3.1.4}$$

where $H(t)$ is the self-adjoint *Hamiltonian operator*, which generates the time evolution of the quantum system.

Proposition 3.1.9. The operator $U(t, t_0)$ satisfying equation (3.1.4), with $H(t) \in \mathrm{B}_{\mathrm{sa}}(\mathbb{H})$ is unitary and unique.

[2]We remind the reader here once more that in this thesis we use natural physical units, such that $\hbar = 1$, which is why this constant does not appear as a factor on the left side of the equation.

Proof. To show the unitarity of $U(t, t_0)$ we can consider, given $|\psi\rangle \in \mathbb{H}$:

$$\begin{aligned}
\frac{d}{dt}||U(t,t_0)\psi||^2 &= \frac{d}{dt}\langle U(t,t_0)\psi|U(t,t_0)\psi\rangle \\
&= \left\langle \frac{d}{dt}U(t,t_0)\psi|U(t,t_0)\psi \right\rangle \\
&\quad + \left\langle U(t,t_0)\psi|\frac{d}{dt}U(t,t_0)\psi \right\rangle \\
&= \langle -iH(t)U(t,t_0)\psi|U(t,t_0)\psi\rangle \\
&\quad + \langle U(t,t_0)\psi| - iH(t)U(t,t_0)\psi\rangle \\
&= i\left(\langle H(t)U(t,t_0)\psi|U(t,t_0)\psi\rangle\right. \\
&\quad \left. -\langle U(t,t_0)\psi|H(t)U(t,t_0)\psi\rangle\right) \\
&= i\left(\langle H(t)^*U(t,t_0)\psi|U(t,t_0)\psi\rangle\right. \\
&\quad \left. -\langle U(t,t_0)\psi|H(t)U(t,t_0)\psi\rangle\right) \\
&= 0.
\end{aligned}$$

Then $||U(t, t_0)\psi||^2$ is costant, but also $||U(t_0, t_0)\psi||^2 = \|\psi\|^2$.

To show the uniqueness of the solution, we can suppose that $V(t, t_0)$ is another solution, then the same calculations with $U(t, t_0) - V(t, t_0)$ instead of $U(t, t_0)$ yields that $||(U(t, t_0) - V(t, t_0))\psi||^2$ is constant, but this time $||(U(t_0, t_0) - V(t_0, t_0))\psi||^2 = 0$, so that $U = V$.

□

In the Chaper 4, we will address the technical details regarding the existence of solutions $t \mapsto U(t, t_0)$. For the time being, we assume that the Hamiltonian $H(t)$ is such that a unique solution always exists.

The operator version of time evolution, as described in Postulate 4, is equivalent to the *Schrödinger equation*:

$$i\frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle,$$

which governs the time evolution of pure states through its action on state vectors. Applying the operator form to the state leads to the Schrödinger equation, and conversely, any solution of the Schrödinger equation for an initial state $|\psi(t_0)\rangle$ provides a solution for $U(t, t_0)$.

The operator $H(t)$ corresponds to the *energy* observable of the system, meaning that the expectation value $\langle H(t)\rangle_\psi$ gives the *expected energy* in state $|\psi\rangle$. When $H(t)$ is time-independent, the system's energy remains constant and is determined by the eigenvalues $\{E_j \,|\, j \in I\}$ of H.

The discreteness of energy levels for certain Hamiltonians is central to the notion of *"quantum"*, a concept introduced by Planck in his study of *black body radiation*. The Hamiltonian $H(t)$ not only represents the system's energy but also dictates its time evolution.

In quantum computing, gates are implemented as unitary operators V, which correspond to specific time evolutions $U(t, t_0)$ generated by a carefully chosen $H(t)$.

3.2 The Pauli Matrices

An important example of observables in quantum computing is the *spin* of an electron, which represents its *intrinsic angular momentum*. This spin is described by three observables S_x, S_y, S_z, collectively denoted as the *spin vector* $\mathbf{S} = (S_x, S_y, S_z)$. Since we are focusing on the spin only, the relevant Hilbert space is two-dimensional, $\mathbb{H} \cong \mathbb{C}^2$. The operators corresponding to the spin components S_j in this space are[3]:

$$S_j = \frac{1}{2}\sigma_j \quad \text{for } j \in \{x, y, z\},$$

where the σ_j are the *Pauli matrices*, defined as follows:

Definition 3.2.1. The Pauli matrices $\sigma_j \in \mathrm{Mat}(2 \times 2, \mathbb{C})$, indexed by $j \in \{1, 2, 3\}$ or equivalently by $j \in \{x, y, z\}$, are defined as:

$$\sigma_x := \sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y := \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z := \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

With $\sigma_0 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ denoting the 2×2 identity matrix, we extend the set to

$$\{\sigma_\alpha\}_{\alpha \in \{0,\dots,3\}} = \{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}.$$

In the context of a two-dimensional Hilbert space $\mathbb{H}$ with a chosen orthonormal basis (ONB), we use the notation

$$\sigma_0 = 1, \quad X = \sigma_1 = \sigma_x, \quad Y = \sigma_2 = \sigma_y, \quad Z = \sigma_3 = \sigma_z,$$

to refer to the corresponding operators in $L(\mathbb{H})$ with these matrix representations.

For the spin states defined as

$$|\uparrow_{\hat{z}}\rangle := |0\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow_{\hat{z}}\rangle := |1\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

the operator S_z has the following effects:

$$S_z|\uparrow_{\hat{z}}\rangle = \frac{1}{2}|\uparrow_{\hat{z}}\rangle, \quad S_z|\downarrow_{\hat{z}}\rangle = -\frac{1}{2}|\downarrow_{\hat{z}}\rangle.$$

[3] In non-natural units $\hbar$ would appear as a factor on the right side.

Thus, S_z possesses eigenvalues $\left\{\pm\frac{1}{2}\right\}$ with eigenvectors $\{|\uparrow_{\hat{z}}\rangle, |\downarrow_{\hat{z}}\rangle\}$, which correspond to the *up* and *down* spin states in the $\hat{z}$ direction. For convenience, we will use $\sigma_j = 2S_j$ as the observables, avoiding the factor of $\frac{1}{2}$.

The notation $|0\rangle$ and $|1\rangle$ for these eigenvectors reflects their association with classical bit values 0 and 1, a standard practice in quantum computing. A general state can be expressed as $a|0\rangle + b|1\rangle$ with $|a|^2 + |b|^2 = 1$. It's important to note that $|0\rangle$ is distinct from the null vector in Hilbert space.

The observable σ_z thus has eigenvalues ± 1 with eigenvectors $|0\rangle$ and $|1\rangle$, leading to expectation values:

$$\langle\sigma_z\rangle|0\rangle = \langle 0|\sigma_z|0\rangle = +1, \quad \langle\sigma_z\rangle|1\rangle = \langle 1|\sigma_z|1\rangle = -1.$$

In the state $|0\rangle$, the uncertainty in σ_z is calculated as follows:

$$\sigma_z - \langle\sigma_z\rangle\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix},$$

yielding

$$\langle 0|(\sigma_z - \langle\sigma_z\rangle|0\rangle)|0\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0.$$

Thus, we find that $\Delta_{|0\rangle}(\sigma_z) = 0$.

Similarly, for the state $|1\rangle$, $\Delta_{|1\rangle}(\sigma_z) = 0$ holds true as these states are eigenstates of σ_z, leading to no measurement uncertainty.

In contrast, σ_x and σ_z are incompatible operators since

$$[\sigma_x, \sigma_z] = -2i\sigma_y.$$

For $|0\rangle$:

$$\langle\sigma_x\rangle|0\rangle = 0 \quad \text{and} \quad \sigma_x - \langle\sigma_x\rangle\mathbf{1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - 0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

leading to

$$\Delta_{|0\rangle}(\sigma_x) = 1,$$

indicating that a measurement of σ_x in the state $|0\rangle$ has non-zero uncertainty. The same conclusion holds for $|1\rangle$ and measurements of σ_x. Consequently, σ_z and σ_x cannot be simultaneously measured with zero uncertainty, a fact that similarly applies to the pairs σ_z, σ_y and σ_x, σ_y.

3.3 Storing Information with Quantum Computers

A *classical bit* represents the smallest unit of information, corresponding to a choice between binary alternatives typically denoted as 0 and 1 (or Yes and No, True and False). Physically, a classical bit can be realized by assigning these alternatives to two distinct states of a physical system, such as *opposite magnetization* on a *hard disk*.

In *quantum computing*, these binary alternatives can be represented by two basis vectors in a quantum state space, which is often an infinite-dimensional Hilbert space. For practical purposes, we can restrict our attention to two-dimensional eigenspaces of appropriately chosen observables. We will now give examples of quantum systems with two-dimensional Hilbert spaces.

3.3.1 Electrons and their Spin

We can ignore the electron's position and momentum, focusing solely on its spin state. As seen before, the binary alternatives can be mapped to the eigenstates of σ_z:

$$|0\rangle = |\uparrow_{\hat{z}}\rangle, \quad |1\rangle = |\downarrow_{\hat{z}}\rangle.$$

This is not the only option, we could have equivalently choose the eigenstates of σ_x:

$$|+\rangle = |\uparrow_{\hat{x}}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_{\hat{z}}\rangle + |\downarrow_{\hat{z}}\rangle), \quad |-\rangle = |\downarrow_{\hat{x}}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_{\hat{z}}\rangle - |\downarrow_{\hat{z}}\rangle),$$

o the eigenstates of σ_y:

$$|\uparrow_{\hat{y}}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_{\hat{z}}\rangle + i|\downarrow_{\hat{z}}\rangle), \quad |\downarrow_{\hat{y}}\rangle = \frac{1}{\sqrt{2}}(i|\uparrow_{\hat{z}}\rangle + |\downarrow_{\hat{z}}\rangle).$$

3.3.2 Photons and Their Polarization

For photons propagating in a specific direction, the *polarization* is described by a two-dimensional complex vector known as the *polarization vector*. We can map the binary alternatives to the vectors:

$$|0\rangle = |H\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{(horizontal polarization)},$$

$$|1\rangle = |V\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{(vertical polarization)}$$

This vectors form an orthonormal basis and are the eigenvectors of the operator $\sigma_z = |H\rangle\langle H| - |V\rangle\langle V|$; the orthogonal projectors $|H\rangle\langle H|$ and $|V\rangle\langle V|$ are also called *horizontal* and *vertical polarizors*.

Another alternative as a base are the eigenstates of the so called *rotated polarizors*:

$$|+\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle), \quad |-\rangle = \frac{1}{\sqrt{2}}(|H\rangle - |V\rangle)$$

expressed as the operators $|+\rangle\langle+|$ and $|-\rangle\langle-|$.

One last commonly used base are the eigenstates of the left and right circular polarizors:

$$|R\rangle = \frac{1}{\sqrt{2}}(|H\rangle + i|V\rangle), \quad |L\rangle = \frac{1}{\sqrt{2}}(i|H\rangle + |V\rangle).$$

When representing classical bit values with, for example, electrons, we can prepare an electron in an eigenstate of σ_z, such as $|0\rangle$ for 0 and $|1\rangle$ for 1. If we isolate the electron from interactions to preserve its state, measuring σ_z will yield the corresponding eigenvalue, indicating the stored binary value.

Maintaining the integrity of the stored bit is crucial, as interactions that alter the electron's state could change the stored information. In classical computers, such as hard disks, external disturbances like light or heat generally do not affect stored bits, allowing for easier maintenance. In contrast, isolating quantum systems from state-changing interactions with their environment presents significant challenges, a key issue in developing quantum computers.

Thus, a classical bit can be represented by an orthonormal basis (ONB) in a two-dimensional Hilbert space, with the specific choice of ONB depending on a suitable observable whose eigenvectors correspond to the ONB vectors. Potential candidates for physical realizations include electrons and photons, but any quantum system with an appropriate two-dimensional Hilbert space can be utilized. Mathematically, we can identify the two-dimensional Hilbert space $\mathbb{H}$ with $\mathbb{C}^2$ by selecting an ONB.

3.4 Qubits

Quantum mechanics also permits states of the form $a|0\rangle + b|1\rangle$ where $a, b \in \mathbb{C}$ and $|a|^2 + |b|^2 = 1$. These linear combinations have no classical analogue and do not exist in classical computing, allowing for a significantly greater information storage capacity in two-dimensional quantum systems. The complexities of writing, reading, or transforming information in such systems necessitate special considerations, motivating the introduction of the term *"qubit"* to denote two-dimensional quantum systems in the context of their information content.

Definition 3.4.1. A *qubit* is a quantum mechanical system represented by a two-dimensional Hilbert space, denoted as ¶$\mathbb{H}$. The information in a qubit is stored in its state, which can be manipulated and read according to quantum mechanics.

In this space, we define an orthonormal basis $\{|0\rangle, |1\rangle\}$ and an observable represented by a self-adjoint operator σ_z. The operator σ_z has the eigenvector $|0\rangle$ with eigenvalue $+1$ and $|1\rangle$ with eigenvalue -1:

$$\sigma_z|0\rangle = +1|0\rangle, \quad \sigma_z|1\rangle = -1|1\rangle.$$

In classical computing, a bit serves as the fundamental unit of information, represented by the binary values {0, 1}. In quantum computing, the equivalent information container is a two-dimensional quantum system characterized by the Hilbert space ¶$\mathbb{H}$. The *"value"* of the quantum information is the state $|\psi\rangle \in$ ¶$\mathbb{H}$ in which the system is.

As a result of the Projection Postulate 3, we can derive the following corollary:

Corollary 3.4.2. A measurement of σ_z on a qubit yields either $+1$ or -1 as the observed value and projects the qubit into the eigenstate $|0\rangle$ or $|1\rangle$ corresponding to the observed value.

The orthonormal eigenvectors $|0\rangle, |1\rangle$ of σ_z form a standard basis in ¶$\mathbb{H}$, allowing us to identify the qubit Hilbert space ¶$\mathbb{H}$ with $\mathbb{C}^2$. From now on, we will use these states to represent the classical bit values 0 and 1. A measurement of σ_z yielding $+1$ corresponds to the classical bit value 0, and according to the Postulate 3, the qubit is in the state $|0\rangle$. Similarly, a measurement yielding -1 corresponds to the bit value 1, and the qubit is in state $|1\rangle$.

Thus, each classical bit value is mapped to a qubit state. However, not every qubit state can represent a classical bit value. A general qubit state is given by

$$|\psi\rangle = a|0\rangle + b|1\rangle \tag{3.4.1}$$

with $a, b \in \mathbb{C}$ and $|a|^2 + |b|^2 = 1$. If both a and b are non-zero, the state is a superposition of $|0\rangle$ and $|1\rangle$, which has no classical counterpart. We won't see this, but such superpositions, unique to quantum systems, are the key to the efficiency of quantum algorithms compared to classical ones.

3.5 Qbytes

Classically, information is represented by bits, and a two-bit word, such as (x_1, x_2), is an element of $\{0, 1\}^2$, where each bit corresponds to 0 or 1. When using qubits instead of bits, we deal with a two-qubit quantum system composed of two subsystems.

In quantum mechanics, systems often consist of multiple parts, each described by its own Hilbert space. For instance, a hydrogen atom consists of a proton and an electron, described by Hilbert spaces $\mathbb{H}^P$ and $\mathbb{H}^E$, respectively. The state space of the entire system is the *tensor product* $\mathbb{H}^P \otimes \mathbb{H}^E$, which combines the sub-systems. More generally, the tensor product of two Hilbert spaces $\mathbb{H}^A \otimes \mathbb{H}^B$ describes the state space of a system composed of two subsystems. In Section A.5 we will recall some important facts about the tensor product of Hilbert spaces, here we will only deal with what is relevant for quantum computing.

Definition 3.5.1. The n-fold tensor product of qubit spaces is defined as

$$\text{¶}\mathbb{H}^{\otimes n} \coloneqq \text{¶}\mathbb{H} \otimes \cdots \otimes \text{¶}\mathbb{H} \quad (n \text{ factors}).$$

We denote the $j+1$-th factor space, counting from the right in $\text{¶}\mathbb{H}^{\otimes n}$, by $\text{¶}\mathbb{H}_j$. In other words, we define

$$\text{¶}\mathbb{H}^{\otimes n} = \text{¶}\mathbb{H}_{n-1} \otimes \cdots \otimes \underbrace{\text{¶}\mathbb{H}_j}_{j+1\text{-th factor}} \otimes \cdots \otimes \text{¶}\mathbb{H}_0.$$

The Hilbert space $\text{¶}\mathbb{H}^{\otimes n}$ is 2^n-dimensional. The reason for counting spaces from the right will become evident in the following, when we will define the computational basis. Remember that every $x \in \mathbb{N}$ with $x < 2^n$ can be expressed in its *bynary representation*

$$x = \sum_{j=0}^{n-1} x_j 2^j \quad \text{with } x_j \in \{0,1\},$$

we can also write

$$(x)_2 = x_{n-1} \dots x_1 x_0.$$

Definition 3.5.2. Let $x \in \mathbb{N}$ with $x < 2^n$, and let $x_0, \dots, x_{n-1} \in \{0,1\}^n$ be the coefficients of its binary representation. For each such x, we define a vector $|x\rangle \in \text{¶}\mathbb{H}^{\otimes n}$ as

$$\begin{aligned} |x\rangle^n \coloneqq |x\rangle &\coloneqq |x_{n-1} \dots x_1 x_0\rangle \\ &\coloneqq |x_{n-1}\rangle \otimes \cdots \otimes |x_1\rangle \otimes |x_0\rangle = \bigotimes_{j=n-1}^{0} |x_j\rangle . \end{aligned}$$

If it is clear in which product space $\text{¶}\mathbb{H}^{\otimes n}$ the vector $|x\rangle^n$ lies, we will simply write $|x\rangle$ instead of $|x\rangle^n$.

In $\text{¶}\mathbb{H}^{\otimes n}$ the smallest and largest representable numbers are 0 and

$2^n - 1$, and for them

$$|2^n - 1\rangle^n = |11\ldots1\rangle = \bigotimes_{j=n-1}^{0} |1\rangle\,,$$
$$|0\rangle^n = |00\ldots0\rangle = \bigotimes_{j=n-1}^{0} |0\rangle\,.$$

Lemma 3.5.3. The set of vectors $\left\{|x\rangle \in {}^{\P}\mathbb{H}^{\otimes n} \mid x \in \mathbb{N},\ x < 2^n\right\}$ forms an ONB of ${}^{\P}\mathbb{H}^{\otimes n}$.

Proof. For $|x\rangle, |y\rangle \in {}^{\P}\mathbb{H}^{\otimes n}$ one has

$$\begin{aligned}\langle x|y\rangle &= \langle x_{n-1}\ldots x_0|y_{n-1}\ldots y_0\rangle \\ &= \prod_{j=0}^{n-1}\langle x_j|y_j\rangle = \prod_{j=0}^{n-1}\delta_{x_j y_j} = \delta_{xy}.\end{aligned}$$

Hence, the set $\left\{|x\rangle \in {}^{\P}\mathbb{H}^{\otimes n} \mid x \in \mathbb{N},\ x < 2^n\right\}$ has 2^n orthonormal vectors, and since $\dim {}^{\P}\mathbb{H}^{\otimes n}$ they form an ONB. □

This ONB is very useful and has its own name.

Definition 3.5.4. The orthonormal basis in ${}^{\P}\mathbb{H}^{\otimes n}$ defined for $x \in \mathbb{N}$, with $x < 2^n$, by $|x\rangle = |x_{n-1}\ldots x_0\rangle$ is called the *computational basis*.

Example 3.5.5. In ${}^{\P}\mathbb{H}$, the computational basis is identical to the standard basis:

$$|0\rangle^1 = |0\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \quad |1\rangle^1 = |1\rangle = \begin{pmatrix}0\\1\end{pmatrix},$$

where the rightmost equalities show the identification with the standard basis in $\mathbb{C}^2 \cong {}^{\P}\mathbb{H}$. The four basis vectors of the computational basis in ${}^{\P}\mathbb{H}^{\otimes 2} \cong \mathbb{C}^4$ are:

$$|0\rangle^2 = |00\rangle = |0\rangle \otimes |0\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \qquad |1\rangle^2 = |01\rangle = |0\rangle \otimes |1\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix},$$
$$|2\rangle^2 = |10\rangle = |1\rangle \otimes |0\rangle = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \qquad |3\rangle^2 = |11\rangle = |1\rangle \otimes |1\rangle = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}.$$

In ¶$\mathbb{H}^{\otimes 3} \cong \mathbb{C}^8$, the computational basis vectors are:

$$|0\rangle^3 = |000\rangle = |0\rangle \otimes |0\rangle \otimes |0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$|1\rangle^3 = |001\rangle = |0\rangle \otimes |0\rangle \otimes |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$|2\rangle^3 = |010\rangle\,, \quad |3\rangle^3 = |011\rangle\,, \quad |4\rangle^3 = |100\rangle\,,$$
$$|5\rangle^3 = |101\rangle\,, \quad |6\rangle^3 = |110\rangle\,, \quad |7\rangle^3 = |111\rangle\,.$$

In ¶$\mathbb{H}$ we may consider

$$|\varphi_1\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$
$$|\varphi_2\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$
$$|\psi_1\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$
$$|\psi_2\rangle = |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then, for example, in ¶$\mathbb{H}^{\otimes 2}$ we can construct:

$$|\varphi_1 \otimes \varphi_2\rangle = |\varphi_1\rangle \otimes |\varphi_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix},$$
$$\langle \psi_1 \otimes \psi_2| = \langle 0| \otimes \langle 1| = \begin{pmatrix} 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix},$$

where the rightmost vectors are expressed in the basis given in Example 3.5.5 and its dual. Using this we can find also that in this basis the matrix

$|\varphi_1 \otimes \varphi_2\rangle \langle \psi_1 \otimes \psi_2|$ is given by

$$|\varphi_1 \otimes \varphi_2\rangle \langle \psi_1 \otimes \psi_2| = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}$$
$$= \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

On the other hand, we could also have obtained that by

$$|\varphi_1 \otimes \varphi_2\rangle \langle \psi_1 \otimes \psi_2| = |\varphi_1\rangle \langle \psi_1| \otimes |\varphi_2\rangle \langle \psi_2|$$
$$= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$
$$= \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

3.6 Quantum Gates

The computational model for quantum computers is analogous to the classical model based on the *Turing Machine*, however, we will not go into the details of such a computational model. Instead of states represented by elements in $\{0,1\}^n$, pure quantum states are vectors in the Hilbert space $\P\mathbb{H}^{\otimes n}$. A quantum computational process transforms a state of n qubits to another, preserving the linear structure and normalization. This process is a unitary transformation $U : \P\mathbb{H}^{\otimes n} \to \P\mathbb{H}^{\otimes n}$, which physically is generated by applying an Hamiltonian for an appropriate period of time.

To extract the result of a quantum computation, measurement is required, and this introduces a non-unitary transition from the prepared quantum state to the final measured state, as we saw in Postulate 3.

Quantum gates, which are analogous to classical gates, are defined as unitary operators that act on the space of qubits.

Definition 3.6.1. A *quantum n-gate* is a unitary operator

$$U : \P\mathbb{H}^{\otimes n} \to \P\mathbb{H}^{\otimes n}.$$

Unary gates correspond to $n = 1$, and *binary gates* to $n = 2$.

Quantum gates are linear transformations and can be represented by matrices in the computational basis. Complex n-gates can be constructed from elementary unary and binary gates. An important example is the following.

Example 3.6.2. The *Quantum NOT gate* is the well known Pauli matrix

$$X \coloneqq \sigma_x.$$

We already encountered this matrix in Section 3.2, and we already know it is unitary. Because of

$$\sigma_x |0\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle ,$$
$$\sigma_x |1\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle .$$

it is considered the quantum analogue of the classical negation and thus termed as the quantum NOT gate.

Example 3.6.3. One of the simplest one qubit non classical gate one can imagine is a fractional power of the NOT gate, such as $\sqrt{\text{NOT}}$:

$$\sqrt{\text{NOT}} \coloneqq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{\frac{1}{2}} = \frac{1}{2} \begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}.$$

The $\sqrt{\text{NOT}}$ gate has the property that a repeated application of the gate, i.e., $\sqrt{\text{NOT}} \cdot \sqrt{\text{NOT}}$, is equivalent to the NOT operation, but a single application results in a quantum state that neither corresponds to the classical bit 0 nor the classical bit 1, in fact

$$\sqrt{\text{NOT}}|0\rangle = \frac{1}{2}\left((1+i)|0\rangle + (1-i)|1\rangle\right)$$
$$\sqrt{\text{NOT}}|1\rangle = \frac{1}{2}\left((1-i)|0\rangle + (1+i)|1\rangle\right).$$

Note that measurements are not quantum gates as intended in Definition 3.6.1, since they are not invertible and so definetely not unitary, but still they are present in many quantum algorithms, so often they are presented with the other quantum gates.

CHAPTER 4

An Application to Quantum Computing

Remember that we had:

Postulate. In a quantum system with Hilbert space $\mathbb{H}$, every change of a pure state over time that is not caused by a measurement is described by the time evolution operator $U(t, t_0) \in \mathcal{U}(\mathbb{H})$.

Let $|\psi(t_0)\rangle$ be the state at time t_0 and $|\psi(t)\rangle$ be the state at time t. The time-evolved state $|\psi(t)\rangle$ originating from the initial state $|\psi(t_0)\rangle$ is given by:

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle. \tag{4.0.1}$$

The time evolution operator $U(t, t_0)$ is the solution of the initial value problem:

$$\begin{aligned} i\frac{d}{dt}U(t, t_0) &= H(t)U(t, t_0), \\ U(t_0, t_0) &= I, \end{aligned} \tag{4.0.2}$$

where $H(t)$ is the self-adjoint Hamiltonian operator, which generates the time evolution of the quantum system.

We now aim to generalize this postulate to handle less regular Hamiltonians, this will enable more complex operations in quantum computing. For instance, we may want to apply quantum gates that depend on parameters varying them with low regularity. This flexibility proves valuable in rapidly expanding fields such as quantum machine learning and optimization. For instance, in [FGG14], a quantum optimization algorithm is introduced, where random "rotations" are applied to an initial random state. The goal is to apply as many of these transformations as possible, since in the limit, one approaches the desired optimal state. Instead of

applying a sequence of multiple operators one after the other, the multiplicative sewing lemma can be employed to obtain a single operator equivalent to the product of all previous operators, in the limit as the number of terms approaches infinity.

4.1 A Generalisation of the Fourth Postulate

From now on we will consider a function $A \in \mathcal{C}^\alpha([0,T], L(\mathbb{H}))$, where $\alpha > 1/2$; we will focus on the differential equation:

$$U_t = I + \int_0^t dA_s \, U_s. \tag{4.1.1}$$

Note that it is equivalent to equation 4.0.2 if we set:

$$\begin{aligned} U_t &= U(t, t_0), \\ A_t &= -i \int_0^t H_s \, ds. \end{aligned}$$

Remark 4.1.1. We can also introduce a "differential notation", for the differential equation (4.1.1), we may write:

$$\begin{cases} dU_t = dA_t \, U_t \\ U_0 = I \end{cases}$$

Meaning that when we integrate the first expression we get the original equation.

4.2 Existence and Uniqueness of Solution

In Theorem 2.3.2 we showed that a solution to the differential equation

$$V_t = I + \int_0^t V_s \, dA_s. \tag{4.2.1}$$

is

$$V_t = \prod_0^t e^{dA_s}.$$

This equation is actually very similar to (4.1.1), and we can use the following lemma to relate the two equations.

Lemma 4.2.1. Let $U \in \mathcal{C}^\alpha([0,T], L(\mathbb{H}))$, then U is a solution to (4.1.1) if and only if U^* is a solution to

$$U_t^* = I + \int_0^t U_s^* \, dA_s^*. \tag{4.2.2}$$

Proof. Note that if U is a solution to (4.1.1) then

$$U_t^* = I + \left(\int_0^t dA_s \, U_s \right)^*,$$

but using Proposition 1.3.4 it is easy to see that

$$\left(\int_0^t dA_s \, U_s \right)^* = \int_0^t U_s^* \, dA_s^*$$

as wanted. □

Then, using this lemma and Theorem 2.3.2, a solution to Equation (4.1.1) is

$$U_t = \left(\prod_0^t e^{dA_s^*} \right)^*.$$

The solution is actually also unique, in the sense that U_t is the only $\mathcal{C}^\alpha$ function that satisfies Equation (4.1.1). To show this we will need the following:

Lemma 4.2.2 (Young-Grönwall)**.** Let $\alpha, \beta \in (0, 1]$, with $\alpha + \beta > 1$ and $\beta > \alpha$, consider the following functions:

$$a \in C^\alpha([0,T]; \mathbb{R}^d), \quad u \in C^\alpha([0,T]; \mathbb{R}^{d\times(k\times d)}), \quad y \in C^\beta([0,T]; \mathbb{R}^k).$$

These functions satisfy the integral equation:

$$a_t = a_0 + \int_0^t ua \, dy, \quad \text{for every } t \in [0,T]. \tag{4.2.3}$$

Then, there exists a constant $\mathbf{c}$ such that:

$$\|a\|_\beta \leq \mathbf{c} \, |a_0|.$$

The proof of the lemma will be in Appendix B.

Note that, although the formulation of the lemma may seem different from the problem we are studying, it is not restrictive. It is in fact sufficient to write the equation in coordinates to get in the same setting.

As an example, let's see how to write Equation (4.1.1) in that form. First, we shall rewrite it as

$$U_t^{ij} = \delta^{ij} + \int_0^t \sum_{k=1}^d dA_s^{ik} U_s^{kj},$$

where δ^{ij} is the Kronecker delta, defined as follows:

$$\delta^{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

Then we can choose β such that $1/2 < \beta < \alpha$, and then define

$$a \in C^\beta([0,T];\mathbb{R}^{d^2}), \quad u \in C^\beta([0,T];\mathbb{R}^{d^2\times(d^2\times d^2)}), \quad y \in C^\alpha([0,T];\mathbb{R}^{d^2}),$$

as

$$a^{i+(j-1)\cdot d} = U^{ij},$$
$$y^{i+(j-1)\cdot d} = A^{ij},$$

We set $u^{lmn} = 1$ if there exists $k \in \{1,\dots,d\}$ such that

$$l = k + (j-1)d, \quad m = i + (j-1)d, \quad n = i + (k-1)d,$$

and $u^{lmn} = 0$ otherwise. Lastly, if we set $a_0^{i+(j-1)\cdot d} = \delta^{ij}$, our original equation will be equivalent to (4.2.3) and we can apply the lemma.

We can now prove:

Proposition 4.2.3. Let $\alpha > 1/2$, and let $A \in \mathcal{C}^\alpha([0,T], L(\mathbb{H}))$. Suppose also that $U, V \in \mathcal{C}^\alpha([0,T], L(\mathbb{H}))$ are such that $U_0 = V_0$ and for every $t \in [0,T]$

$$U_t = U_0 + \int_0^t dA_s\, U_s, \quad V_t = V_0 + \int_0^t V_s\, dA_s.$$

Then, the solutions are unique, i.e., $U = V$.

Proof. Let $P = U - V$, note that P is α-Hölder continuous and satisfies the equation

$$P_t = \int_0^t dA_s\, P_s.$$

A simple application of Lemma 4.2.2 then leads to

$$\|P\|_\alpha \le \mathbf{c}\ |P_0| = 0,$$

but this means P is identically zero and so $U = V$ as desired.

□

Remark 4.2.4. Note that with a very similar approach we could show uniqueness also for the equation

$$U_t = U_0 + \int_0^t U_s\ dA_s$$

and more generally for all equations of the form

$$U_t = U_0 + \int_0^t f(dA_s, U_s),$$

where $f(dA_s, U_s)$ is some linear combination of the entries of dA_s and U_s; it is in fact sufficient to write the right u to apply Lemma 4.2.2.

4.3 Unitarity of the Solutions

Now that we have established existence and uniqueness for the solutions of Equation (4.1.1), we still have to ensure that the solutions are unitary, otherwise we won't have any physical interpretation for non-unit vectors of $\mathbb{H}$.

To do so we can ask that $A^* = -A$, this is motivated by the fact that to get the original Equation (4.0.2) we need to set

$$A_t = -i \int_0^t H_s \, ds,$$

and since H is Hermitian we get exactly $A_t^* = -A_t$ for every t.

Proposition 4.3.1. Let A as in proposition 4.2.3 be such that $A_t^* = -A_t$ for every $t \in [0, T]$, and let U be a solution to

$$U_t = I + \int_0^t dA_s \; U_s,$$

then U_t is unitary for every $t \in [0, T]$.

Proof. To show that U_t is unitary, we will consider U_t^*. Note that, because of Lemma 4.2.1, it satisfies the equation

$$U_t^* = I + \int_0^t U_s^* \; dA_s^*,$$

and since $A_t^* = -A_t$ we get

$$U_t^* = I - \int_0^t U_s^* \; dA_s.$$

Using now the integration by parts formula for the Young integral, shown in Theorem 1.4.4, we get

$$U_t U_t^* = U_0 U_0^* + \int_0^t U_s \; dU_s^* + \int_0^t dU_s \; U_s^*,$$

and using the transitivity of Young integral (Propositions 1.4.2 and 1.4.3):

$$\begin{aligned} U_t U_t^* &= U_0 U_0^* - \int_0^t U_s U_s^* \; dA_s + \int_0^t dA_s \; U_s U_s^* \\ &= I + \int_0^t (dA_s \; U_s U_s^* - U_s U_s^* \; dA_s). \end{aligned} \tag{4.3.1}$$

Where the last integral is, as in Remark 4.2.4, to be understood as a linear combination of the entries of dA_s and $U_s U_s^*$. Because of Remark

4.2.4, we can apply Young-Grönwall's lemma and obtain that Equation (4.3.1) has a unique solution.

Note that $U_t U_t^* = I$ for every t is a solution to (4.3.1), so U_t is unitary for every $t \in [0, T]$.

□

Remark 4.3.2. If we write the equations in the differential form we can get the same result, integration by part yields

$$d(U_t U_t^*) = U_t \; dU_t^* + dU_t \; U_t^*,$$

and then formally substituting the equation for dU^* and dU, which is equivant to using transitivity of Young integral, we get

$$d(U_t U_t^*) = -(U_t U_t^*) \; dA_t + dA_t \; (U_t U_t^*),$$

this can be written in a nice way as

$$d(U_t U_t^*) = [dA_t, (U_t U_t^*)],$$

meaning that the commutator must be formally carried out. The last expression is equivalent to equation (4.3.1).

APPENDIX A

Notation and Useful Results

A.1 Discrete differential calculus

Given a metric space (X, d) and a function $f : X \to \mathbb{R}^k$, we write $f(x) := f_x$ for $x \in X$ and define

$$\delta f : X^2 \to \mathbb{R}^k, \quad \delta f_{xy} := f_y - f_x \quad \text{for } x, y \in X.$$

Notice that discrete Leibniz rules hold in the following form. For functions $f, g : X \to \mathbb{R}$, let $(fg)_x := f_x g_x$, then for $x, y \in X$

$$\delta(fg)_{xy} = (\delta f_{xy})g_y + f_x(\delta g_{xy}) = f_x(\delta g_{xy}) + (\delta f_{xy})g_x + (\delta f_{xy})(\delta g_{xy})$$

With a slight abuse of notation, when $X \subset \mathbb{R}^k$ and f is the identity map, we write $\delta_{xy} = y - x$, hence $d(x, y) = |y - x| = |\delta_{xy}|$.

A.2 Hölder functions

For a metric space (X, d) and a function $f : X \to \mathbb{R}^k$, we define

$$[f]_0 := \sup_{x \in X} |f_x|.$$

For a function $\omega : X^2 \to \mathbb{R}^k$ and $\alpha \geq 0$, we define

$$[\omega]_\alpha := \sup_{\substack{x,y \in X \\ x \neq y}} \frac{|\omega_{xy}|}{d(x, y)^\alpha} \in [0, +\infty].$$

We have the following properties:

$$[\omega + \omega']_\alpha \leq [\omega]_\alpha + [\omega']_\alpha, \quad [f\omega]_\alpha \leq [f]_0[\omega]_\alpha,$$

$$[\omega]_\alpha \leq [\omega]_\beta \operatorname{diam}(X)^{\beta-\alpha} \quad \text{if } \alpha \leq \beta. \tag{A.2.1}$$

Moreover, for any fixed $x \in X$, it holds that

$$[f - f_x]_0 \leq [\delta f]_\alpha \operatorname{diam}(X)^\alpha. \tag{A.2.2}$$

We say f is α-Hölder continuous if $[\delta f]_\alpha < +\infty$, moreover we say that $f \in C^\alpha(X; \mathbb{R}^k)$, if

$$\|f\|_\alpha := [f]_0 + [\delta f]_\alpha < \infty. \tag{A.2.3}$$

It is possible to prove that this is indeed a norm on $C^\alpha(X, \mathbb{R}^k)$. Note that for $\alpha = 1$, this defines the space of bounded Lipschitz functions on X with values in $\mathbb{R}^k$, rather than the usual space of continuously differentiable functions. When $k = 1$, we simply write $C^\alpha(X)$ instead of $C^\alpha(X; \mathbb{R})$.

A.3 Hilbert Spaces

Definition A.3.1. A Hilbert space $\mathbb{H}$ is a complex vector space with a (positive-definite) scalar product

$$\langle \cdot | \cdot \rangle : \mathbb{H} \times \mathbb{H} \to \mathbb{C}, \quad (\psi, \varphi) \mapsto \langle \psi | \varphi \rangle,$$

such that for all $\varphi, \psi, \varphi_1, \varphi_2 \in \mathbb{H}$ and $a, b \in \mathbb{C}$, the following hold:

$$\langle \psi | \varphi \rangle = \overline{\langle \varphi | \psi \rangle},$$

$$\langle \psi | \psi \rangle \geq 0,$$

$$\langle \psi | \psi \rangle = 0 \iff \psi = 0,$$

$$\langle \psi | a\varphi_1 + b\varphi_2 \rangle = a\langle \psi | \varphi_1 \rangle + b\langle \psi | \varphi_2 \rangle.$$

The scalar product induces a norm

$$\| \cdot \| : \mathbb{H} \to \mathbb{R}, \quad \psi \mapsto \sqrt{\langle \psi | \psi \rangle},$$

in which $\mathbb{H}$ is complete.

Remark A.3.2. Finite dimensional complex vector spaces, with a complex norm, which are the only cases relevant for us in this thesis, are always complete.

With the help of the scalar product every vector $\psi \in \mathbb{H}$ defines a linear map:

$$\langle \psi | : \mathbb{H} \to \mathbb{C}, \quad \varphi \mapsto \langle \psi | \varphi \rangle.$$

Conversely, one can show that every linear and continuous[1] map from $\mathbb{H}$ to $\mathbb{C}$ can be expressed with a $\psi \in \mathbb{H}$ in the form $\langle \psi |$. This means

[1] Continuity needs to be mentioned separately only in the infinite-dimensional case. In finite dimensional spaces every linear map is necessarily continuous.

that there is a natural bijection between $\mathbb{H}$ and its dual space, $\mathbb{H}^*$. This identification[2] motivates the *bra* and *ket* notation, derived from the word *bracket* and introduced by Dirac. *Bra-vectors* are elements of $\mathbb{H}^*$ and are written as $\langle\psi|$, while *ket-vectors* are elements of $\mathbb{H}$ and are written as $|\psi\rangle$. The application of the bra $\langle\varphi|$ on the ket $|\psi\rangle$ is the *bracket* $\langle\varphi|\psi\rangle \in \mathbb{C}$.

If we fix an orthonormal basis, or ONB, $\{e_j\}$ of $\mathbb{H}$, then for any $\psi \in \mathbb{H}$:

$$|\psi\rangle = \sum_j |e_j\rangle \langle e_j|\psi\rangle .$$

A.4 Operators on Hilbert spaces

Definition A.4.1. A linear map $A : \mathbb{H} \to \mathbb{H}$ is called an *operator* on the Hilbert space $\mathbb{H}$. The set of all operators on $\mathbb{H}$ is denoted by $L(\mathbb{H})$.

The operator $A^* : \mathbb{H} \to \mathbb{H}$ that satisfies

$$\langle A^*\psi|\varphi\rangle = \langle\psi|A\varphi\rangle \quad \forall\, |\psi\rangle, |\varphi\rangle \in \mathbb{H},$$

is called the adjoint operator to A. If $A^* = A$, then A is called *self-adjoint*.

In the finite dimensional case self-adjoint is the same as *Hermitian*, to be precise, A^* is actually a map $A^* : \mathbb{H}^* \to \mathbb{H}^*$, but as mentioned before we can identify $\mathbb{H}^*$ with $\mathbb{H}$.

Definition A.4.2. An operator U on $\mathbb{H}$ is called *unitary* if

$$\langle U\psi|U\varphi\rangle = \langle\psi|\varphi\rangle \quad \forall\, |\psi\rangle, |\varphi\rangle \in \mathbb{H},$$

The set of all unitary operators on $\mathbb{H}$ is denoted by $\mathcal{U}(\mathbb{H})$.

Unitary operators does not change the norm, i.e. $\|U\psi\| = \|\psi\|$, and it is easy to show that they have their adjoint operator as their inverse.

Definition A.4.3. Let A be an operator on a Hilbert space $\mathbb{H}$. A vector $|\psi\rangle \in \mathbb{H} \setminus \{0\}$ is called an eigenvector of A with eigenvalue $\lambda \in \mathbb{C}$ if

$$A|\psi\rangle = \lambda|\psi\rangle.$$

The eigenspace of A for λ, denoted $\mathrm{Eig}(A, \lambda)$, is the subspace spanned by all eigenvectors corresponding to λ. An eigenvalue λ is non-degenerate if its eigenspace is one-dimensional, and degenerate otherwise. The spectrum of A, $\sigma(A)$, is the set of $\lambda \in \mathbb{C}$ for which $(A - \lambda I)^{-1}$ does not exist.

[2] Identified with each other are the sets, but not the linear structures of the vector spaces, since the bijection $\mathbb{H} \ni |\psi\rangle \mapsto \langle\psi| \in \mathbb{H}^*$ is anti-linear.

Eigenvalues of an operator A are part of its spectrum. In infinite-dimensional Hilbert spaces, the spectrum can include a so called continuous part, but in this thesis we focus only on finite-dimensional spaces, so we can identify the spectrum with the set of the eigenvalues.

For self-adjoint operators, eigenvalues are real, and for unitary operators, they have absolute value 1.

Definition A.4.4. The *commutator* of two operators A and B is defined as

$$[A, B] \coloneqq AB - BA,$$

we say that A and B commute if their commutator vanishes, that is, if $[A, B] = 0$.

A.5 Tensor products of Hilbert spaces

Here we give a more informal definition of the tensor product of two finitely dimensional Hilbert spaces, this is sufficient for our purposes.

Let $\mathbb{H}^A$ and $\mathbb{H}^B$ be Hilbert spaces. Let $|\varphi\rangle \in \mathbb{H}^A$ and $|\psi\rangle \in \mathbb{H}^B$, we define the map

$$|\varphi\rangle \otimes |\psi\rangle : \mathbb{H}^A \times \mathbb{H}^B \to \mathbb{C}, \quad (\xi, \eta) \mapsto \langle \xi|\varphi\rangle_{\mathbb{H}^A} \langle \eta|\psi\rangle_{\mathbb{H}^B}.$$

This map is anti-linear in ξ and η and continuous. We define the set of all such maps as

$$\mathbb{H}^A \otimes \mathbb{H}^B := \{\Psi : \mathbb{H}^A \times \mathbb{H}^B \to \mathbb{C} \,|\, \text{anti-linear and continuous}\}.$$

This forms a vector space over $\mathbb{C}$ thus, $|\varphi\rangle \otimes |\psi\rangle$ is a vector in the space of anti-linear and continuous maps $\mathbb{H}^A \otimes \mathbb{H}^B$ as defined. In order to simplify the notation we shall also write:

$$|\varphi \otimes \psi\rangle \coloneqq |\varphi\rangle \otimes |\psi\rangle .$$

For vectors $|\varphi_k \otimes \psi_k\rangle \in \mathbb{H}^A \otimes \mathbb{H}^B$, with $k \in \{1, 2\}$, we can define the scalar product as:

$$\langle \varphi_1 \otimes \psi_1 | \varphi_2 \otimes \psi_2 \rangle \coloneqq \langle \varphi_1 | \varphi_2 \rangle_{\mathbb{H}^A} \langle \psi_1 | \psi_2 \rangle_{\mathbb{H}^B} , \tag{A.5.1}$$

where in the following we shall often omit the subscripts when it is clear in which space the scalar product is to be calculated.

Let $\{|e_a\rangle\} \subseteq \mathbb{H}^A$ and $\{|f_b\rangle\} \subseteq \mathbb{H}^B$ be two ONB of the two spaces, then the set $\{|e_a \otimes f_b\rangle\} \in \mathbb{H}^A \otimes \mathbb{H}^B$ is orthonormal, this is clear because of (A.5.1). Considering now an arbitrary vector $\Psi \in \mathbb{H}^A \otimes \mathbb{H}^B$, then for

this anti linear map:

$$\begin{aligned}\Psi(\xi,\eta) &= \Psi\left(\sum_a |e_a\rangle\langle e_a|\xi\rangle, \sum_b |f_b\rangle\langle f_b|\eta\rangle\right)\\ &= \sum_{a,b} \Psi\left(|e_a\rangle, |f_b\rangle\right)\langle\xi|e_a\rangle\langle\eta|f_b\rangle\\ &= \sum_{a,b} \Psi_{ab}\cdot\langle\xi|e_a\rangle\langle\eta|f_b\rangle\\ &= \sum_{a,b} \Psi_{ab}\cdot(|e_a\rangle\otimes|f_b\rangle)(\xi,\eta)\\ &= \sum_{a,b} \Psi_{ab}\cdot|e_a\otimes f_b\rangle(\xi,\eta),\end{aligned}$$

where we defined $\Psi_{ab} \coloneqq \Psi(|e_a\rangle, |f_b\rangle) \in \mathbb{C}$.

This proves that every vector $|\Psi\rangle \in \mathbb{H}^A \otimes \mathbb{H}^B$ can be written as a linear combination of the form

$$|\Psi\rangle = \sum_{a,b} \Psi_{ab}|e_a\rangle\otimes|f_b\rangle.$$

We can the extend the definition of the scalar product in (A.5.1) to every $\Psi, \Phi \in \mathbb{H}^A \otimes \mathbb{H}^B$ as:

$$\begin{aligned}\langle\Psi|\Phi\rangle &= \sum_{a_1,b_1}\sum_{a_2,b_2} \overline{\Psi_{a_1b_1}}\Phi_{a_2b_2}\langle e_{a_1}\otimes f_{b_1}|e_{a_2}\otimes f_{b_2}\rangle\\ &= \sum_{a,b} \overline{\Psi_{ab}}\Phi_{ab}. \qquad \text{(A.5.2)}\end{aligned}$$

One can show that (A.5.2) does not depend on the choice of the ONBs, and also that it is indeed a scalar product. Since $\mathbb{H}^A \otimes \mathbb{H}^B$ is finitely dimensional it is also complete, so it is an Hilbert space.

Definition A.5.1. The Hilbert space $\mathbb{H}^A \otimes \mathbb{H}^B$ with the scalar product (A.5.2) is called the *tensor product* of $\mathbb{H}^A$ and $\mathbb{H}^B$.

Also, we can easily construct an ONB for this space, using the following proposition:

Proposition A.5.2. Let $\{|e_a\rangle\} \subseteq \mathbb{H}^A$ be an ONB in $\mathbb{H}^A$ and $\{|f_b\rangle\} \subseteq \mathbb{H}^B$ be an ONB in $\mathbb{H}^B$. The set $\{|e_a\rangle\otimes|f_b\rangle\} = \{|e_a\rangle\otimes|f_b\rangle\}$ forms an ONB in $\mathbb{H}^A \otimes \mathbb{H}^B$, and for finite-dimensional $\mathbb{H}^A$ and $\mathbb{H}^B$ one has

$$\dim(\mathbb{H}^A\otimes\mathbb{H}^B) = \dim(\mathbb{H}^A)\dim(\mathbb{H}^B).$$

APPENDIX B

Proof of Young-Grönwall's Lemma

We will now state and prove the Young-Grönwall's Lemma that we used in Chapter 4. We will state and prove a more precise version of Lemma 4.2.2, in which we specify the dependence of the costant from the other parameters. We will also make a heavier use of the notation introduced in Appendix A.

Lemma B.0.1 (Young-Grönwall)**.** Let $\alpha, \beta \in (0, 1]$, with $\alpha + \beta > 1$ and $\beta > \alpha$, consider the following functions:

$$a \in C^\alpha([0,T];\mathbb{R}^d), \quad u \in C^\alpha([0,T];\mathbb{R}^{d\times(k\times d)}), \quad y \in C^\beta([0,T];\mathbb{R}^k).$$

These functions satisfy the integral equation:

$$a_t = a_0 + \int_0^t ua\,dy, \quad \text{for every } t \in [0,T]. \tag{B.0.1}$$

Then, there exists a constant $\mathbf{c} = \mathbf{c}(\alpha, \beta, T, \|u\|_\alpha, [\delta y]_\beta)$ such that:

$$\|a\|_\beta \leq \mathbf{c} \cdot |a_0|. \tag{B.0.2}$$

Proof. From an application of the Sewing Lemma, just like in 1.4.1 we can easily get that

$$|\delta a_{st}| = \left| \int_s^t ua\,dy \right| \leq \tilde{\mathbf{c}} \cdot \|u\|_\alpha \|a\|_\alpha [\delta y]_\beta |\delta_{st}|^\beta, \tag{B.0.3}$$

with, here and below, $\tilde{\mathbf{c}} \coloneqq \mathbf{c}(\alpha, \beta) \cdot (1 + T^\alpha)$, hence $\|a\|_\beta < +\infty$. To get (B.0.2) we can write

$$\delta a_{st} = \int_s^t ua_0\,dy + \int_a^b u(a - a_0)\,dy$$

and estimate the two terms individually. For the first one, we can proceed as in (B.0.3) and get

$$\left|\int_s^t ua_0\,dy\right| \leq \tilde{\mathbf{c}} \cdot \|u\|_\alpha |a_0| [\delta y]_\beta |\delta_{st}|^\beta.$$

For the second one, we have in a similar way

$$\begin{aligned}\left|\int_a^b u(a-a_0)\,dy\right| &\leq \tilde{\mathbf{c}} \cdot \|u(a-a_0)\|_\alpha [\delta y]_\beta |\delta_{st}|^\beta \\ &\leq \tilde{\mathbf{c}} \cdot \|u\|_\alpha [\delta a]_\beta T^{\beta-\alpha} [\delta y]_\beta |\delta_{st}|^\beta.\end{aligned}$$

Assume now that T is small enough so that

$$\mathbf{c} \cdot (1+T^\alpha)\|u\|_\alpha T^{\beta-\alpha}[\delta y]_\beta \leq \frac{1}{2}, \tag{B.0.4}$$

then we get

$$\left|\int_a^b u(a-a_0)\,dy\right| \leq \frac{1}{2}[\delta a]_\beta |\delta_{st}|^\beta \leq \frac{1}{2}\|a\|_\beta |\delta_{st}|^\beta.$$

In this case, putting all together, we get

$$|\delta a_{st}| \leq \tilde{\mathbf{c}} \cdot \|u\|_\alpha |a_0| [\delta y]_\beta |\delta_{st}|^\beta + \frac{1}{2}\|a\|_\beta |\delta_{st}|^\beta,$$

so that

$$[a]_\beta \leq \tilde{\mathbf{c}} \cdot \|u\|_\alpha |a_0| [\delta y]_\beta + \frac{1}{2}\|a\|_\beta$$

and

$$\|a\|_\beta \leq |a_0| + \tilde{\mathbf{c}} \cdot \|u\|_\alpha |a_0| [\delta y]_\beta + \frac{1}{2}\|a\|_\beta.$$

Then (B.0.2) holds with $\mathbf{c} \coloneqq 2(1 + \tilde{\mathbf{c}} \cdot \|u\|_\alpha [\delta y]_\beta)$.

For a general T, we introduce a partition $\{t_0 = 0, t_1, \ldots, t_n = T\}$ such that for each $I_i = [t_i, t_{i+1}]$, (B.0.4) holds with $\delta_{t_i t_{i+1}}$ instead of T. This can be achieved with n depending on $T, \|u\|_\alpha, [\delta y]_\beta$ (over the entire interval) only. We can also choose the partition such that $\delta_{t_i t_{i+1}} \leq 1$ for every i.

Let also $\|a\|_{\beta,i}$ be the norm of a restricted to the interval I_i, we will now prove by induction that $\|a\|_{\beta,i} \leq \mathbf{c}^{i+1} \cdot |a_0|$ for every $i < n$.

The base case is exactly what we showed in the first part of the proof, also, with the same reasoning, we get for $i > 0$:

$$\|a\|_{\beta,i} \leq \mathbf{c} \cdot |a_{t_i}|.$$

Now we treat the two cases $|a_{t_i}| \leq |a_{t_{i-1}}|$ and $|a_{t_i}| > |a_{t_{i-1}}|$ individually.

In the first case, we immediately get

$$\|a\|_{\beta,i} \leq \mathbf{c} \cdot |a_{t_i}| \leq \mathbf{c} \cdot |a_{t_{i-1}}| \leq \mathbf{c} \cdot (|a_{t_{i-1}}| + [\delta a]_{\beta,i-1}) \leq \mathbf{c} \cdot \|a\|_{\beta,i-1},$$

while for the second one, we can use that $|a_{t_i}| - |a_{t_{i-1}}| = |a_{t_i} - a_{t_{i-1}}|$ and

$$|a_{t_i} - a_{t_{i-1}}| \leq [\delta a]_{\beta,i-1} \cdot \delta_{t_i t_{i-1}} \leq [\delta a]_{\beta,i-1},$$

since $\delta_{t_i t_{i-1}} \leq 1$, so that

$$\|a\|_{\beta,i} \leq \mathbf{c} \cdot |a_{t_i}| \leq \mathbf{c} \cdot (|a_{t_{i-1}}| + [\delta a]_{\beta,i-1}) \leq \mathbf{c} \cdot \|a\|_{\beta,i-1}.$$

In both cases, using the inductive hypothesis, we obtain

$$\|a\|_{\beta,i} \leq \mathbf{c}^{i+1} \cdot |a_0|,$$

as desired.

Finally, given $s \in I_i$ and $t \in I_j$ with $s \neq t$, assuming without loss of generality that $i < j$, we have

$$\begin{aligned}
|\delta a_{st}| &\leq |\delta a_{st_{i+1}}| + \sum_{k=i+1}^{j-1} |\delta a_{t_k t_{k+1}}| + |\delta a_{t_j t}| \\
&\leq \mathbf{c} \cdot \|a\|_{\beta,i} |\delta_{st_{i+1}}|^\beta + \sum_{k=i+1}^{j-1} \mathbf{c} \cdot \|a\|_{\beta,k} |\delta_{t_k t_{k+1}}|^\beta + \mathbf{c} \cdot \|a\|_{\beta,j} |\delta_{t_j t}|^\beta \\
&\leq \sum_{k=i}^{j} \mathbf{c} \cdot \|a\|_{\beta,k} |\delta_{st}|^\beta \leq n\mathbf{c}^{n+1} \cdot |a_0| |\delta_{st}|^\beta,
\end{aligned}$$

hence the thesis follows with the costant $1 + n\mathbf{c}^{n+1}$.

□

Bibliography

[FdLP06] Denis Feyel and Arnaud de La Pradelle, *Curvilinear Integrals Along Enriched Paths*, Electronic Journal of Probability **11** (2006), no. none, 860 – 892.

[FdLPM08] Denis Feyel, Arnaud de La Pradelle, and Gabriel Mokobodzki, *A non-commutative sewing lemma.*, Electronic Communications in Probability [electronic only] **13** (2008), 24–34 (eng).

[FGG14] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann, *A quantum approximate optimization algorithm*, 2014.

[Gub04] M. Gubinelli, *Controlling rough paths*, Journal of Functional Analysis **216** (2004), no. 1, 86–140.

[Kon37] V. Kondurar, *Sur l'intégrale de Stieltjes*, Rec. Math. Moscou, n. Ser. **2** (1937), 361–366 (French).

[Sch19] Wolfgang Scherer, *Mathematics of quantum computing. An introduction. Translated from the German*, Cham: Springer, 2019 (English).

[ST22] Eugene Stepanov and Dario Trevisan, *On exterior differential systems involving differentials of Hölder functions*, J. Differ. Equations **337** (2022), 91–137 (English).

[You36] L. C. Young, *An inequality of the Hölder type, connected with Stieltjes integration*, Acta Mathematica **67** (1936), no. none, 251 – 282.