UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Quantum Wasserstein distance of order 1 between channels

by

Mathumo Mapaya

Under the supervision of Prof. Rocco Duvenhage

This is a dissertation submitted in fulfillment for the degree of

Magister Scientae

in the

Faculty of Natural and Agricultural Sciences

University of Pretoria

October 26, 2023

Declaration

I declare that this dissertation was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

A significant portion of the work presented here forms part of an article that appears in volume 26, issue number 23 of the Infinite Dimensional Analysis, Quantum Probability and Related Topics journal. This mostly includes work covered in Chapters 6, 7, 8 and 9. It also includes work from Sections 5.3 and 5.4. The article is titled "Quantum Wasserstein distance of order 1 between channels" and it was written by my supervisor Prof. Rocco Duvenhage with me as a co-author.

W. Mapaya.
Mr. Mathumo Mapaya
Student Number: 18026312

26/10/2023
Date

Abstract

We analyse a formulation of the quantum Wasserstein distance of order 1 and set up a general theory leading to a Wasserstein distance of order 1 between the unital maps from one specific algebra to another specified algebra. This gives us a metric on the set of unital maps from one composite system to another, which is deeply connected to the reductions of the unital maps. We use the fact that channels are unital maps with extra structure, to systematically define a quantum Wasserstein distance of order 1 between channels, i.e., a metric on the set of channels. Lastly, the additivity and stability properties of this metric are studied.

Keywords: quantum optimal transport; quantum Wasserstein distance of order 1; quantum channels; quantum states; composite systems

Acknowledgements

I thank everyone who has contributed to my academic journey thus far. First and foremost, I thank my supervisor, Professor Rocco Duvenhage, for accepting me into his research group. During the three years I've worked with him, he has made an invaluable contribution to my growth as a student researcher. His continuous support and humble approach to research is an inspiration.

Additionally, I thank my family and friends for their patience, financial assistance and moral support over the last few years.

Lastly, I would like to acknowledge the financial assistance from my supervisor, South Africa's National Institute for Theoretical and Computational Sciences (NITheCS) and the South African National Research Foundation (NRF).

Index of Symbols and Conventions

$\|\cdot\|$	:	Standard norm, sometimes denoted with a subscript.
$\mathbb{R}$	:	Field of real numbers.
$\mathbb{C}$	:	Field of complex numbers.
$\mathbb{N}$	:	Standard set of all natural numbers.
$\mathbb{F}$	:	Abitrary field, in this dissertation, this represents $\mathbb{C}$ or $\mathbb{R}$.
$\inf(A)$	:	Shorthand notation for the infimum of set A.
$\sup(A)$	:	Shorthand notation for the supremum of set A
$\lim_{n\to\infty} x_n$	:	limit of (x_n) as n tends to ∞.
$\|\cdot\|$	:	Modulus of complex numbers.
$M_n(A)$	:	The algebra of $n \times n$ matrices with entries from a C*-algebra A.
M_n	:	The algebra of $n \times n$ matrices with complex entries.
$M_{n,m}$	:	The algebra of $n \times m$ matrices with complex entries.
$1_A, I_n$	:	1_A is a unit of an algebra A and I_n is the unit of M_n.
x^*	:	The adjoint of $x \in A$ for A some *-algebra.
$\mathrm{Inv}(A)$	:	The set of all invertible elements of some algebra A.
$\sigma(x)$	:	Spectrum of an element x of some algebra.
$L(A, B)$	:	Space of linear maps $A \to B$.
$L(A)$	:	Shorthand notation for $L(A, A)$.
$\eta_i \circ \eta_j$	:	The composite map of linear maps η_i and η_j, see for example Eq. 6.1.
A''	:	Algebraic dual of an algebra A.
c.p.	:	Short for completely positive map.
u.c.p.	:	Short for unital completely positive map.
$\oplus$	:	Direct sum.
$\mathrm{span}_{\mathbb{F}}$	:	Span of some set over scalars from field $\mathbb{F}$. If there is no subscript, take the span over the field of the vector space containing the set.

$\langle \cdot, \cdot \rangle$	:	Inner product.
$\mathcal{H}, \mathcal{H}_{\mathcal{A}}$	:	A Hilbert space and a Hilbert space of some system $\mathcal{A}$ respectively.
$B(\mathcal{H})$	:	The set of all bounded operators of some Hilbert space $\mathcal{H}$.
$A \odot$ B	:	Algebraic tensor product between spaces or algebras A and B.
$A \otimes B$	:	Complete tensor product between spaces or algebras A and B.
$a \otimes b$	:	Elemenetary tensor of two elements from some vector spaces over $\mathbb{F}$.
$A \times B$	:	Cartesian product of spaces A and B.
$\mathbb{C}^d$	:	Standard d-dimensional complex vector space.
$\|\psi\rangle, \langle\psi\|$	:	A vector and its adjoint in Dirac notation.
$\langle \cdot \| \cdot \rangle$	:	Inner product in Dirac notation.
$\|\psi\rangle \langle\phi\|$	:	An operator in Dirac notation defined by $(\|\psi\rangle \langle\phi\|) \|\delta\rangle = \langle\phi\|\delta\rangle \|\psi\rangle$ for any $\|\delta\rangle$ in some inner product space.
A^+	:	The set of the positive elements of some algebra A.
Tr	:	The usual trace of a matrix.
$\mathbf{Tr}_i$	:	The partial trace, see Eq. 4.9.
Conv C	:	The convex hull of some set C.
W_1	:	Wasserstein distance of order 1.
$A_{\widehat{j}}$	:	Tensor product space with jth factor left out.
$\|_{A_{\widehat{j}}}$	:	Restriction of a map to the subspace $A_{\widehat{j}}$.
$\eta_{\widehat{j}}$	:	A map η restricted to a map between tensor product spaces that have the jth factor left out, see Eq.4.11.
A^T	:	The usual transpose of a matrix A.
$T_{\mathcal{B}}$	:	Partial transpose, see Eq.5.6.
δ_{jn}	:	The Kronecker delta.
id_A	:	The identity map on algebra A.
$K(A, B)$	:	Space of u.c.p. maps between C*-algebras A and B.
$L_u(A, B)$	:	Set of pointed maps, see Defn. 6.1.2.
id_A	:	The identity map on algebra A.
u_A	:	A distinguished element of the space A.
$(A, B, \nu, \mathcal{L})$	:	The pointed (W_1, n) structure, see Defn. 6.3.3.
$x \bowtie y$	:	Operator defined through $(x \bowtie y)z = x\nu(\langle y, z\rangle_\eta)$ for all $z \in G_\eta$, see the proof of Prop. 7.1.1.
$\mathcal{P}_n$	:	Collection of non-empty proper subsets of $\{1, \ldots, n\}$.
J'	:	The complement of a set $J \in \mathcal{P}_n$.
$\|J\|$	:	The cardinality of a set $J \in \mathcal{P}_n$.

A_J	:	Tensor product of A_j's for all $j \in J$ and any $J \in \mathcal{P}_n$.
η^J	:	The reduction of a map to any $J \in \mathcal{P}_n$.
$\mathcal{S}^J$	:	A set of maps that are reduced to any $J \in \mathcal{P}_n$.
$\vec{\odot}, \vec{\otimes}$	:	Tensor product indicating a specific order, see page 72.
$\alpha_{IJ}, \beta^0_{IJ}$	:	An ordering maps for tensor product algebras, see page 72.

Contents

Chapter 1

Introduction

Recent developments in quantum computing stimulate the growth of quantum information theory, with one of the aims being the comparison of quantum states. In classical information theory, two probability distributions can be compared through something called optimal transport. Optimal transport deals with finding the optimal cost of changing a probability distribution A into another probability distribution B. A nice way of visualizing how the transport part comes into the picture (making this a dynamical situation) is by letting the probability distribution A be a pile of balls while B is a pile of holes and assuming that moving a ball from A into a hole in B is associated with some unique cost. In this case, the main aim of optimal transport would be to compute the minimum (or lowest) cost of moving the balls (all of them) from A into the holes in B. The minimum cost is what is used to define the Wasserstein metrics (or distances).

In classical optimal transport, Wasserstein metrics have a wide range of applications, from physics to artificial intelligence. This leads to the expectation that these metrics will be equally important in quantum physics.

A number of approaches to formulating the quantum versions of the Wasserstein metrics and quantum optimal transport are under active development and being intensively investigated. See for example [12], [13], [14], [18], [16], [20], [29], [30], [33] and [68]. Every approach has its own advantages (the major differences will be covered towards the end of this introduction).

The framework provided by operator algebras and functional analysis is a natural setting for much of this development. For finite dimensional state spaces, the special case provided by linear algebra is sufficient as a framework for developing quantum optimal transport. For the infinite-dimensional case,

the theory of operator algebras and functional analysis become indispensable (this dissertation is a clear example of this).

Reference [29] (referred to as the DMTL in the remainder of this dissertation) introduces a quantum Wasserstein distance of order 1 between the states of a composite system with a finite dimensional state space, as a natural and very effective distinguishability measure with a number of desirable properties.

This dissertation is devoted to devising a distance between channels which is natural in the context of composite systems by analysing [29]'s approach. The ability to distinguish and quantify how far or different a real quantum channel is from an idealised channel is of great interest in quantum information theory. There are various other measures for this, each with its pros and cons [37]. In the DMTL approach the quantum W_1 distance has properties that are desired for quantum channels [11].

The metric resulting from the DMTL approach can be viewed as a quantum version of Ornstein's $\bar{d}$ distance [58]. While we use some of the core ideas of the DMTL approach, the extension to infinite dimensions and to unital maps (including channels), involves significant changes in the approach, including concepts and techniques not present in the finite dimensional case for states. Although the term Wasserstein distance is typically used for states (or probability distributions in the classical case), we nevertheless continue to use the term in the cases of unital maps and channels as well.

The infinite dimensional setting of this dissertation is expressed in terms of C*-algebras as a generalization of the matrix algebras M_d (Chapter 4 is a clear outline of this). We obtain a metric on the set of all channels from one composite system to another, where both consist of n systems. The case of states is obtained when the channels are taken to map to the complex numbers. In particular, the DMTL approach is recovered when restricting to finite dimensions, specifically to the algebras M_d.

The DMTL approach starts with the concept of neighbouring states, a form of which was introduced earlier in [1]. Broadly speaking, our approach is to extend the idea of neighbouring states to neighbouring channels. This is done through the reduction of channels. Given a channel from one composite system to another, choose the j'th system in each, and reduce the channel to the remaining systems. If two channels give the same channel after this reduction for some j is applied to both, we consider them to be neighbouring

channels. I.e., two channels are called neighbouring if they coincide after the removal of two corresponding systems from the two composite systems respectively. The neighbouring channels are then used to build a unit ball leading to a norm on a certain space of linear maps and ultimately to a distance between any two channels. More precisely, we obtain a metric on the set of channels from the one specific composite system to the other. All this is illustrated concretely in finite dimensions in Chapter 5.

One of our intentions is to highlight the general mathematical structure behind the quantum Wasserstein distance of order 1, through an abstract approach (with the metric obtained in Chapter 2 as the starting point). This involves first setting up the theory in general vector spaces, and subsequently algebras, without reference to any positivity conditions on the maps between these spaces. The C*-algebraic framework, with the maps taken as channels (unital completely positive maps between C*-algebras), is obtained as a special case of the abstract setup. This clarifies the overall structure of our approach, in particular, where and how complete positivity is used, and potentially allows for cases other than C*-algebras and channels between them.

In order to clarify our basic approach, we give an abstract DMTL approach in the next chapter. We then discuss the finite dimensional cases of states and channels. The C*-algebraic framework is a generalization of the finite dimensional cases, and is very natural and relevant from the view of quantum physics. The dissertation is written in such a way that the main thread and results in the C*-algebraic case can be followed without going through the more abstract approach mentioned above. The finite dimensional situation is in turn obtained as a simple special case of the C*-algebraic framework. The proofs of our results do depend on the abstract development, though.

Chapter 2 defines an abstract norm, as a result of analysing the mathematical structure of the DMTL approach. This includes the necessary background in relation to unit balls and norms that features throughout the rest of this dissertation. Chapter 3 introduces the operator algebraic framework (a summary, for cases where a reader is not that familiar with operator algebra theory).

For the case of states, Chapter 4 gives an alternative but equivalent way of defining the quantum W_1 distance between states (when compared to the DMTL approach).

Chapter 5 gives an outline of our approach for channels in finite dimensions, in a form that allows for a clear comparison to the DMTL approach. It can be viewed as an extension of this introduction, explaining some of the goals and motivation for this dissertation, but it also serves as an overview for readers whose main interest is the finite dimensional case. Section 5.1 reviews basic background related to channels, while Sections 5.3 and 5.4 take an initial step in developing the theory behind the Wasserstein distance of order 1 between channels.

Chapter 6 and chapter 7 develop the abstract theory to obtain our Wasserstein distance of order 1. The reader can in fact skip these two chapters upon initial reading, and go directly to Section 7.3. There the definition and main results leading to the Wasserstein distance of order 1 between channels in the C*-algebraic framework, are presented with no reference to Chapters 6 and 7, although the proofs depend on those two chapters. Note that we'll usually simply say "Wasserstein distance" instead of "quantum Wasserstein distance", since our theory contains the special case of abelian C*-algebras.

The dissertation then proceeds to the behaviour of the Wasserstein distance of order 1 with respect to subsystems of the composite systems, i.e., when we consider smaller composite systems consisting of a subset of the original systems. The main result here is that this Wasserstein distance is additive over tensor products of channels between such subsystems, with stability as a special case. Again this result is first approached abstractly in general vector spaces, before the C*-algebraic case is presented. As before, the reader can page directly to Section 8.4 after Section 7.3, to see the C*-algebraic results with a minimum of direct reference to the abstract theory, although the proofs again rely on the latter.

To conclude this introduction, we briefly discuss previous work.

The study of quantum channels, including mathematical techniques to analyse them theoretically, remains undeniably important. Refer for example to [15, 43, 50] for an overview of a variety of aspects of quantum channels and their significance. In particular, there have been other approaches to distances between channels, in relation to channel discrimination. For example, the diamond norm, also called the completely bounded trace norm; see [48, 3] for early work in connection to quantum computation, and [69, Section 3.3] for a more general finite dimensional overview in the context of quantum information. Refer to [37] for a broader perspective on distances between channels.

On the other hand, there has also been much effort to obtain quantum (or noncommutative) Wasserstein distances between states. Papers obtaining Wasserstein distances (with a focus on order 2) which are actual metrics on sets of quantum states, include [5, 12, 13, 14, 16, 18, 20, 33, 44, 70]. These papers follow different approaches from this dissertation and [29], and to a large extent from one another. Broadly speaking there have been two main approaches, in analogy to the classical case: a coupling (or transport plan) approach [5, 20, 33] and a dynamical approach [12, 13, 14, 16, 18, 44, 70]. The book [67] includes a nice introduction to the former, while [4] is the origin of the latter. The DMTL approach and this dissertation, can roughly be classified as part of the coupling approach. Other papers using the coupling approach, though not obtaining all the usual metric properties, include [30, 38, 39, 40]. Papers treating various other approaches than [29] to noncommutative or quantum Wasserstein distances of order 1, include [2, 14, 17, 36, 57, 60, 63, 68]. Some of these use a dual formulation of the coupling approach, which is closely related to Connes' spectral distance in noncommutative geometry, first introduced in [21], and studied further in [22, 23, 27, 59], among others.

It seems that a theory of Wasserstein distances, by any approach, has not before been extended from states to channels in the literature. This is of course aside from attempting to apply the former directly to the latter via the Choi-Jamiołkowski duality between states and channels. This duality holds under quite general conditions and could be applied beyond just the finite dimensional case; see in particular [34, Section 3]. However, the intention of this dissertation is rather to systematically rebuild such a theory for the channels, which is not equivalent to merely translating via the Choi-Jamiołkowski duality.

Chapter 2

An abstract approach to norms from unit balls

One should note that some background in functional analysis and linear algebra is assumed throughout this chapter and the rest of this dissertation. This chapter sets up some basic notions and elementary results for later use. This is mainly done for easy reference because the references providing this background in the preferred or needed form appear to be scarce. A somewhat different presentation of the following chapter is offered in [7, Section I.1] and [55, Chapters 4 and 5].

One of the main aims of this chapter is to define a norm from an abstract unit ball. It can be viewed as an abstract version of the corresponding ideas in [29]. Most of the general background definitions and results regarding analysis that are stated in this chapter, come from [51, 55, 62]. However, the notions "absorbing set" and "ray-wise bounded set" are introduced in a slightly different manner here than in other texts [7, 55], while the definition and analysis of the norm $\|\cdot\|_{\mathcal{C}}$ are adapted in part from [29], all motivated by the goal of obtaining an abstract form of [29]'s approach. We start with the definition of a norm and a metric.

2.1 Norms and metrics

This section is merely a review of some basic definitions and facts about seminorms, norms, pseudometrics and metrics.

Definition 2.1.1 (Seminorm and norm). A *seminorm* $\|\cdot\|$ on a vector space $\mathcal{X}$ over some field $\mathbb{F}$ is a real-valued function on $\mathcal{X}$ whose value at each $x \in \mathcal{X}$ is denoted by $\|x\|$ and has the following properties for any $x, y \in \mathcal{X}$ and $\alpha \in \mathbb{F}$:

N.1. $\|x\| \in [0, \infty)$

N.2. $\|\alpha x\| = |\alpha| \|x\|$ 'Homogeneity'

N.3. $\|x + y\| \leq \|x\| + \|y\|$ 'Triangle inequality'.

If in addition $\|\cdot\|$ has the following property,

N.4. if $\|x\| = 0$ then $x = 0$,

then $\|\cdot\|$ is called a *norm* on $\mathcal{X}$. In this case, the pair $(\mathcal{X}, \|\cdot\|)$, or even $\mathcal{X}$ itself when there is no ambiguity, is called a *normed space*.

We are interested in using our norm to induce a distance and we know from functional analysis and topology that the abstract definition of a distance is that of a metric, and this is what we define below.

Definition 2.1.2 (pseudometric and metric). A *pseudometric* on a set $\mathcal{X}$ is a real-valued function d on $\mathcal{X}$ whose value at each pair $x, y \in \mathcal{X}$ is denoted by $d(x, y)$ and has the following properties for any $x, y, z \in \mathcal{X}$:

M.0. $d(x, x) = 0$

M.1. $d(x, y) \in [0, \infty)$

M.2. $d(x, y) = d(y, x)$ 'Symmetric'

M.3. $d(x, y) \leq d(x, z) + d(z, y)$ 'Triangle inequality'.

If in addition d has the following property,

M.4. if $d(x, y) = 0$ then $x = y$,

then d is called a *metric*. In this case, the pair $(\mathcal{X}, d)$, or even $\mathcal{X}$ itself when there is no ambiguity, is called a *metric space*.

Remark 2.1.3. A note to make here is that the field of scalars $\mathbb{F}$ represents $\mathbb{R}$ or $\mathbb{C}$ throughout this dissertation. In this chapter, we mostly use $\mathbb{F}$ to represent $\mathbb{R}$, the reason why we do not use $\mathbb{R}$ directly is to emphasise that the definition or property in question also applies to $\mathbb{C}$. By *real* or *complex* vector space we refer to a vector space over $\mathbb{R}$ or $\mathbb{C}$ respectively.

In standard functional analysis literature they often show how a norm induces a metric, but they rarely show how a seminorm induces a pseudometric, hence we will show both in the following theorem.

Theorem 2.1.4. *Let $\mathcal{X}$ be a real vector space with a seminorm $\|\cdot\|_C$ defined on it, then*

$$d(x,y) := \|x-y\|_C$$

is a pseudometric on $\mathcal{X}$ for all $x, y \in \mathcal{X}$. If $\|\cdot\|_C$ is a norm, then d is a metric on $\mathcal{X}$.

Proof. Assume that $\|\cdot\|_C$ is a seminorm defined on some real or complex vector space $\mathcal{X}$, and define

$$d(x,y) := \|x-y\|_C\,,$$

for any $x, y \in \mathcal{X}$. A vector space is by definition closed under addition, hence $x-y \in \mathcal{X}$. **M**.0 follows from **N**.2 since $\|x-x\|_C = \|0\|_C = |0|\|0\|_C = 0$. **M**.1 follows from **N**.1. We also have that **N**.2 implies **M**.2 because $d(x,y) = \|x-y\|_C = \big\|(-1)(y-x)\big\|_C = |-1|\|y-x\|_C = d(y,x)$. Lastly, we also have that **N**.3 implies **M**.3 since

$$\begin{aligned}
d(x,y) &= \|x-y\|_C && \text{by the definition of } d\\
&= \|x-y-z+z\|_C && \text{for any } z \in \mathcal{X}\\
&= \big\|(x-z)+(z-y)\big\|_C\\
&\leq \|x-z\|_C + \|z-y\|_C && \text{by } \mathbf{N}.3 \text{ of } \|\cdot\|_C\\
&= d(x,z)+d(z,y) && \text{by the definition of } d.
\end{aligned}$$

Since **M**.0 - **M**.3 hold true, Definition 2.1.2 implies that d is a pseudometric on $\mathcal{X}$. If we assume that $\|\cdot\|_C$ is a norm, i.e., **N**.4 holds true, then we have that if $d(x,y) = 0$ then $\|x-y\|_C = 0$ with **N**.4 implying that $x-y=0$, hence $x=y$. So in conclusion, if $\|\cdot\|_C$ is a norm, then d is metric on $\mathcal{X}$. □

2.2 Unit balls

This section introduces the notion of an abstract unit ball.

In [29], they use the so-called set of neighbouring states to define their norm, so the first question one may ask is what properties must a ball have (in an abstract sense) to give us a of a norm? The following definitions are the answer to this question, we in fact need to define a set similar to what's known as unit ball (similar in a general sense) in introductory topology (see [25]).

Definition 2.2.1 (Absorbing set). A subset $\mathcal{C}$ of a real vector space $\mathcal{X}$ is said to be *absorbing* (for $\mathcal{X}$) if for every $x \in \mathcal{X}$ there is a number $r > 0$ such that $rx \in \mathcal{C}$.

Definition 2.2.2 (Convex set). A subset $\mathcal{C}$ of a real vector space $\mathcal{X}$ is said to be *convex* if for any $x, y \in \mathcal{C}$ we have that $ax + (1-a)y \in \mathcal{C}$ for $\forall a \in [0,1] \subseteq \mathbb{R}$.

Definition 2.2.3 (Symmetric set). A subset $\mathcal{C}$ of a real vector space $\mathcal{X}$ is said to be *symmetric* if $\mathcal{C} = -\mathcal{C}$, in other words if $x \in \mathcal{C}$ then $-x \in \mathcal{C}$.

Definition 2.2.4 (Ray-wise bounded). A subset $\mathcal{C}$ of a real vector space $\mathcal{X}$ is said to *ray-wise (or radially) bounded* if for any non-zero $x \in \mathcal{X}$ there is a number $s_0 > 0$ such that $sx \notin \mathcal{C}$ for all $s > s_0$.

Using the above definitions we can define an abstract *semi unit ball* and *unit ball* as the following.

Definition 2.2.5 (Semi unit ball and unit ball). A non-empty subset $\mathcal{C}$ of a real vector space $\mathcal{X}$, is called a *semi unit ball* of $\mathcal{X}$ if it is absorbing, convex and symmetric. If in addition $\mathcal{C}$ is ray-wise bounded, then we call it a *unit ball* of $\mathcal{X}$.

We will later show the connection between a standard definition of a unit ball from [25] and our Definition 2.2.5. Our definition is more general, in the sense that it doesn't require a prior norm or metric to be defined on the vector space in question.

Lastly, we mention a result that will allow us to make the above connection.

Proposition 2.2.6. *Let $\mathcal{C}$ be a semi unit ball in some real vector space $\mathcal{X}$, then $0 \in \mathcal{C}$.*

Proof. Let $x \in \mathcal{C}$ since $\mathcal{C}$ is non-empty by Definition 2.2.5. By the same definition, we get that $\mathcal{C}$ is symmetric and convex, hence $-x \in \mathcal{C}$ and $ax + (1-a)(-x) \in \mathcal{C}$ for any $a \in [0,1]$ respectively. Thus, for $a = 1/2$, we get that

$$ax + (1-a)(-x) = \frac{1}{2}x + \left(1 - \frac{1}{2}\right)(-x) = 0 \in \mathcal{C},$$

hence $0 \in \mathcal{C}$ as required. □

2.3 Norms from unit balls

The main aim of this section is to obtain a norm from a unit ball through what is known as a *Minkowski functional* or *gauge* function in a topological space (see [55]). Given that our focus is on vector spaces, we define a gauge function in our preferred setting as below.

Definition 2.3.1 (Gauge function). If $\mathcal{C}$ is a subset of a real vector space $\mathcal{X}$, then the *Minkowski functional* or *gauge* of $\mathcal{C}$ is defined to be the function $\|\cdot\|_{\mathcal{C}} : \mathcal{X} \to \mathbb{R}_+$, given by

$$\|x\|_{\mathcal{C}} := \inf\{t \geq 0 : x \in t\mathcal{C}\}$$

for every $x \in \mathcal{X}$, where $\mathbb{R}_+ = \{t \in \mathbb{R} : t \geq 0\}$.

The following definitions and lemma are needed to prove that the above gauge function is a norm under certain conditions and to show that our norm is more general when compared to [29, Definition 6] (in a mathematical sense).

Definition 2.3.2 (Lower bound). The *lower bound* b of a non-empty subset E of $\mathbb{R}$ is $b \in \mathbb{R}$ such that $b \leq x, \forall x \in E$.

Definition 2.3.3 (Infimum). The *infimum* of a non-empty subset E of $\mathbb{R}$ is the greatest lower bound of E, denoted as $\inf(E)$.

Remark 2.3.4. In a similar way (as in Definition 2.3.2), we define the *upper bound* by replacing '$\leq$' with '$\geq$'. Then the *supremum* (denoted as sup) is just defined as the least upper bound.

Remark 2.3.5. The *minimum* (denoted as min) of a non-empty subset E of $\mathbb{R}$ is the greatest lower bound b of E such that $b \in E$, i.e. when an infimum is contained in the set that it bounds it is called the minimum element of that set.

A note on the generality of our norm, Definition 2.3.1 uses an infimum of the set while [29, Definition 6] uses a minimum instead. The above remark shows that from a mathematical standpoint Definition 2.3.1 is a generalisation of [29, Definition 6] in the sense that a minimum of a set is a special case of the infimum.

Lemma 2.3.6. *If $P \subseteq \mathbb{R}$ and the $\inf(P)$ exists then there exists a sequence $(x_n) \in P$ such that (x_n) converges to $\inf(P)$.*

Proof. Assuming $P \subseteq \mathbb{R}$ has a lower bound, [41, The completeness axiom] implies that P has an infimum in $\mathbb{R}$, i.e., $m := \inf(P)$. Definition 2.3.3 implies that for any $n \in \mathbb{N}$, there exists $x_n \in P$ such that

$$m \leq x_n < m + \frac{1}{n},$$

now taking note of the following

$$\lim_{n\to\infty} m + \frac{1}{n} = m \quad \textit{and} \quad \lim_{n\to\infty} m = m,$$

[41, Sandwich rule] implies that

$$\lim_{n\to\infty} x_n = m = \inf(P).$$

So there is a sequence that converges to the infimum. □

Proposition 2.3.7. *Let $\mathcal{C}$ be a semi unit ball in a real vector space $\mathcal{X}$, then its gauge function $\|\cdot\|_{\mathcal{C}}$ is a seminorm on $\mathcal{X}$. If $\mathcal{C}$ is a unit ball in $\mathcal{X}$, then its gauge function $\|\cdot\|_{\mathcal{C}}$ is a norm on $\mathcal{X}$.*

Proof. Assuming that $\mathcal{C}$ is a semi umit ball in some real vector space $\mathcal{X}$, then for all $x, y \in \mathcal{X}$ and $\alpha \in \mathbb{R}$ we prove the following about the gauge function $\|\cdot\|_{\mathcal{C}}$:

N.1. $\|x\|_{\mathcal{C}} \in [0, \infty)$:

For any $x \in \mathcal{X}$ there is an $r > 0$ such that $rx \in \mathcal{C}$, this follows from Definition 2.2.5 and Definition 2.2.1. Hence $x \in \frac{1}{r}\mathcal{C}$, and therefore $\{t \geq 0 : x \in t\mathcal{C}\}$ is a non-empty set with 0 as a lower bound. Hence the infimum will be greater or equals to 0 because the infimum is the greatest lower bound. In addition, $\inf\{t \geq 0 : x \in t\mathcal{C}\} \leq \frac{1}{r} < \infty$.

N.2. $\|\alpha x\|_{\mathcal{C}} = |\alpha| \|x\|_{\mathcal{C}}$:

Let $t = |\alpha| s$ for $s \geq 0$ and $\alpha \neq 0$ then

$$\begin{aligned}
\|\alpha x\|_{\mathcal{C}} &= \inf\{t \geq 0 : \alpha x \in t\mathcal{C}\} \\
&= \inf\{|\alpha| s \geq 0 : \alpha x \in |\alpha| s\mathcal{C}\} \\
&= |\alpha| \inf\{s \geq 0 : \alpha x \in |\alpha| s\mathcal{C}\} \\
&= |\alpha| \inf\{s \geq 0 : x \in \frac{|\alpha|}{\alpha} s\mathcal{C}\} \\
&= |\alpha| \inf\{s \geq 0 : x \in s\mathcal{C}\} \quad \mathcal{C} \text{ is a symmetric set by Definition 2.2.5} \\
&= |\alpha| \|x\|_{\mathcal{C}}.
\end{aligned}$$

For $\alpha = 0$, it follows that $\|0x\|_{\mathcal{C}} = \|0\|_{\mathcal{C}}$, but $0 \in 0\mathcal{C}$ thus $\|0\|_{\mathcal{C}} = 0 = |0|\|x\|_{\mathcal{C}}$, so the homogeneity of the norm holds true.

N.3. $\|x+y\|_{\mathcal{C}} \leq \|x\|_{\mathcal{C}} + \|y\|_{\mathcal{C}}$:

Definition 2.3.1 implies that

$$\|x\|_{\mathcal{C}} = \inf\{t \geq 0 : x \in t\mathcal{C}\} \quad \text{and} \quad \|y\|_{\mathcal{C}} = \inf\{u \geq 0 : y \in u\mathcal{C}\}.$$

If $x = 0$ or $y = 0$, then $x \in 0\mathcal{C}$ or $y \in 0\mathcal{C}$ implying that $\|x\|_{\mathcal{C}} = 0$ or $\|y\|_{\mathcal{C}} = 0$, hence $\|x+y\|_{\mathcal{C}} = \|x\|_{\mathcal{C}} + \|y\|_{\mathcal{C}}$. Thus, without any loss of generality we may assume that $x \neq 0$ and $y \neq 0$. By Lemma 2.3.6 we have that there exists sequences (t_n) and (u_n) in $\{t \geq 0 : x \in t\mathcal{C}\}$ and $\{u \geq 0 : y \in u\mathcal{C}\}$ (respectively) that converge to $\|x\|_{\mathcal{C}}$ and $\|y\|_{\mathcal{C}}$ respectively. For $t_n \neq 0$ and $u_n \neq 0$, we have that $\frac{x}{t_n}$, $\frac{y}{u_n} \in \mathcal{C}$. Given that $\mathcal{C}$ is a semi unit ball, i.e., a convex set by Definition 2.2.5 we have that

$$\frac{t_n}{t_n + u_n}\left(\frac{x}{t_n}\right) + \left(\frac{u_n}{t_n + u_n}\right)\left(\frac{y}{u_n}\right) \in \mathcal{C},$$

which simplifies to

$$x + y \in (t_n + u_n)\mathcal{C}.$$

So by the properties of the infimum we have the following

$$\|x+y\|_{\mathcal{C}} \leq t_n + u_n \quad \forall n \in \mathbb{N}$$

hence,

$$\begin{aligned} \|x+y\|_{\mathcal{C}} &\leq \lim_{n\to\infty}(t_n + u_n) \\ &= \lim_{n\to\infty} t_n + \lim_{n\to\infty} u_n \\ &= \|x\|_{\mathcal{C}} + \|y\|_{\mathcal{C}}. \end{aligned}$$

For $t_n = 0$ or $u_n = 0$ we have that $x \in 0\mathcal{C}$ or $y \in 0\mathcal{C}$, implying that $x = 0c = 0$ or $y = 0c = 0$ for some $c \in \mathcal{C}$ which clearly leads to $\|x+y\|_{\mathcal{C}} = \|x\|_{\mathcal{C}} + \|y\|_{\mathcal{C}}$ as explained before.

Since **N**.1 - **N**.3 hold true, we have by Definition 2.1.1 that $\|\cdot\|_{\mathcal{C}}$ is a seminorm in $\mathcal{X}$ when $\mathcal{C}$ is a semi unit ball. If in addition, we have that $\mathcal{C}$ is also ray-wise bounded (i.e., a unit ball by Definition 2.2.5), then we can show that the following is true for any $x, y \in \mathcal{X}$:

N.4. If $\|x\|_{\mathcal{C}} = 0$ then $x = 0$:

We can follow the contrapositive statement and assume that $x \neq 0$ to prove that if $x \neq 0$ then $\|x\|_{\mathcal{C}} \neq 0$. Since $\mathcal{C}$ is ray-wise bounded, there is an $r_0 > 0$ such that $rx \notin \mathcal{C}$ for all $r > r_0$. Hence $x \notin \frac{1}{r}\mathcal{C}$ for $\frac{1}{r_0} > \frac{1}{r} > 0$, implying that $\|x\|_{\mathcal{C}} \geq \frac{1}{r_0} > 0$, i.e., $\|x\|_{\mathcal{C}} \neq 0$.

This now implies that **N**.1 - **N**.4 hold true in this case, hence by Definition 2.1.1 we thus have that $\|\cdot\|_{\mathcal{C}}$ is a norm on $\mathcal{X}$ when $\mathcal{C}$ is a unit ball. □

To summarise Proposition 2.3.7, $\mathcal{C}$ being absorbing ensures that $\|\cdot\|_{\mathcal{C}}$ is well-defined and finite at every point of the vector space, convexity implies the triangle inequality, symmetry implies the homogeneity, while ray-wise boundedness guarantees that $x = 0$ whenever $\|x\|_{\mathcal{C}} = 0$.

The assumption that $\mathcal{C}$ is a unit ball in a real vector space $\mathcal{X}$ implies that $\mathcal{X}$ is a normed space with the norm $\|\cdot\|_{\mathcal{C}}$ defined on $\mathcal{X}$ as in Definition 2.3.1. The next proposition mentions an interesting property and result of this norm and unit ball.

Proposition 2.3.8. *Let $\mathcal{C}$ be a unit ball in some real vector space $\mathcal{X}$ and define $\|\cdot\|_{\mathcal{C}}$ as in Definition 2.3.1, then*

$$\{x \in \mathcal{X} : \|x\|_{\mathcal{C}} < 1\} \subseteq \mathcal{C} \subseteq \{x \in \mathcal{X} : \|x\|_{\mathcal{C}} \leq 1\}.$$

Proof. Let $a \in \{x \in \mathcal{X} : \|x\|_{\mathcal{C}} < 1\}$, then $\|a\|_{\mathcal{C}} < 1$ implying that there exists some $t \in [0, 1)$ such that $a \in t\mathcal{C}$. For $t = 0$ we have that $a = 0 \in \mathcal{C}$ (see Proposition 2.2.6). Assuming that $t > 0$ implies that there is a vector $b \in \mathcal{C}$ such that $b = \frac{1}{t}a$. Since $\mathcal{C}$ is a convex set we have that

$$a = tb + (1 - t)0 \in \mathcal{C}.$$

Hence, $\{x \in \mathcal{X} : \|x\|_{\mathcal{C}} < 1\} \subseteq \mathcal{C}$.

Assume $a \in \mathcal{C}$, then Definition 2.3.1 implies that

$$\|a\|_{\mathcal{C}} = \inf\{t \geq 0 : \ a \in t\mathcal{C}\} \leq 1,$$

implying that $a \in \{x \in \mathcal{X} : \|x\|_{\mathcal{C}} \leq 1\}$. Therefore,
$\mathcal{C} \subseteq \{x \in \mathcal{X} : \|x\|_{\mathcal{C}} \leq 1\}$. □

Remark 2.3.9. From Proposition 2.3.8 it is clear that $\overline{\mathcal{C}} = \{x \in \mathcal{X} : \|x\|_{\mathcal{C}} \leq 1\}$ is the closure of $\mathcal{C}$, while $\mathring{\mathcal{C}} = \{x \in \mathcal{C} : \|x\|_{\mathcal{C}} < 1\}$ is the interior of $\mathcal{C}$. The definitions and properties of the closure and interior that make these results clear are in [64, Chapter 2]. This can be viewed as part of the motivation behind our definition of the abstract unit ball.

Chapter 3

The operator algebraic framework

This chapter is devoted to gathering (or summarising) some of the important definitions and results for our operator algebraic framework. In the first section of this chapter we focus on important definitions from the theory of operator algebras. Our focus in the subsequent sections is on the relevant results. The majority of this chapter comes from [6, 10, 54, 66], where operator algebras are discussed in more detail.

A note to make here (as mentioned in Remark 2.1.3) is that every vector space or algebra discussed in this chapter is over $\mathbb{C}$, that is $\mathbb{F} = \mathbb{C}$ in cases where the field $\mathbb{F}$ is used. In those cases, $\mathbb{F}$ is mainly used to emphasise that the definition or result also works for $\mathbb{R}$.

3.1 Elementary definitions

We start with a few definitions. These are algebraic structures that result from vector spaces that have more structure to them.

Definition 3.1.1 (Algebra, commutative algebra and unital algebra). An *algebra* is a vector space A over $\mathbb{C}$ that is equipped with a multiplication law or product $A \times A \to A : (x, y) \mapsto xy$, with the following distributive and associative properties:

A.1. $x\,(y + z) = xy + xz$

A.2. $\alpha\,(xy) = (\alpha x)\,y = x\,(\alpha y)$

A.3. $x\,(yz) = (xy)\,z$

for all $x, y, z \in A$ and $\alpha \in \mathbb{C}$. In addition, if $xy = yx, \forall x, y \in A$, then A is called a *commutative algebra* and if A has a unit (or multiplicative identity) denoted 1_A, i.e., $1_A \in A$ such that for all $x \in A$ we have that $1_A x = x 1_A = x$, then A is called a *unital algebra*.

Definition 3.1.2 (Inverse and spectrum). Let A unital algebra over $\mathbb{C}$. An element $x \in A$ is said to be *invertible* if there exists another element $y \in A$ such that

$$xy = yx = 1_A.$$

y is unique in this case and is called the *inverse* of x denoted by x^{-1}. The set of all invertible elements of A is denoted by

$$\text{Inv}(A) = \{x \in A : x \text{ is invertible}\}.$$

The *spectrum* of $x \in A$ is defined as the set

$$\sigma(x) = \{\lambda \in \mathbb{C} : \lambda 1_A - x \notin \text{Inv}(A)\}.$$

Definition 3.1.3 (*-algebra). An *involution* on an algebra A over $\mathbb{C}$ is a map $^* : A \to A : x \mapsto x^*$ which satisfies the following properties:

1. $(x^*)^* = x$,
2. $(xy)^* = y^* x^*$,
3. $(\alpha x)^* = \overline{\alpha} x^*$,

for all $x, y \in A$ and $\alpha \in \mathbb{C}$. A **-algebra* is an algebra A with an involution defined on it. If in addition, $1_A \in A$ such that $1_A^* = 1_A$, then A is called a *unital *-algebra*.

Let (x_n) be any sequence in some metric space $(\mathcal{X}, d)$. (x_n) is a said to be a *Cauchy sequence* if for every $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that for all natural numbers $n, m > N_\epsilon$, we have that $d(x_n, x_m) < \epsilon$. A metric space $(\mathcal{X}, d)$ is called a *complete metric space* if every Cauchy sequence converges to some point in $\mathcal{X}$ under d. A metric d is called a *complete metric* if $(\mathcal{X}, d)$ is a complete metric space and a *complete norm* is a norm that induces a complete metric.

Every finite dimensional vector space is complete (see [51, Theorem 2.4-2]). This is not always true for infinite dimensional spaces. In this dissertation, we are interested in spaces that are not necessarily finite dimensional. Hence the role of completeness will be discussed clearly (for example, see Section 6.2).

Definition 3.1.4 (Submultiplicative norm and normed algebra). A norm $\|\cdot\|$ defined on some algebra A is said to be *submultiplicative* if

$$\|xy\| \leq \|x\|\|y\|$$

for all $x, y \in A$. A *normed algebra* is an algebra A with a submultiplicative norm $\|\cdot\|$ defined on it.

Definition 3.1.5 (Banach *-algebra). A *Banach *-algebra* is a *-algebra A with a complete submultiplicative norm $\|\cdot\|$ defined on it such that

$$\|x^*\| = \|x\|$$

for all $x \in A$. In addition, if A has a unit $1_A \in A$ such that $\|1_A\| = 1$, then A is called a *unital Banach *-algebra*.

Definition 3.1.6 (C*-algebra). A *C*-algebra* is a Banach *-algebra A which satisfies

$$\|x^*x\| = \|x\|^2 \tag{3.1}$$

for all $x \in A$. Equation (3.1) is sometimes referred to as the *C*-property*, while a norm that satisfies this property is called a *C*-norm*. Additionally, if A is a unital Banach *-algebra with the C*-property, then A is said to be a *unital C*-algebra*.

The following examples are standard but important results that will also play a role in Chapter 4.

Example 3.1.7. $\mathbb{C}$ is a unital C*-algebra.

Proof. By [51, Theorem 1.4-4] we have that $\mathbb{C}$ is a metric space with a metric induced by the modulus. It's straightforward to verify that the multiplication of complex numbers satisfy Definition 3.1.1's **A**.1 - **A**.3. The involution is simply the complex conjugate of each complex number (i.e., for any $z \in \mathbb{C}$ we have that $z^* = \overline{z}$). The modulus is the norm, and it has the following properties:

$$|z| = |\overline{z}| = |z^*| \text{ and } |yz| = |y||z|,$$

for any $y, z \in \mathbb{C}$. Hence we have that, $|z^*z| = |z^*||z| = |z|^2$. Hence the modulus is a complete submultiplicative C*-norm, implying that $\mathbb{C}$ is a unital C*-algebra with $1_A = 1$ the unit (or multiplicative identity). □

Remark 3.1.8. Looking at the definition of a field [41, Section 2.2], we can say that every field is an algebra, but an algebra is not necessarily a field.

In the following example, one should note that we define and denote $M_n(A)$ as the set of all $n \times n$ matrices with the entries being elements of A.

Example 3.1.9. Let A be any C*-algebra, then $M_n(A)$ is a C*-algebra.

Proof. See the proof of [54, Theorem 3.4.2.]. □

In particular, the set $M_n(\mathbb{C})$ (which we simply denote as M_n) of $n \times n$ complex matrices is a unital C*-algebra.

Lastly, from the examples in [9, p. 20] and [66, p. 4], we arrive at the well-known conclusion that $B(\mathcal{H})$, the set of all bounded linear maps on some Hilbert space $\mathcal{H}$ is a unital C*-algbera, when it's equpped with the norm

$$\|T\|_{B(\mathcal{H})} = \sup\left\{\|Tx\|_{\mathcal{H}} \ : \ x \in \mathcal{H},\ \|x\|_{\mathcal{H}} = 1\right\}, \tag{3.2}$$

for any $T \in B(\mathcal{H})$, with $\|\cdot\|_{\mathcal{H}}$ being a norm defined on $\mathcal{H}$.

Definition 3.1.10 (Adjoint, self-adjoint and positive). Let A be a C*-algebra. For any element $x \in A$, the element $x^* \in A$ is called the *adjoint* of x. If $x = x^*$ then x is said to be *self-adjoint*. If in addition to being self-adjoint, $\sigma(x) \subset [0, \infty)$, then x is said to be positive and is denoted by $x \geq 0$.

Remark 3.1.11. Let A be given as in Definition 3.1.10, then for any $x \in A$, we have that $x^*x \geq 0$ and for any $y \in A$ such that $y \geq 0$, we have that there exists $z \in A$ such that $y = z^*z$. These results come from [54, Section 2.2].

We are now interested in maps between any of the above algebraic structures, the reasons behind this interest will become more evident in Chapters 4 and 5.

Definition 3.1.12 (Linear maps). Consider any two vector spaces A and B over the same field of scalars $\mathbb{F}$. A *linear map* from A to B is a function λ that maps the elements of A into B with the following properties for any $x, y \in A$ and $\alpha \in \mathbb{F}$,

1. $\lambda(x + y) = \lambda(x) + \lambda(y)$,
2. $\lambda(\alpha x) = \alpha\lambda(x)$.

We denote the set of all such linear maps by $L(A, B)$. A linear map is sometimes referred to as a *linear operator*. If $B = A$ then $L(A, B)$ is simply denoted as $L(A)$, and is called the set of all linear operators on A.

Remark 3.1.13. If we consider three vector spaces A, B and C over the same field of scalars $\mathbb{F}$, then a *bilinear map* $\phi : A \times B \to C$ is defined as a function ϕ that maps the cartesian product of two vector spaces A and B into another vector space C, such that ϕ is a linear map in each of its arguments separately. This is analogously extended to a definition of *multilinear maps*.

The following definitions will play an important role when it comes to connecting the DMTL approach [29] to our work in Chapter 4.

Definition 3.1.14 (Linear functionals and algebraic dual space). A *linear functional* is a linear map from a vector space A to its field of scalars $\mathbb{F}$. The *algebraic dual space* of A is the set of all such linear functionals denoted by A''.

Remark 3.1.15. It is easy to show that $L(A, B)$ is a vector space over the field $\mathbb{F}$, one just needs to use the properties from Definition 3.1.12 to show that the axioms of a vector space as given in [51, pp. 50-51] are true for this set. This also applies to the algebraic dual space (see [51, pp. 106-107]).

Next, we expand on our notions of maps as follows,

Definition 3.1.16 (morphisms). Let A and B be algebras, then a *homomorphism* π is defined as a linear map from A into B that preserves multiplication (i.e., $\pi(a_1a_2) = \pi(a_1)\pi(a_2)$ for all $a_1, a_2 \in A$). In addition, when A and B are *-algebras, if π preserves the involution, then it's called a $*$*-homomorphism*. Additionally, a $*$*-isomorphism* is defined as a bijective $*$*-homomorphism*.

Definition 3.1.17 (Positive map). Let A, B be C*-algebras. A *positive map* is linear map $\phi \in L(A, B)$ such that $\phi(x) \geq 0$ for any $x \in A$ such that $x \geq 0$.

Remark 3.1.18. An alternative but equivalent definition of a positive map is that it is any $\phi \in L(A, B)$ such that $\phi(x^*x) \geq 0$ for any $x \in A$ (This is a clear consequence of the results mentioned in Remark 3.1.11).

Definition 3.1.19 (Unital map). Let A, B be unital algebras. A *unital map* is linear map $\omega \in L(A, B)$ such that $\omega(1_A) = 1_B$.

In the case of C*-algebras, we are mainly interested in the following linear maps.

Definition 3.1.20 (States). A *state* ω of a unital C*-algebra A is a linear functional that maps A into $\mathbb{C}$ satisfying the following properties:

1. ω is a positive map,
2. ω is a unital map.

Remark 3.1.21. We use the term "state" in Definition 3.1.20 because those linear maps represent the states of a physical quantum system (see Theorem 4.1.4).

Definition 3.1.22 (Completely positive map). For A and B some C*-algebras, let ϕ be a linear map from A to B, then ϕ is said to be *completely positive* if the induced map

$$\phi^{(n)} : M_n(A) \to M_n(B),$$

which is obtained by applying ϕ entrywise is positive for all n.

Remark 3.1.23. It's clear that a completely positive map is a positive map (when $n = 1$, since 1×1 matrices can be treated as the field of entries). We will sometimes use c.p. to refer to completely positive maps. If in addition ϕ is a unital map, then we will use u.c.p. to denote a unital completely positive map. According to [6, II.6.9.6], a state is a u.c.p. map to $\mathbb{C}$, hence states are a special case of u.c.p. maps. In a C*-algbera setting, [6, II.6.2.2 and II.6.9.2] implies that states and completely positive maps are bounded linear maps. Hence, they are continuous (see [51, Theorem 2.7-9]).

The following is a crucial result. It simplifies working with c.p. maps in the context of Chapter 5. Before that, one should note that by $M_{n,m}$ we denote the set of all $n \times m$ complex matrices and by M_n, we denote $M_n(\mathbb{C})$ (which is a C*-algebra by Example 3.1.7).

Theorem 3.1.24. *Let ϕ be a linear map from M_n to M_m then ϕ is completely positive if and only if it admits the expression*

$$\phi(X) = \sum_{i=1}^{k} V_i^* X V_i$$

for all $X \in M_n$, where $\{V_1, \ldots, V_k\} \subseteq M_{n,m}$ for any $k \in \mathbb{N}$.

Proof. See [19] for the proof. □

Remark 3.1.25. From [19, Remark 4], we get that we may in the above theorem, require that $\{V_1, \ldots, V_k\}$ be linearly independent. In such a case, we get that $k \leq nm$.

3.2 Composite spaces and algebras

Following the DMTL approach, we are also interested in composite spaces (or algebras), i.e., larger vector spaces that are a result of combining smaller vector spaces. In this section we introduce two ways in which one can construct these composite spaces. The first part focuses on direct sums and representations while the second part is on tensor products. Tensor products play a big role in the DMTL approach, so we will use them for most of our composite spaces. We will only use direct sums to combine representations into larger representations.

3.2.1 Direct sums

This subsection is devoted to summarising *direct sums* to an extent that is sufficient for the remainder of this dissertation (this is mostly aimed at addressing some of the definitions that we need in the next section). We start with a definition that's motivated by [49, Section II.6].

Definition 3.2.1 (Direct sum). Let A and B be some vector spaces over $\mathbb{F}$, then the (external) *direct sum* of A and B, denoted $A \oplus B$, is a vector space that is defined as follows: The underlying set is the cartesian product, that is the set $A \times B$ of ordered pairs (a, b) for $a \in A$ and $b \in B$. Addition and scalar multiplication are defined by:

1. $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$,
2. $\alpha(a_1, b_1) = (\alpha a_1, \alpha b_1)$,

for $a_i \in A$ and $b_i \in B$ and $\alpha \in \mathbb{F}$. It is thus clear that the above is a vector space that can be defined as

$$A \oplus B := \text{span}_{\mathbb{F}} \left\{ (a, b) \ : \ a \in A, b \in B \right\}.$$

Remark 3.2.2. Direct sums are associative, i.e., $A_1 \oplus A_2 \oplus A_3 = (A_1 \oplus A_2) \oplus A_3 = A_1 \oplus (A_2 \oplus A_3)$, with the elements of $A_1 \oplus A_2 \oplus A_3$ being the triples (a_1, a_2, a_3) for all $a_i \in A_i$.

From [46, p. 121], we can conclude that, if $\mathcal{H}_i$ for $i \in \Lambda = \{1, 2, ..., n\}$ is a family of Hilbert spaces, then elements of the direct sum $\mathcal{H} = \bigoplus_{i \in \Lambda} \mathcal{H}_i = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \cdots \oplus \mathcal{H}_n$ are denoted by the n-tuple $(h_i)_{i \in \Lambda} = (h_1, h_2, \ldots, h_n)$. This direct sum space is itself a Hilbert space with the following inner product

$$\langle x, y \rangle = \sum_{i \in \Lambda} \langle x_i, y_i \rangle_{\mathcal{H}_i}, \tag{3.3}$$

for any $x = (x_i)_{i \in \Lambda} \in \mathcal{H}$ and $y = (y_i)_{i \in \Lambda} \in \mathcal{H}$, with $\langle \cdot, \cdot \rangle_{\mathcal{H}_i}$ as the inner product on $\mathcal{H}_i$ for each $i \in \Lambda$. The case of an infinite number of Hilbert spaces is discussed in [24, p. 24] and [46, p. 123]. In such a case, we simply set $\Lambda = \{1, 2, ...\}$ in the above arguments and assume that $\sum_{i \in \Lambda} \|x_i\|_{\mathcal{H}_i} < \infty$, for any $x = (x_i)_{i \in \Lambda} \in \mathcal{H}$. A note to make here is that $\|\cdot\|_{\mathcal{H}_i}$ is a norm on $\mathcal{H}_i$ for each $i \in \Lambda$ and $\mathcal{H}$ is still a Hilbert space (see [24, Proposition 6.2] and [46, pp. 123-124]).

The ideas about (3.3) apply to bilinear maps as well, that is if ϕ_i is a bilinear map for each $i \in \Lambda$, then equation (3.3) gives a bilinear map on $\mathcal{H}$. The above is easy to verify, we will show the linearity of one argument, the second one follows in a similar fashion. Assume that ϕ_i is a bilinear map for each $i \in \Lambda$, then for any $x = (x_i)_{i \in \Lambda} \in \mathcal{H}$, $y = (y_i)_{i \in \Lambda} \in \mathcal{H}$ and $z = (z_i)_{i \in \Lambda} \in \mathcal{H}$ define

$$\phi(x, y) := \sum_{i \in \Lambda} \phi_i(x_i, y_i), \tag{3.4}$$

then

$$\begin{aligned}\phi(x+y,z) &= \sum_{i\in\Lambda}\phi_i(x_i+y_i,z_i)\\ &= \sum_{i\in\Lambda}\big(\phi_i(x_i,z_i)+\phi_i(y_i,z_i)\big) \text{ since } \phi_i \text{ is a bilinear map}\\ &= \sum_{i\in\Lambda}\phi_i(x_i,z_i)+\sum_{i\in\Lambda}\phi_i(y_i,z_i)\\ &= \phi(x,z)+\phi(y,z).\end{aligned} \tag{3.5}$$

From the above ideas (3.5), we can also conclude that $\phi(\alpha x,z)=\alpha\phi(x,z)$ for any $\alpha\in\mathbb{F}$.

3.2.2 Representations

This subsection describes and summarises representations of algebras, this includes the role played by constructions when it comes to representations. As mentioned before, this is not an in-depth discussion, so we will mostly focus on relevant definitions and results.

Definition 3.2.3 (Representation). A *representation* of a C*-algebra A is a pair $(\mathcal{H},\pi)$ where $\mathcal{H}$ is some Hilbert space and $\pi : A\to B(\mathcal{H})$ is a $*$-homomorphism. $(\mathcal{H},\pi)$ is said to be a *faithful representation* if π is injective. If there is a vector Ω in $\mathcal{H}$ for which the linear subspace

$$\pi(A)\Omega=\big\{\pi(a)\Omega \;:\; a\in A\big\}$$

is dense in $\mathcal{H}$, then the triple $(\mathcal{H},\pi,\Omega)$ is called a *cyclic representation* of A, and Ω is termed a *cyclic vector* (or *generating vector*) for π.

In the above definition, one should recall that a set S in some metric space $(\mathcal{M},d)$ is said to be *dense* in $\mathcal{M}$, if for all $m\in\mathcal{M}$ and any $\epsilon>0$ there exists $s\in S$ such that $d(m,s)<\epsilon$.

Remark 3.2.4. If A is just an algebra then the pair $(\mathcal{H},\pi)$ will be considered a represention if π is a homomorphism (a note to make in this case is that $\mathcal{H}$ is not necessarily a Hilbert space, in Chapter 7 we will construct such a representation).

One may now ask if any arbitrary C*-algebra will always have such a representation or not, the *GNS construction* is a well-known answer to this question (see the following results).

Theorem 3.2.5. *Let A be a C*-algebra and consider ϕ as a state of A, then there exists a representation $(\mathcal{H}_\phi, \pi_\phi)$ associated with ϕ called the* GNS representation *of A.*

Proof. Let A be a C*-algebra and assume that ϕ is a state of A, then ϕ is a positive linear functional on A by Definition 3.1.20. According to [6, II.6.4.1] and [54, pp. 93-94], since ϕ is a positive linear functional on A, we can always use the GNS construction to manufacture a GNS representation $(\mathcal{H}_\phi, \pi_\phi)$ associated with ϕ. □

Remark 3.2.6. Following [6, II.6.4.2], we have that in the above theorem, there also exists a cyclic vector $\Omega_\phi \in \mathcal{H}_\phi$ associated with the state ϕ. I.e., $\pi_\phi(A)\Omega_\phi$ is dense in $\mathcal{H}_\phi$, implying that the triple $(\mathcal{H}_\phi, \pi_\phi, \Omega_\phi)$ is a cyclic representation of A.

According to Theorem 3.2.5, every state of a C*-algebra A has its own representation. Let Λ be the set of all states of A, then for each $\lambda \in \Lambda$ we get that $(\mathcal{H}_\lambda, \pi_\lambda)_{\lambda\in\Lambda}$ is a set of the representations of A associated with each state. The direct sum $(\mathcal{H}, \pi)$ of these representations is itself a representation of A defined by setting

$$\mathcal{H} = \bigoplus_{\lambda\in\Lambda} \mathcal{H}_\lambda \text{ and } \pi(a)x = (\pi_\lambda(a)x_\lambda)_{\lambda\in\Lambda} \in \mathcal{H},$$

for any $x = (x_\lambda)_{\lambda\in\Lambda} \in \mathcal{H}$ and any $a \in A$. $(\mathcal{H}, \pi)$ is called the *universal representation* of A.

Remark 3.2.7. It's easy to show that $(\mathcal{H}, \pi)$ is a representation in the above arguments. From the arguments leading up to (3.3), the direct sum is an inner product space, hence $\mathcal{H}$ is simply its completion. π is clearly a $*$-homomorphism, for the sake of argument we will show that π preserves addition (the rest of the properties can be proven in a similar way). Let $a, b \in A$ for any $x = (x_\lambda)_{\lambda\in\Lambda} \in \mathcal{H}$, then

$$\begin{aligned}
\pi(a+b)x &= (\pi_\lambda(a+b)x_\lambda)_{\lambda\in\Lambda} \\
&= (\pi_\lambda(a)x_\lambda + \pi_\lambda(b)x_\lambda)_{\lambda\in\Lambda} \text{ since } \pi_\lambda \text{ is a } *\text{-homomorphism} \\
&= (\pi_\lambda(a)x_\lambda)_{\lambda\in\Lambda} + (\pi_\lambda(b)x_\lambda)_{\lambda\in\Lambda} \text{ by Definition 3.2.1} \\
&= \pi(a)x + \pi(b)x.
\end{aligned}$$

We can at last conclude the following from the GNS construction.

Theorem 3.2.8. *Let A be a C*-algebra, then A has faithful representation, namely its universal representation.*

Proof. See the proof of Theorem 3.4.1 in [54]. □

Remark 3.2.9. A clear note to make about universal representations is that they are not limited to C*-algebras, in other words, given a set of representations for some algebra, the direct sum will always give us a universal representation of that algebra (for an example, see Chapter 7).

3.2.3 Tensor products

We start this subsection with a basic non-constructive definition of *algebraic tensor products*, followed by some key interesting properties. These properties are going to play important roles in the rest of the dissertation. A note to keep in mind here, is that every algebra in this subsection is over the same $\mathbb{F}$, unless we specify otherwise. We also make the assumption that the reader is familiar with tensor product spaces when it comes to finite dimensional vector spaces, this includes finite dimensional Hilbert spaces. A detailed discussion of the above topics is given in [26, Section 4]. Another assumption that the reader must bear in mind is that this is nowhere near a complete discussion of tensor product algebras, we are merely defining some of the words that will show up in Section 7.3 and stating some relevant results. For a more rigorous treatment of these ideas, see [10, 6].

Definition 3.2.10 (Algebraic tensor product space). Let A and B be vector spaces over the same field $\mathbb{F}$, then the vector space

$$A \odot B := \operatorname{span}\{a \otimes b \ : \ a \in A, b \in B\},$$

is their *algebraic tensor product space*, where

$$\odot : A \times B \to A \odot B : (a, b) \mapsto a \otimes b$$

is a bilinear map, such that $A \odot B$ satisfies the *universal property* of tensor products in the following sense:
For C being any vector space over $\mathbb{F}$ and $\phi : A \times B \to C$ any bilinear map, there exists a unique linear map $\phi' : A \odot B \to C$ such that $\phi'(a \otimes b) = \phi(a, b)$ for all $a \in A$ and $b \in B$.

Elements of the form $a \otimes b$ (read "a tensor b") for $a \in A$ and $b \in B$ are called *simple or elementary tensors*. $A \odot B$ is defined in a way such that the following properties hold true for all $a_1, a_2 \in A, b_1, b_2 \in B$ and $c \in \mathbb{F}$:

T.1. $c\,(a_1 \otimes b_1) = (ca_1) \otimes b_1 = a_1 \otimes (cb_1)$,

T.2. $(a_1 + a_2) \otimes b_1 = a_1 \otimes b_1 + a_2 \otimes b_1$,

T.3. $a_1 \otimes (b_1 + b_2) = a_1 \otimes b_1 + a_1 \otimes b_2$.

If in addition A and B are *-algebras then $A \odot B$ is a *-algebra where the adjoints and multiplication satisfy the following:

T.4. $(a_1 \otimes b_1)^* = a_1^* \otimes b_1^*$,

T.5. $(a_1 \otimes b_1)\,(a_2 \otimes b_2) = a_1 a_2 \otimes b_1 b_2$.

Remark 3.2.11. The elements in $A_1 \odot A_2$ cannot all be expressed as simple tensors but in many cases as linear combinations of simple tensors. One clear consequence of **T**.1 is that if $x \in A \odot B$, then

$$x = \sum a_i \otimes b_i,$$

for $a_i \in A$ and $b_i \in B$. The other consequence is that for $a_1 \in A_1$ and $a_2 \in A_2$, if $a_1 = 0$ or $a_2 = 0$, then $a_1 \otimes a_2 = 0$.

One of the basic ideas of tensor products is the notion of a *tensor product map*. This idea is a clear consequence of linear algebra and the universal property. It states that, if A_1, A_2, B_1, B_2 are vector spaces over the same field $\mathbb{F}$ and $\phi_i : A_i \to B_i$ for $i = 1, 2$ are linear maps, then there exists a unique linear map

$$\phi_1 \odot \phi_2 : A_1 \odot A_2 \to B_1 \odot B_2,$$

which is called the tensor product of ϕ_1 and ϕ_2, such that

$$\phi_1 \odot \phi_2\,(a_1 \otimes a_2) = \phi_1\,(a_1) \otimes \phi_2\,(a_2)$$

for all $a_1 \in A_1$ and $a_2 \in A_2$.

We are interested in tensor products of vector spaces, unital algebras and unital C*-algebras. In the case of unital C*-algebras, reference [10, p. 72] tells us that $A \odot B$ may carry multiple C*-norms (this was first proven by Takesaki [65]).

A C*-norm $\|\cdot\|_\alpha$ on $A \odot B$ is defined in the usual way, i.e., it's a norm that satisfies all the properties of a C*-norm as stated in Definition 3.1.6. It's a fact that C*-norms always exist for algebraic tensor products.

In general, for vector spaces A and B over the same field $\mathbb{F}$, we denote $A \otimes B$ as the completion of $A \odot B$ under some arbitrary norm defined on $A \odot B$.

Remark 3.2.12. Tensor products are also associative, i.e., for vector spaces A_1, A_2 and A_3 over the same field $\mathbb{F}$, we have that $A_1 \odot A_2 \odot A_3 = (A_1 \odot A_2) \odot A_3$. Similarly for the completions, we also have that $A_1 \otimes A_2 \otimes A_3 = (A_1 \otimes A_2) \otimes A_3$.

For C*-algebras, the following are the most natural definitions of norms on $A \odot B$.

Definition 3.2.13 (Maximal C*-norm). Let A and B be C*-algebras, then we define the *maximal C*-norm* on $A \odot B$ as

$$\|x\|_{max} = \sup \left\{ \left\| \pi(x) \right\|_{B(\mathcal{H})} \mid \pi : A \odot B \to B(\mathcal{H}) \text{ a } *\text{-homomorphism} \right\},$$

where the supremum is taken over all representations $(\mathcal{H}, \pi)$ of $A \odot B$. We will denote the completion of $A \odot B$ with respect to $\|\cdot\|_{max}$ as $A \otimes_{max} B$, which is also a C*-algebra normally referred to as the *maximal tensor product* of A and B.

Definition 3.2.14 (Minimal or spatial C*-norm). Let A and B be C*-algebras with faithful representations $\pi_A : A \to B(\mathcal{H}_A)$ and $\pi_B : B \to B(\mathcal{H}_B)$, then the minimal C*-norm on $A \odot B$ is defined by

$$\|x\|_{min} = \left\| \sum \pi_A(a_i) \otimes \pi_B(b_i) \right\|_{B(\mathcal{H}_A \otimes \mathcal{H}_B)},$$

for all $x = \sum a_i \otimes b_i \in A \odot B$. We denote the completion of $A \odot B$ with respect to $\|\cdot\|_{min}$ as $A \otimes_{min} B$, which is also a C*-algebra normally referred to as the *minimal tensor product* of A and B.

The remainder of [10, Chapter 3] verifies some of the properties related to $\|\cdot\|_{max}$ and $\|\cdot\|_{min}$.

Chapter 4

An outline for the case of quantum states

In this dissertation, we are interested in formulating a distance between channels. We define the notion of a channel in Chapter 5 and proceed to show that channels are unital completely positive maps between two unital C*-algebras. From Remark 3.1.23, we know that states are special cases of u.c.p. maps, hence they are a special case of channels, so a distance between channels gives us a distance between states as a special case.

The main aim of this chapter and Chapter 5 is to motivate the method that we will use to formulate our quantum Wasserstein distance of order 1 between channels. This chapter focuses on the case of states. It will further illustrate the connection between the states of a physical quantum system and the states of a C*-algebra (clarifying the use of the word "state" in the context of C*-algebras).

4.1 The setting

In this section, we review some basic quantum theory. We will assume that the reader is familiar with both the Hilbert space and density formalism of quantum mechanics. A detailed discussion of these is given in [56].

The incorporation of time evolution in a quantum system gives rise to multiple mathematical formulations. The *Schrödinger picture* and the *Heisenberg picture* are the two most important such formulations. In the next paragraph, we give a basic summary of the main assumptions in each formulation. For a detailed discussion on the pros and cons of each formulation, includ-

ing the equivalence between them, see [53, Chapter 14 (Quantum dynamics)].

In the more commonly used Schrödinger picture, the states of a quantum system are set to be time-dependent, while the observables are time-independent. In the Heisenberg picture, the states of a quantum system are set to be time-independent, while the observables are time-dependent.

In this dissertation, we will focus on the Heisenberg picture, i.e., the states of our quantum systems are time-independent.

Using a few ideas from Chapter 3, let

$$\mathcal{H}_{\mathcal{A}} = \mathbb{C}^d$$

be a finite dimensional Hilbert space associated with some isolated physical quantum system $\mathcal{A}$. A standard result of linear algebra implies that the unital C*-algebra of linear maps (or operators) on $\mathcal{H}_{\mathcal{A}}$ denoted $L(\mathcal{H}_{\mathcal{A}})$ can be represented by complex matrices M_d, i.e., $L(\mathcal{H}_{\mathcal{A}}) = M_d$. Let

$$A = M_d,$$

then A is a unital C*-algebra sometimes referred to as the *observable algebra* of $\mathcal{A}$.

For simplicity, in the case of Hilbert spaces, we will use Dirac notation to represent vectors. Let the canonical basis of $\mathbb{C}^d$ be the set of vectors $\{|1\rangle, ..., |d\rangle\}$ such that a pure state $|\psi\rangle$ of the quantum system can be written as

$$|\psi\rangle = \psi_1 |1\rangle + \cdots + \psi_d |d\rangle$$

with

$$\langle\psi| = \overline{\psi_1} \langle 1| + \cdots + \overline{\psi_d} \langle d|,$$

for $\psi_1, \psi_2, ..., \psi_d \in \mathbb{C}$. In this case we denote the inner product between the normalised states $|\psi\rangle, |\phi\rangle \in \mathcal{H}_{\mathcal{A}}$ as

$$\langle\psi|\phi\rangle,$$

with

$$|\psi\rangle \langle\psi|$$

defining a projection operator on $\mathcal{H}_{\mathcal{A}}$ given by

$$\left(|\psi\rangle \langle\psi|\right) |\phi\rangle = \langle\psi|\phi\rangle |\psi\rangle.$$

Since $T|\alpha\rangle \in \mathcal{H}_\mathcal{A}$, for all $|\alpha\rangle \in \mathcal{H}_\mathcal{A}$, it follows that $|\psi\rangle\langle\psi| \in L(\mathcal{H}_\mathcal{A})$. We also have that the inner product between $|\psi\rangle$ and $T|\alpha\rangle$ is denoted as

$$\langle\psi|T|\alpha\rangle .$$

According to the fundamental theorem of invertible matrices, we know that a matrix $T \in A$ is invertible if and only if $\det(T) \neq 0$. Hence, by Definition 3.1.2 the spectrum of T is given by

$$\sigma(T) = \{\lambda \in \mathbb{C} : \det(\lambda 1_A - T) = 0\},$$

which is a set of all the eigenvalues of T. The eigenvalue equation in Dirac notation, can be stated as

$$T|i\rangle = \lambda_i |i\rangle , \tag{4.1}$$

where the λ_is are the eigenvalues (i.e., $\lambda_i \in \sigma(T)$) with their associated eigenvectors $|i\rangle \in \mathcal{H}_\mathcal{A}$ (which are normalisable). To summarise [49, Theorem 3.21] further, we have by the spectral theorem that if T is self-adjoint (i.e., $T^* = T$), then there exists an orthonormal set of eigenvectors of T that span $\mathcal{H}_\mathcal{A}$ such that T admits the expression:

$$T = \sum_{i=1}^{d} \lambda_i |i\rangle\langle i| ,$$

with λ_i being an eigenvalue of T associated with the eigenvector $|i\rangle \in \mathcal{H}_\mathcal{A}$ for each i. By orthornormal, we simply mean that the elements of the set satisfy the following property:

$$\langle i|j\rangle = \delta_{ij} = \begin{cases} 1 & \text{when } \ i = j, \\ 0 & \text{when } \ i \neq j, \end{cases}$$

with δ_{ij} being the standard Kronecker delta.

The following are definitions of specific elements of A.

Definition 4.1.1 (Positive matrix). A matrix $X \in A = M_d$ is said to be a *positive matrix* if X is self-adjoint and

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

for any $|\psi\rangle \in \mathcal{H}_\mathcal{A}$. The set of all positive matrices in $A = M_d$ is denoted $A^+ = M_d^+$.

Considering the above definition, since A is a unital C*-algebra, take $X \in A$ such that $X \geq 0$. Definition 3.1.10 implies X is self-adjoint (i.e., $X^* = X$) and $\sigma(X) \subset [0, \infty)$. From (4.1) we get that

$$X |i\rangle = \lambda_i |i\rangle ,$$

where the λ_is are the eigenvalues of X (i.e., $\lambda_i \in \sigma(X)$ such that $\lambda_i \geq 0$) with their associated eigenvectors $|i\rangle \in \mathcal{H}_{\mathcal{A}}$ being an orthonormal basis of $\mathcal{H}_{\mathcal{A}}$ (this is guaranteed by [49, Spectral theorem]). Thus for any $|\psi\rangle \in \mathcal{H}_{\mathcal{A}}$ we have that

$$|\psi\rangle = \sum_i \psi_i |i\rangle ,$$

with each $\psi_i \in \mathbb{C}$. This implies that

$$\begin{aligned}
\langle\psi| X |\psi\rangle &= \sum_i \overline{\psi}\psi \langle i| X |i\rangle \\
&= \sum_i \overline{\psi_i}\psi_i \langle i| \lambda_i |i\rangle \text{ by (4.1)} \\
&= \sum_i \overline{\psi_i}\psi_i\lambda_i \langle i|i\rangle \\
&= \sum_i \overline{\psi_i}\psi_i\lambda_i \text{ the } |i\rangle\text{'s are orthonormal} \\
&\geq 0 \text{ since } \overline{\psi_i}\psi_i \geq 0 \text{ and } \lambda_i \geq 0 \text{ for each } i.
\end{aligned}$$

Thus Definition 3.1.10 implies Definition 4.1.1. The converse is also true. Let $X \in A$ such that X is a positive matrix, Definition 4.1.1 implies that X is self-adjoint and

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

for any $|\psi\rangle \in \mathcal{H}_{\mathcal{A}}$. Let λ_i be the eigenvalues of X with the associated eigenvectors given by $|i\rangle \in \mathcal{H}_{\mathcal{A}}$ (we again assume that these form an orthonormal basis), then (4.1) implies that $X |i\rangle = \lambda_i |i\rangle$ and since $|i\rangle \in \mathcal{H}_{\mathcal{A}}$, it follows that

$$0 \leq \langle i|X|i\rangle = \lambda_i \langle i|i\rangle = \lambda_i,$$

implying that $\sigma(X) \subset [0, \infty)$.

Remark 4.1.2. The above arguments imply that Definitions 3.1.10 and 4.1.1 are in agreement (in the sense that one implies the other). We can always drop the condition that X is self-adjoint in Definition 4.1.1 (see [49, Proposition 3.17]), but that is not necessary for our goal.

One should also note that in the physics literature, the adjoint X^* of any $X \in A$ is sometimes denoted as $X^\dagger$ with a self-adjoint matrix sometimes being called a *Hermitian matrix*. We will stick with the notation that we have been using thus far.

Definition 4.1.3 (Density matrix). $\rho \in A = M_d$ is called a *density matrix* if it satisfies the following properties:

1. $\mathbf{Tr}(\rho) = 1$ 'Trace condition'
2. $\rho \geq 0$ 'positivity condition'

The **Tr** in the above definition (including the rest of this dissertation) is the usual trace for matrices.

According to [56, Part I, Section 2.4], a density matrix represents a state of some physical quantum system. One can then make the conclusion that the set of density matrices in A is the set of the physical states of system $\mathcal{A}$. Under this representation of physical states, let $\rho \in A = M_d$ be a physical state (in other words, a density matrix) and let $X \in M_d$ be an observable of quantum system $\mathcal{A}$, then

$$\mathbf{Tr}(\rho X)$$

is the expected value of the observable X, when the system is in the state ρ.

The next property tells us that any density matrix induces a state on the unital C*-algebra M_d. This result is what moves us into the algebraic framework. Before we start, let B_{jk} denote the j,k'th entry of any matrix $B \in M_d$ and write $S = M_d$ for the matrix algebra containing the density matrices of system $\mathcal{A}$. Another note to make here is that the multiplicative unit of A is given by I_d, i.e., $1_A = I_d$.

Proposition 4.1.4. *For the unital C*-algebra M_d of $d \times d$ complex matrices, let ω be any map from M_d into $\mathbb{C}$, then ω is a state on M_d if and only if there exists a density matrix $\rho \in M_d$ such that $\omega(A) = \mathbf{Tr}(\rho A)$ for every $A \in M_d$.*

Proof. Assume that ω is a state and let E_{ij}'s be the basis elements of the complex vector space M_d, where E_{ij} is a $d \times d$ matrix with 1 in the i,j position and zeroes in the rest of its entries, then any matrix $A \in M_d$ can be written as

$$A = \sum_{ij} A_{ij} E_{ij}. \tag{4.2}$$

Given that ω is a state on M_d, Definition 3.1.20 implies that $\omega(I_d) = 1$ and $\omega(B) \geq 0$ for I_d the unit of M_d and any $B \in M_d$ such that $B \geq 0$. Let $\rho \in M_d$, such that

$$\rho = \sum_{ij} \omega\left(E_{ji}\right) E_{ij}. \tag{4.3}$$

Since ω is a linear functional, (4.2) implies that

$$\omega(A) = \sum_{ij} A_{ij}\omega\left(E_{ij}\right). \tag{4.4}$$

Let us consider the following trace,

$$\begin{aligned} \mathbf{Tr}(\rho A) &= \sum_{k} (\rho A)_{kk} \\ &= \sum_{jk} \rho_{kj} A_{jk} \\ &= \sum_{jk} \omega\left(E_{jk}\right) A_{jk} \text{ by (4.2) and (4.3)} \\ &= \omega(A) \text{ by (4.4)}. \end{aligned} \tag{4.5}$$

Since ω is a positive map, then by [6, II.6.2], we have for any $B \in M_d$ that $\omega(B^*) = \overline{\omega(B)}$. Hence, given $E_{ij} = E_{ji}^*$, we get that $\overline{\omega\left(E_{ij}\right)} = \omega\left(E_{ij}^*\right) = \omega\left(E_{ji}\right)$. But from (4.3) we have that

$$\rho^* = \sum_{ij} \overline{\omega\left(E_{ij}\right)} E_{ij} = \sum_{ij} \omega\left(E_{ji}\right) E_{ij},$$

thus $\rho = \rho^*$, i.e., ρ is self-adjoint. Since ρ is self-adjoint, the spectral decomposition of ρ is

$$\rho = \sum_{i} \lambda_i P_i \text{ for } P_i \in M_d, \tag{4.6}$$

where the λ_i's are the eigenvalues of ρ while the P_i's are projection operators onto 1-dimensional orthogonal subspaces of the eigenspaces. In other words, the P_i's are self-adjoint, orthogonal, indempotent and have a trace of 1, i.e., $P_i^* = P_i$, $P_iP_j = 0$ for $i \neq j$ while $P_iP_j = P_i$ for $i = j$ and $\mathbf{Tr}(P_i) = 1$. These properties imply the following

$$\rho P_i = \lambda_i P_i P_i = \lambda_i P_i \tag{4.7}$$

Hence,

$$\begin{aligned} \lambda_i = & \ \mathbf{Tr}\left(\rho P_i\right) \\ = & \ \omega\left(P_i\right) \text{ follows from (4.5)} \\ = & \ \omega\left(P_i P_i\right) \quad \text{because } P_i \text{ is indempotent} \\ = & \ \omega\left(P_i^* P_i\right) \text{ because } P_i \text{ is self-adjoint} \\ \geq & \ 0 \text{ see Remark 3.1.18 .} \end{aligned} \tag{4.8}$$

So each eigenvalue λ_i of ρ is in $[0, \infty)$. This implies that $\sigma(\rho) \subset [0, \infty)$, thus by Definition 3.1.10, we have that $\rho \geq 0$. We also have that $\omega\left(I_d\right) = 1$, this implies that $\mathbf{Tr}\left(\rho I_d\right) = \mathbf{Tr}\left(\rho\right) = 1$, hence by the definition of a density matrix we have that $\rho \in S$.

For the converse, suppose ρ is any density matrix (i.e., $\rho \in S$), and consider the linear functional $\omega(A) = \mathbf{Tr}\left(\rho A\right)$ for any $A \in M_d$. Let $B \in S$, then by Remark 3.1.18, we get that there exists $C \in M_d$ such that $B = C^* C$. Hence for B we have

$$\begin{aligned} \omega(B) = & \ \mathbf{Tr}\left(\rho C^* C\right) \\ = & \ \mathbf{Tr}\left(C \rho C^*\right) \text{ by the cyclic property of } \mathbf{Tr}. \end{aligned}$$

By [6, Corollary II.3.1.5], we have that $C\rho C^* \geq 0$. This means that the eigenvalues of $C\rho C^*$ are non-negative, hence their sum is non-negative. Since the trace of a matrix is equal to the sum of the eigenvalues of the matrix, we have that $\omega(B) \geq 0$. We also have that $\omega(I_d) = \mathbf{Tr}\left(\rho I_d\right) = \mathbf{Tr}\left(\rho\right) = 1$, by Definition 4.1.3. By Definition 3.1.20, ω is indeed a state. □

Remark 4.1.5. The above proposition essentially states that the density matrices of S can be compared (distance-wise) by comparing the states of M_d. For instance, if we have a metric, say W_1 on the states of M_d then we can use it to compare any two known or given density operators from S.

We can thus conclude that the physical states of some quantum system are equivalent to the states of its observable algebra.

A note to make here is that in classical information theory the basic (or smallest) unit of information is a binary digit often called a *bit*, and at any

point a bit has a value of 0 or 1 but not both (i.e., the state of a bit can only be a 0 or a 1). In quantum information theory the story is a little different in that the quantum version of a bit is called a *qubit* and is at any time a linear combination of the physical states $|0\rangle$ and $|1\rangle$. Hence a qubit is a physical state of the quantum system $\mathbb{C}^2$, which is spanned by $\{|0\rangle, |1\rangle\}$. A *qudit* is a generalised qubit, i.e., it is physical state of the quantum system $\mathbb{C}^d$.

For the remainder of this chapter and the next, we will make use of the *partial trace*, [56, Part I, Section 2.4.3] defined as the normal trace over a specific factor of a tensor product space, i.e., if $A_i = M_{q_i}$ such that $A = A_1 \otimes A_2$ then, for a simple tensor $X \otimes Y \in A$, we have that

$$\mathbf{Tr}_1(X \otimes Y) = \mathbf{Tr}(X)Y \text{ and } \mathbf{Tr}_2(X \otimes Y) = X\mathbf{Tr}(Y). \tag{4.9}$$

Our focus from here onwards is on composite spaces as defined through the tensor products of Subsection 3.2.3. According to [56, Part I, Subsection 2.2.8] a *composite system* is a quantum system made up of two or more distinct quantum systems, with the composite system simply defined as their tensor product.

The main advantage of representing the physical states of a quantum system as density matrices is how it describes the subsystems of a composite system (see the following remark).

Remark 4.1.6. The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if we have systems $\mathcal{A}$ and $\mathcal{B}$, and system $\mathcal{A}$ is prepared in the state ρ while $\mathcal{B}$ is prepared in the state σ, then the joint state of the composite system $\mathcal{AB}$ is given by:

$$\rho^{AB} = \rho \otimes \sigma. \tag{4.10}$$

One of the deepest applications of formulating quantum mechanics using density matrices is how they describe the subsystems of a composite quantum sysetm. Given a state ρ^{AB} on the composite system, we can define a *reduced density matrices* ρ^A and ρ^B of systems $\mathcal{A}$ and $\mathcal{B}$ respectively by

$$\rho^A := \mathbf{Tr}_B\left(\rho^{AB}\right) \text{ and } \rho^B := \mathbf{Tr}_A\left(\rho^{AB}\right).$$

If we take the state given in (4.10) into consideration, we get by the trace condition of density matrices and (4.9) that

$$\rho^A = \mathbf{Tr}_B\left(\rho^{AB}\right) = \mathbf{Tr}_B\left(\rho \otimes \sigma\right) = \mathbf{Tr}\left(\sigma\right)\rho = \rho$$

and

$$\rho^B = \mathbf{Tr}_A\left(\rho^{AB}\right) = \mathbf{Tr}\left(\rho\right)\sigma = \sigma.$$

It is thus clear that if we know the state of a composite system, we can desribe the states of the component physical systems.

Next, we move towards the DMTL approach and Chapter 2. For consistency between our work and [29], we will introduce some notation from [29]. Let the canonical basis of $\mathbb{C}^d$ be the set of vectors

$$\{|1\rangle, ..., |d\rangle\}$$

and by taking the tensor product n times, set

$$\mathcal{H}_n = \mathbb{C}^d \otimes \cdots \otimes \mathbb{C}^d$$

to be the Hilbert space of n qudits. I.e., $\mathcal{H}_n = \left(\mathbb{C}^d\right)^{\otimes n}$. Let $\mathcal{O}_n$ be the set of the self-adjoint operators on $\mathcal{H}_n$, denote $\mathcal{O}_n^T \subset \mathcal{O}_n$ as the subset of traceless self-adjoint operators and let $\mathcal{S}_n \subset \mathcal{O}_n$ be the subset of density operators.

The Kronecker product [26, Section 4] implies that $\mathcal{H}_n$ can be treated as $\mathbb{C}^{d^n}$ with the set of all operators on $\mathcal{H}_n$ contained in the unital C*-algebra M_{d^n}, in particular, $\mathcal{S}_n \subset M_{d^n}$ and $\mathcal{O}_n^T \subset \mathcal{O}_n \subset M_{d^n}$.

4.2 Neighbouring states

In this section we use the notation from the previous section to introduce the notion of *neighbouring states*.

The following definition from [29] plays an important role, the definition is stated in a slightly different manner here to avoid confusion.

Definition 4.2.1 (Neighbouring density operators). $\rho, \sigma \in \mathcal{S}_n$ are called *neighbouring density operators* if $\mathbf{Tr}_i\rho = \mathbf{Tr}_i\sigma$ for some $i \in [n] = \{1, ..., n\}$, i.e., they coincide after discarding a suitable qudit. Let $\mathcal{N}_n \in \mathcal{O}_n^T$ be set of the differences between couples of neighbouring density operators as follows:

$$\mathcal{N}_n = \bigcup_{i=1}^{n} \mathcal{N}_n^{(i)}, \qquad \mathcal{N}_n^{(i)} = \{\rho - \sigma \ ; \ \rho, \sigma \in \mathcal{S}_n, \mathbf{Tr}_i\rho = \mathbf{Tr}_i\sigma\}, \qquad i \in [n]$$

and write

$$\mathcal{B}_n = \left\{ \sum_{i=1}^{n} p_i \left(\rho^{(i)} - \sigma^{(i)}\right) \ : \ p_i \geq 0, \ \sum_{i=1}^{n} p_i = 1, \ \rho^{(i)}, \sigma^{(i)} \in \mathcal{S}_n, \ \mathbf{Tr}_i\rho^{(i)} = \mathbf{Tr}_i\sigma^{(i)} \right\}$$

for the *convex hull* of $\mathcal{N}_n$.

Remark 4.2.2. According to [8, Part I, Section 2.1.4], A *convex hull* of some set C, denoted $\operatorname{conv} C$, is the set of all convex combinations of points in C, i.e.,

$$\operatorname{conv} C = \left\{ \sum_{i=1}^{k} \theta_i x_i \ : \ x_i \in C, \theta_i \geq 0, \sum_{i=1}^{k} \theta_i = 1 \right\}.$$

In addition, $\operatorname{conv} C$ is the smallest convex set containing C. The above clearly implies that we can simply write $\mathcal{B}_n = \operatorname{conv} \mathcal{N}_n$.

The remark above allows us to note an elementary result which will be relevant in Section 6.1 when proving the absorbing property.

Lemma 4.2.3. *Let $\mathcal{X}$ be a real vector space. Consider any symmetric subset $\mathcal{N}$ of $\mathcal{X}$ such that $\mathcal{X} = \operatorname{span} \mathcal{N}$. Then its convex hull $\mathcal{C} = \operatorname{conv} \mathcal{N}$ is absorbing for $\mathcal{X}$.*

Proof. For any non-zero $x \in \mathcal{X}$ we can write $x = s_1 x_1 + \cdots + s_k x_k$ for some $k < \infty$, $x_j \in \mathcal{N}$ and $s_j > 0$. Let $t = s_1 + \cdots + s_k$ and $p_j = s_j / t$ to have $x = ty$ with $y = p_1 x_1 + \cdots + p_k x_k \in \mathcal{C}$. The case $x = 0$ is trivial. □

In the DMTL approach [29], they use Definition 4.2.1 and the trace norm to define a metric on $\mathcal{S}_n$. It is possible to do this without a prior norm. One just needs to show that $\mathcal{O}_n^T$ is a real vector space and $\mathcal{B}_n$ is a unit ball (as defined in Definition 2.2.5), then Definition 2.3.1 and Proposition 2.3.7 will give us a norm on $\mathcal{O}_n^T$, which induces a metric on $\mathcal{S}_n$ (because for any $\rho, \sigma \in \mathcal{S}_n$ we clearly have that $\rho - \sigma \in \mathcal{O}_n^T$, implying that $\mathcal{B}_n \subset \mathcal{O}_n^T$). What you end up with is the same metric (or a more general version of it) as given in [29, Definition 7].

One may now wonder if the metric induced by Proposition 2.3.7 does indeed produce a quantum version of the classical Wasserstein distance of order 1. The simplest answer is a yes and it follows from [29, Proposition 6], where they prove that this induced metric can indeed be reduced to the classical version under some assumptions. Hence, we can proceed to use the term "Wasserstein distance of order 1" in the remainder of this dissertation.

4.3 The W_1 metric for states

We can now define a metric W_1 on the set of states as follows. Let

$$A_j = M_d \text{ and } B_j = \mathbb{C}.$$

Then as explained before, via the Kronecker product, we get

$$A = \bigotimes_{j=1}^{n} A_j = M_{d^n} \text{ and } B = \bigotimes_{j=1}^{n} B_j = \mathbb{C}.$$

The reason for formulating things in this way, is to clarify the transition to channels (which will be discussed in the next chapter) instead of states, where (in the current context) the B_j's will be matrix algebras. Ultimately the A_j's and B_j's will be arbitrary unital C*-algebras (see Section 7.3).

It should also be clear that $I_{d^n} = \bigotimes_{j=1}^{n} I_d$ and $1 = \bigotimes_{j=1}^{n} 1$ are the units of A and B respectively. We can now consider the space of maps

$$L(A, B) = L(M_{d^n}, \mathbb{C}),$$

and define

$$L_S(A, B) := \left\{\eta \in L(A, B) \ : \ \eta(I_{d^n}) = 1\right\}.$$

Let $\mathcal{L}$ be any set of states on A, that is $\mathcal{L} \subset L_S(A, B)$ by Definition 3.1.20. Note now that for any $\eta, \theta \in \mathcal{L}$, we have $(\eta - \theta)(I_{d^n}) = 0$, thus one can define the real vector space

$$\mathcal{O} = \{\lambda \in \operatorname{span}_{\mathbb{R}} \mathcal{L} : \lambda(I_{d^n}) = 0\}.$$

We now obviously have that, $\eta - \theta \in \mathcal{O}$, so defining a norm on $\mathcal{O}$ induces a metric on the set of states. This is the route we will follow to define the metric W_1. The notion of neighbouring states is what allows us to define the norm, which will be denoted $\|\cdot\|_{W_1}$.

Note that a partial trace as in (4.9) removes a factor of the tensor product, e.g., for a simple tensor $a_1 \otimes a_2 \otimes a_3 \otimes \cdots \otimes a_n \in A$, $\mathbf{Tr}_2$ gives

$$\begin{aligned}\mathbf{Tr}_2(a_1 \otimes a_2 \otimes a_3 \otimes a_4 \otimes \cdots \otimes a_n) &= a_1 \otimes \mathbf{Tr}(a_2) \otimes a_3 \otimes a_4 \otimes \cdots \otimes a_n \\ &= \mathbf{Tr}(a_2)\ a_1 \otimes a_3 \otimes a_4 \otimes \cdots \otimes a_n \\ &\in A_1 \otimes A_3 \otimes A_4 \otimes \cdots \otimes A_n,\end{aligned}$$

where $A_1 \otimes A_3 \otimes A_4 \otimes \cdots \otimes A_n$ is A with A_2 left out of the tensor product. More generally, we write

$$A_{\widehat{j}} = A_1 \otimes ... \widehat{A}_j ... \otimes A_n \quad \text{and} \quad B_{\widehat{j}} = B_1 \otimes ... \widehat{B}_j ... \otimes B_n,$$

where these are A and B with A_j and B_j respectively left out of the tensor products, for any j. Note that $\widehat{A}_j$ is used to indicate the factor that is left out of the tensor product (the same applies to $\widehat{B}_j$). We correspondingly use the same notation for elementary tensors. Any $\eta \in L(A, B)$, can be restricted to the map

$$\eta_{\widehat{j}} = \eta|_{A_{\widehat{j}}}(a_1 \otimes ...\widehat{a}_j ... \otimes a_n) = \eta(a_1 \otimes ... \otimes a_{j-1} \otimes 1_{A_j} \otimes a_{j+1} \otimes ... \otimes a_n). \quad (4.11)$$

The above (4.11) is sufficient for this case, unlike in the later, more general cases (see Chapter 6 and Section 5.3). The reason for this being the fact that $B_{\widehat{j}} = \mathbb{C} = B$ (see the Kronecker product [26]), hence $\eta_{\widehat{j}} : A_{\widehat{j}} \to B_{\widehat{j}}$. The notation $\widehat{a}_j$ indicates the absence of a_j in the elementary tensor. We can now bring in a notion of neighbouring states (similar to Definition 4.2.1, but more abstract), and say $\eta, \theta \in \mathcal{L} \subset L(A, B)$ are neighbouring states if and only if $\eta_{\widehat{j}} = \theta_{\widehat{j}}$ for some jth factor of the tensor product. We can now define

$$\mathcal{N}_j = \{\eta - \theta \ : \ \theta, \eta \in \mathcal{L} \text{ with } \theta_{\widehat{j}} = \eta_{\widehat{j}}\} \quad (4.12)$$

and

$$\mathcal{N} = \bigcup_{j=1}^{n} \mathcal{N}_j.$$

Allowing us to then define $\|\cdot\|_{W_1}$ as the norm on $\mathcal{O}$ which has the convex hull

$$\mathcal{C} = \operatorname{conv} \mathcal{N}$$

of $\mathcal{N}$ as its unit ball. I.e.,

$$\|\omega\|_{W_1} = \inf\{t \geq 0 : \omega \in t\mathcal{C}\} \quad (4.13)$$

where $t\mathcal{C} = \{t\omega : \omega \in \mathcal{C}\}$ for any real number t. According to Proposition 2.3.7, the above is only a norm when $\mathcal{C}$ is indeed a unit ball (as given in Definition 2.2.5). The proof that $\mathcal{C}$ is a unit ball will be given in a more general setup, in later chapters. This norm clearly induces a metric on $\mathcal{L}$, since the difference between any two elements of $\mathcal{L}$ is in $\mathcal{O}$.

An alternative but similar (and equivalent) approach, which defines a norm on the set of density matrices (or operators) $\mathcal{S}_n$ will be discussed in the next chapter. This is done in a more general setting (see the end of Section 5.4), where we start with channels and show how this reduces to the case of density operators.

Chapter 5

W_1 between quantum channels

This chapter is aimed developing a quantum Wasserstein distance of order 1 between quantum channels by extending the approach for states. We do this in a finite dimensional setting (see Section 5.3 and Section 5.4). A much more general setup is covered Section 7.3. In this chapter, we try to give a complete presentation of the finite dimensional W_1, we leave the major proofs of the W_1 to the general setup that will follow after this chapter.

5.1 Quantum channels

This section defines the notion of a *quantum channel* and related background, while the subsequent sections are aimed at formulating the Wasserstein distance of order 1 between quantum channels. This is all done in a finite dimensional setup.

In quantum information theory we may be required to transmit quantum information, say from one quantum system $\mathcal{B}$ to another $\mathcal{A}$. Since the states of a quantum can be described by density matrices (see Chapter 4), we know that a linear map $\mathcal{E}$ that maps any density matrix of system $\mathcal{B}$ to a denstiy matrix in system $\mathcal{A}$ may be sufficient. The question is what other properties must $\mathcal{E}$ have, in order to achieve this goal? We will briefly discuss some ideas that will lead us to an answer for this question and more. Most of the ideas presented here are a summary of the work done in [26] and [56].

Fix an orthonormal basis $|1^{\mathcal{B}}\rangle, ..., |r^{\mathcal{B}}\rangle$ for the Hilbert space $\mathcal{H}_{\mathcal{B}}$ of the system $\mathcal{B}$. Similarly fix $|1^{\mathcal{A}}\rangle, ..., |q^{\mathcal{A}}\rangle$ for system $\mathcal{A}$. Represent the observable algebra of $\mathcal{B}$ in terms of the given basis as the matrix algebra $B = M_r$, and similarly use $A = M_q$ for $\mathcal{A}$.

To keep the distinction between states (or density matrices in our case) and observables conceptually and notationally clear, we also write $S_{\mathcal{A}} = M_q$ for the matrix algebra containing the density matrices of $\mathcal{A}$, and similarly $S_{\mathcal{B}} = M_r$ for $\mathcal{B}$.

The first property that $\mathcal{E}$ needs is positivity because Definition 4.1.3 tells us that density matrices are positive, so $\mathcal{E}$ has to satisfy Definition 3.1.17. However, quantum systems are not necessarily isolated, they can be entangled with other systems, i.e., they can be a part of some composite system (in other words a tensor product). Hence, positivity is not always sufficient, in such cases we need complete positivity.

The second property also follows from the need to guarantee that a density matrix is mapped to a density matrix, the property is that $\mathcal{E}$ must be *trace-preserving*, i.e., for any $X \in B$ we require that $\mathbf{Tr}(X) = \mathbf{Tr}(\mathcal{E}(X))$.

Hence, the map $\mathcal{E}$ such that

$$\mathcal{E} : S_{\mathcal{B}} \to S_{\mathcal{A}},$$

has to be completely positive and trace-preserving.

We start with a simpler definition of a completely positive map (simpler in this context, when compared to Definition 3.1.22). The definition is equivalent to the one given in Chapter 3 via Theorem 3.1.24. We will then add the trace-preserving property, in order to get a map that satisfies our two properties.

Definition 5.1.1 (Completely positive map). A map $\mathcal{E} : S_{\mathcal{B}} \to S_{\mathcal{A}}$ is said to be *completely positive* if it is given by

$$\mathcal{E}(X) = \sum_{i=1}^{k} V_i X V_i^*$$

for all $X \in B$, where $\{V_1, \ldots, V_k\} \subseteq M_{q,r}$ and any $k \in \mathbb{N}$.

Remark 5.1.2. Definition 5.1.1 makes use of $V_i (\cdot) V_i^*$ instead of $V_i^* (\cdot) V_i$ as in Theorem 3.1.24, but the former is equivalent to $(V_i^*)^* (\cdot) V_i^*$, because if we introduce $W_i = V_i^*$ for each i, then we have that $W_i^* (\cdot) W_i$.

From here onwards $\mathcal{E}$ refers to a completely positive map as given in the above definition. It is clear that $\mathcal{E}$ is a linear map and it maps $S_{\mathcal{B}}$ into $S_{\mathcal{A}}$, this follows from matrix addition and multiplication. We now verify that this map indeed satisfies our required properties. The following proposition is for the first property.

Proposition 5.1.3. *$\mathcal{E}$ as given in Definition 5.1.1 is a positive map, i.e., if $X \in S_{\mathcal{B}}^{+}$ then $\mathcal{E}(X) \in S_{\mathcal{A}}^{+}$.*

Proof. Let $X \in S_{\mathcal{B}}^{+} = M_r^{+}$, then by Definition 4.1.1 we have that $X = X^*$ and $\langle\psi_{\mathcal{B}}| X |\psi_{\mathcal{B}}\rangle \geq 0$ for any $|\psi_{\mathcal{B}}\rangle \in \mathcal{H}_{\mathcal{B}}$. It is also clear by matrix multiplication that $V_i X V_i^* \in S_{\mathcal{A}} = M_q$, for any $V_i \in M_{q,r}$. Note that $V_i^* |\psi_{\mathcal{A}}\rangle \in \mathcal{H}_{\mathcal{B}}$ for any $|\psi_{\mathcal{A}}\rangle \in \mathcal{H}_{\mathcal{A}}$, so denote $|\psi_{\mathcal{B}}\rangle = V_i^* |\psi_{\mathcal{A}}\rangle$, which implies that $\langle\psi_{\mathcal{B}}| = \langle\psi_{\mathcal{A}}| V_i$. Consider

$$\begin{aligned}(V_i X V_i^*)^* &= \left((V_i)\,(X V_i^*)\right)^* \\ &= \left((X V_i^*)^*\,(V_i)^*\right) \\ &= \left((V_i^*)^*\,X^*\right) V_i^* \\ &= V_i X V_i^* \qquad \text{by Definition 3.1.3,}\end{aligned} \tag{5.1}$$

and

$$\begin{aligned}\langle\psi_{\mathcal{A}}| V_i X V_i^* |\psi_{\mathcal{A}}\rangle &= \langle\psi_{\mathcal{B}}| X |\psi_{\mathcal{B}}\rangle \\ &\geq 0.\end{aligned} \tag{5.2}$$

Hence, 5.1 and 5.2 imply that $V_i X V_i^* \in S_{\mathcal{A}}^{+}$. A sum of positive matrices is clearly positive so $\mathcal{E}(X) \in S_{\mathcal{A}}^{+}$. □

For the second property, we get that assuming

$$\sum_{i=1}^{k} V_i^* V_i = 1_B$$

in Definition 5.1.1 guarantees that $\mathcal{E}$ preserves the trace. This is summarised by the following proposition.

Proposition 5.1.4. *For $\mathcal{E}$ as given in Definition 5.1.1, if*

$$\sum_{i=1}^{k} V_i^* V_i = 1_B,$$

then $\mathbf{Tr}(\mathcal{E}(X)) = \mathbf{Tr}(X)$ for any $X \in S_{\mathcal{B}}$.

Proof. Let $\mathcal{E}$ be as given in Definition 5.1.1. For any $X \in S_{\mathcal{B}}$ assume

$$\sum_{i=1}^{k} V_i^* V_i = 1_B,$$

then by the properties of the trace (trace of a sum and the cyclic property) we have that

$$\begin{aligned}
\mathbf{Tr}(\mathcal{E}(X)) &= \mathbf{Tr}\left(\sum_{i=1}^{k} V_i X V_i^*\right) \\
&= \sum_{i=1}^{k} \mathbf{Tr}(V_i X V_i^*) \\
&= \sum_{i=1}^{k} \mathbf{Tr}(V_i^* V_i X) \\
&= \mathbf{Tr}\left(\left(\sum_{i=1}^{k} V_i^* V_i\right) X\right) \\
&= \mathbf{Tr}(1_B X) \\
&= \mathbf{Tr}(X).
\end{aligned}$$

□

We now use some of these properties to define a map that satisfies all our required properties.

Definition 5.1.5 (Quantum channel). A map $\mathcal{E} : B \to A$ is said to be a *quantum channel* if it is completely positive and satisfies Proposition 5.1.4. I.e., it is given by

$$\mathcal{E}(X) = \sum_{i=1}^{k} V_i X V_i^*$$

for all $X \in B$, where $\{V_1, \ldots, V_k\} \subseteq M_{q,r}$ with $k \in \mathbb{N}$ and

$$\sum_{i=1}^{k} V_i^* V_i = 1_B.$$

Hence we say a quantum channel is a completely positive trace-preserving map (denoted c.p.t.p. map).

Let $\mathcal{E}$ be a quantum channel as in Definition 5.1.5, then $\mathcal{E}$ has a dual representation in terms of the observable algebras. To see this, we apply the sum and cyclic properties of the trace to the following,

$$\begin{aligned}
\mathbf{Tr}\big(Y\mathcal{E}(X)\big) &= \mathbf{Tr}\left(Y\sum_{i=1}^{k} V_i X V_i^*\right) \\
&= \sum_{i=1}^{k}\big(\mathbf{Tr}(YV_iXV_i^*)\big) \\
&= \sum_{i=1}^{k}\big(\mathbf{Tr}(V_i^*YV_iX)\big) \\
&= \mathbf{Tr}\left(\left(\sum_{i=1}^{k} V_i^*YV_i\right)X\right) \\
&= \mathbf{Tr}\big(E(Y)X\big),
\end{aligned} \tag{5.3}$$

for any $Y \in A$, $X \in S_{\mathcal{B}}$ with

$$E : A \to B$$

such that

$$E(Y) = \sum_{i=1}^{k} V_i^*YV_i,$$

It is clear from Theorem 3.1.24 that E is a a c.p. map. We can now consider E acting on the unit 1_A of A as follows,

$$\begin{aligned}
E(1_A) &= \sum_{i=1}^{k} V_i^* 1_A V_i \\
&= \sum_{i=1}^{k} V_i^* V_i \\
&= 1_B \text{ by Definition 5.1.5.}
\end{aligned} \tag{5.4}$$

From (5.4), we get that E is a unital completely positive map and a dual representation of $\mathcal{E}$ since (5.3) works both ways, i.e., starting with any unital completely positive map we can use (5.3) to get a quantum channel. From

here onwards, we will refer to both E and $\mathcal{E}$ as *channels* since they are dual representations of one another, but the term quantum channel will be used when discussing $\mathcal{E}$.

By the above definitions and our earlier discussion (in this chapter), we now know that a quantum channel is a completely positive map that transmits quantum information from one quantum system to another. This chapter is interested in formulating a distance between any two quantum channels from one specific quantum system to another specified quantum system. We want to cover this in a general sense, hence we will consider composite quantum systems in Sections 5.3, 5.4.

For a general setup, we know from Chapter 3 that B and A are unital C*-algebras, so Section 7.3 formulates a distance between unital completely positive maps (i.e., channels) from one specific composite C*-algebra to another specified composite C*-algebra. In this setting of C*-algebras, we see that Definition 5.1.5 generalises Definition 3.1.20 (see Remark 3.1.23 for a better explanation of this). This implies that states are a special case of channels, hence a distance between channels will also give us a distance between states.

The next section shows a duality between density matrices and quantum channels.

5.2 The Choi-Jamiołkowski duality

Following the setup of the previous section, we fix an orthonormal basis $|1^{\mathcal{B}}\rangle, ..., |r^{\mathcal{B}}\rangle$ for the Hilbert space $\mathcal{H}_{\mathcal{B}}$ of quantum system $\mathcal{B}$ and similarly fix $|1^{\mathcal{A}}\rangle, ..., |q^{\mathcal{A}}\rangle$ for quantum system $\mathcal{A}$. We also represent the observable algebra of $\mathcal{B}$ in terms of the given basis as the matrix algebra $B = M_r$, and similarly use $A = M_q$ for $\mathcal{A}$. Let $\mathcal{E}$ be a quantum channel from B to A.

From [56, Subsection 8.2.3] we can define the maximally entangled state vector of the composite system $\mathcal{B}\mathcal{B}$ (i.e., a system consisting of two copies of $\mathcal{B}$) as

$$|\Omega\rangle = \frac{1}{\sqrt{r}} \sum_{i=1}^{r} |i^{\mathcal{B}}\rangle \otimes |i^{\mathcal{B}}\rangle \in \mathcal{H}_{\mathcal{B}} \otimes \mathcal{H}_{\mathcal{B}}.$$

The above state is represented by the following density matrix:

$$|\Omega\rangle\langle\Omega| = \frac{1}{r}\sum_{i=1}^{r}\sum_{j=1}^{r}|i^{\mathcal{B}}\rangle\langle j^{\mathcal{B}}| \otimes |i^{\mathcal{B}}\rangle\langle j^{\mathcal{B}}|.$$

It is easy to show that $|\Omega\rangle\langle\Omega|$ is a density matrix. $|\Omega\rangle\langle\Omega| \geq 0$ because it is a self-adjoint sum and tensor product of positive operators. Taking note of the fact that the $\mathbf{Tr}(X_1 \otimes X_2) = \mathbf{Tr}(X_1)\mathbf{Tr}(X_2)$ and other properties of the trace we have that

$$\begin{aligned}\mathbf{Tr}\big(|\Omega\rangle\langle\Omega|\big) &= \frac{1}{r}\sum_{i=1}^{r}\sum_{j=1}^{r}\mathbf{Tr}\Big(|i^{\mathcal{B}}\rangle\langle j^{\mathcal{B}}|\Big)^2 \\ &= \frac{1}{r}\sum_{i=1}^{r}\sum_{j=1}^{r}\delta_{ij} \\ &= 1,\end{aligned}$$

where δ_{ij} is the standard *Kronecker delta*. So $|\Omega\rangle\langle\Omega|$ is a denstiy matrix by Definition 4.1.3.

For any quantum channel $\mathcal{E}$, we define the *Choi-matrix* representation of $\mathcal{E}$ in $\mathcal{BA}$ as

$$\kappa_{\mathcal{E}} := \frac{1}{r}\sum_{i=1}^{r}\sum_{j=1}^{r}|i^{\mathcal{B}}\rangle\langle j^{\mathcal{B}}| \otimes \mathcal{E}(|i^{\mathcal{B}}\rangle\langle j^{\mathcal{B}}|). \tag{5.5}$$

It is clear from the arguments about $|\Omega\rangle\langle\Omega|$ being a density matrix and $\mathcal{E}$ being a quantum channel (preserving trace and positivity) that $\kappa_{\mathcal{E}}$ is also a density matrix in $\mathcal{BA}$.

The above result is called the *Choi-Jamiołkowski duality* which states that any quantum channel $\mathcal{E}$ from system $\mathcal{B}$ to $\mathcal{A}$ corresponds to a density matrix $\kappa_{\mathcal{E}}$ in the composite system $\mathcal{BA}$. This implies that one way to compare two quantum channels $\mathcal{E}_1, \mathcal{E}_2$ (i.e., find the distance between them) is to do this by comparing $\kappa_{\mathcal{E}_1}$ and $\kappa_{\mathcal{E}_2}$ in the composite system. Our aim is to formulate a distance between $\mathcal{E}_1$ and $\mathcal{E}_2$ that avoids this relationship.

In our approach, we will instead adapt the DMTL directly to channels, while using the Choi-Jamiołkowski duality to highlight the correspondence of our approach to the special case of states.

For the above to work we need to make sure that every quantum channel has its own density matrix, in other words, mapping a quantum channel to a density matrix is injective (or one to one).

To do this we need to note that for any quantum channel $\mathcal{E}$ with its associated $\kappa_{\mathcal{E}}$ as given in (5.5), we have that

$$\begin{aligned}\kappa_{\mathcal{E}}^{\mathrm{T}_{\mathcal{B}}} &= \frac{1}{r}\sum_{m=1}^{r}\sum_{n=1}^{r}\left(|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|\right)^{\mathrm{T}}\otimes\mathcal{E}(|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|)\\ &= \frac{1}{r}\sum_{m=1}^{r}\sum_{n=1}^{r}\left(|n^{\mathcal{B}}\rangle\langle m^{\mathcal{B}}|\right)\otimes\mathcal{E}(|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|),\end{aligned} \tag{5.6}$$

where $\mathrm{T}_{\mathcal{B}}$ is the partial transposition over $\mathcal{B}$ as shown above. Using (5.6) and noting that we have an orthonormal basis, we get that

$$\begin{aligned}&r\left(\langle j^{\mathcal{B}}|\otimes 1_B\right)\kappa_{\mathcal{E}}^{\mathrm{T}_{\mathcal{B}}}\left(|i^{\mathcal{B}}\rangle\otimes 1_B\right)\\ &= \left(\langle j^{\mathcal{B}}|\otimes 1_B\right)\left(\sum_{m=1}^{r}\sum_{n=1}^{r}|n^{\mathcal{B}}\rangle\langle m^{\mathcal{B}}|\otimes\mathcal{E}(|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|)\right)\left(|i^{\mathcal{B}}\rangle\otimes 1_B\right)\\ &= \sum_{m=1}^{r}\sum_{n=1}^{r}\langle j^{\mathcal{B}}|n^{\mathcal{B}}\rangle\langle m^{\mathcal{B}}|i^{\mathcal{B}}\rangle\,\mathcal{E}(|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|)\\ &= \sum_{m=1}^{r}\sum_{n=1}^{r}\delta_{jn}\delta_{mi}\mathcal{E}(|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|)\\ &= \mathcal{E}(|i^{\mathcal{B}}\rangle\langle j^{\mathcal{B}}|),\end{aligned} \tag{5.7}$$

where 1_B the unit of $B = M_r$. (5.6) and (5.7) imply that we can recover a quantum channel from its Choi-matrix, so any two distinct quantum channels cannot have the same choi-matrix.

Remark 5.2.1. It can be shown that the Choi-matrix $\kappa_{\mathcal{E}}$ of a quantum channel $\mathcal{E}$ from system $\mathcal{B}$ to $\mathcal{A}$ reduces to a maximally mixed state (density matrix) of $\mathcal{B}$. In fact, we can show that the association of a Choi-matrix and its quantum channel is surjective (i.e., any density matrix that reduces to a maximally mixed state of $\mathcal{B}$ defines a quantum channel via (5.7)). This is the reason why this association is called the Choi-Jamiołkowski duality. We are not going to show these results in detail because they don't play any role here.

As mentioned before, our ultimate aim is to side-step the Choi-Jamiołkowski duality, in other words, formulate a W_1 distance that compares quantum channels more directly. Before we get there we need to discuss a few results of the Choi-matrix. (5.3) defines a channel from $E : A \mapsto B$ for any quantum channel $\mathcal{E}$. It is straightforward to show that the action of E can be defined via the Choi-matrix $\kappa_{\mathcal{E}}$. Similar to before, taking the partial transposition over $\mathcal{B}$ in the second order of $|\Omega\rangle\langle\Omega|$ implies that

$$\begin{aligned}\kappa_{\mathcal{E}}^{\mathrm{T}_{\mathcal{A}}} &= \frac{1}{r}\sum_{m=1}^{r}\sum_{n=1}^{r}|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|\otimes\mathcal{E}\left(\left(|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|\right)^{\mathrm{T}}\right)\\ &= \frac{1}{r}\sum_{m=1}^{r}\sum_{n=1}^{r}|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|\otimes\mathcal{E}(|n^{\mathcal{B}}\rangle\langle m^{\mathcal{B}}|).\end{aligned} \tag{5.8}$$

When we consider the following,

$$\begin{aligned}&\mathbf{Tr}_{\mathcal{A}}\left(r\kappa_{\mathcal{E}}^{\mathrm{T}_{\mathcal{A}}}(1_B\otimes Y)\right) =\\ &\qquad\mathbf{Tr}_{\mathcal{A}}\left(\left(\sum_{m=1}^{r}\sum_{n=1}^{r}|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|\otimes\mathcal{E}(|n^{\mathcal{B}}\rangle\langle m^{\mathcal{B}}|)\right)(1_B\otimes Y)\right)\\ &\qquad=\mathbf{Tr}_{\mathcal{A}}\left(\left(\sum_{m=1}^{r}\sum_{n=1}^{r}|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|\otimes\mathcal{E}(|n^{\mathcal{B}}\rangle\langle m^{\mathcal{B}}|)Y\right)\right)\\ &\qquad=\sum_{m=1}^{r}\sum_{n=1}^{r}|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|\,\mathbf{Tr}\left(\mathcal{E}(|n^{\mathcal{B}}\rangle\langle m^{\mathcal{B}}|)Y\right)\\ &\qquad=\sum_{m=1}^{r}\sum_{n=1}^{r}|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|\,\mathbf{Tr}\left(Y\mathcal{E}(|n^{\mathcal{B}}\rangle\langle m^{\mathcal{B}}|)\right)\\ &\qquad=\sum_{m=1}^{r}\sum_{n=1}^{r}|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|\,\mathbf{Tr}\left(E(Y)\,|n^{\mathcal{B}}\rangle\langle m^{\mathcal{B}}|\right)\text{ by (5.3)}\\ &\qquad=\sum_{m=1}^{r}\sum_{n=1}^{r}\left(E(Y)_{mn}\right)|m^{\mathcal{B}}\rangle\langle n^{\mathcal{B}}|\\ &\qquad=E(Y),\end{aligned} \tag{5.9}$$

with $E(Y)_{mn}$ being the m,n entry of $E(Y)$ and $\mathbf{Tr}_{\mathcal{A}}$ denoting the partial trace over the system $\mathcal{A}$. Based on what we have covered so far, it is straightforward to show that $\mathbf{Tr}_{\mathcal{A}}\left(r\kappa_{\mathcal{E}}^{\mathrm{T}_{\mathcal{A}}}\right) = 1_B$. With the above in mind we can consider the set $K(A,B)$ of channels $E : A \to B$, and the set

$$C(\mathcal{B}\mathcal{A}) = \{\delta \in B\otimes S_{\mathcal{A}} : \delta \geq 0 \text{ and } \mathbf{Tr}_{\mathcal{A}}\delta = 1_B\},$$

we have a one-to-one correspondence between $K(A,B)$ and $C(\mathcal{BA})$ given by

$$E(a) = \mathbf{Tr}_{\mathcal{A}}[\delta(1_B \otimes a)], \tag{5.10}$$

for any $a \in A$. Denoting the dual of E by $\mathcal{E} : S_{\mathcal{B}} \to S_{\mathcal{A}}$, (5.9) implies that

$$\delta = r\kappa_{\mathcal{E}}^{\mathrm{T}_{\mathcal{A}}}$$

in this one-to-one correspondence. Because of the usual interpretation of the Choi-Jamiołkowski duality, one might be tempted to rather view the elements of $C(\mathcal{BA})$ as $\delta \in S_{\mathcal{B}} \otimes S_{\mathcal{A}}$ such that δ/r is a density matrix of the composite system $\mathcal{BA}$ reducing to the maximally mixed state of $\mathcal{B}$, but strictly speaking $E(Y)$ as given by (5.10) would then be in $S_{\mathcal{B}}$ instead of B. When $r > 1$, it is therefore in fact conceptually better if we do not view δ/r as a density matrix representing some state. On the other hand, in the special case where $\mathcal{B}$ is a trivial system, i.e., $r = 1$ and $B = \mathbb{C}$, the set $C(\mathcal{BA})$ is precisely all the density matrices of $\mathcal{A}$, and we simply recover the usual representation of expectation values of a state, $E(Y) = \mathbf{Tr}(\delta Y)$, in terms of the density matrix δ or a state of an algebra as given in Definition 3.1.20.

5.3 The W_1 norm

Using this representation of channels as elements of $C(\mathcal{BA})$, we can formulate an extension of the Wasserstein distance of order 1 between states in the DMTL approach, to channels. This is a distance between channels acting from one composite system, $\mathcal{A}$, to another, $\mathcal{B}$. Here we assume that

$$A = A_1 \otimes ... \otimes A_n \quad \text{and} \quad B = B_1 \otimes ... \otimes B_n$$

with $A_j = M_{q_j}$ and $B_j = M_{r_j}$. The latter are simply the observable algebras of systems $\mathcal{A}_j$ and $\mathcal{B}_j$ respectively. One then defines the real vector space

$$\mathcal{O} = \{X \in \mathrm{span}_{\mathbb{R}}\, C(\mathcal{BA}) : \mathbf{Tr}_{\mathcal{A}} X = 0\}, \tag{5.11}$$

where $\mathrm{span}_{\mathbb{R}}$ again refers to finite linear combinations with real coefficients. The goal is to define a certain norm $\|\cdot\|_{W_1}$ on $\mathcal{O}$, which when applied to differences $\delta - \varepsilon$ of elements of $C(\mathcal{BA})$, will in turn define a metric on $C(\mathcal{BA})$. This metric will be the Wasserstein distance of order 1 on $C(\mathcal{BA})$, or equivalently on $K(A,B)$ via the one-to-one correspondence from before, extending the construction in the DMTL approach to channels.

The definition of the norm $\|\cdot\|_{W_1}$ entails extending the idea of neighbouring states in the DMTL approach to channels. We write

$$A_{\widehat{j}} = A_1 \otimes ... \widehat{A}_j ... \otimes A_n \quad \text{and} \quad B_{\widehat{j}} = B_1 \otimes ... \widehat{B}_j ... \otimes B_n,$$

i.e., these are A and B with A_j and B_j respectively left out of the tensor products. Similarly we set

$$A_{\leq j} = A_1 \otimes ... \otimes A_j \quad \text{and} \quad A_{\geq j} = A_j \otimes ... \otimes A_n,$$

and likewise for $B_{\leq j}$ and $B_{\geq j}$. In terms of the following notation (which is chosen to fit in with that of later chapters and sections),

$$\nu_j = \frac{1}{r_j}\mathbf{Tr}$$

where this $\mathbf{Tr}$ is the usual trace on $B_j = M_{r_j}$, we can then reduce a channel $E : A \to B$ to the channel

$$E_{\widehat{j}} = (\mathrm{id}_{B_{\leq j-1}} \otimes \nu_j \otimes \mathrm{id}_{B_{\geq j+1}}) \circ E|_{A_{\widehat{j}}}$$

from $A_{\widehat{j}}$ to $B_{\widehat{j}}$, where id denotes the identity map on the indicated algebra, and with $E|_{A_{\widehat{j}}}$ defined via

$$E|_{A_{\widehat{j}}}(a_1 \otimes ... \widehat{a}_j ... \otimes a_n) = E(a_1 \otimes ... \otimes a_{j-1} \otimes 1_{A_j} \otimes a_{j+1} \otimes ... \otimes a_n).$$

I.e., we restrict E to $A_{\widehat{j}}$, and evaluate the "partial expectation" of the result over B_j. Note that via the one-to-one correspondence given by (5.10), this reduction of E is equivalent to the reduction

$$\delta_{\widehat{j}} = (\mathrm{id}_{B_{\leq j-1}} \otimes \nu_j \otimes \mathrm{id}_{B_{\geq j+1}}) \otimes \mathbf{Tr}_{\mathcal{A}_j}$$

of the corresponding $\delta \in C(\mathcal{BA})$, where $\mathbf{Tr}_{\mathcal{A}_j}$ denotes the partial trace on $S_{\mathcal{A}} = M_{q_1} \otimes ... \otimes M_{q_n}$ over $S_{\mathcal{A}_j} = M_{q_j}$ with δ corresponding to E via (5.10). In terms of this notation, we view $\delta, \varepsilon \in C(\mathcal{BA})$ as representing neighbouring channels when $\delta_{\widehat{j}} = \varepsilon_{\widehat{j}}$ for some j. We define

$$\mathcal{N}_j = \{\delta - \varepsilon : \delta, \varepsilon \in C(\mathcal{BA}) \text{ with } \delta_{\widehat{j}} = \varepsilon_{\widehat{j}}\} \tag{5.12}$$

and

$$\mathcal{N} = \bigcup_{j=1}^{n} \mathcal{N}_j.$$

We then define $\|\cdot\|_{W_1}$ as the norm on $\mathcal{O}$ which has the convex hull

$$\mathcal{C} = \operatorname{conv} \mathcal{N}$$

of $\mathcal{N}$ as its unit ball. I.e.,

$$\|X\|_{W_1} = \inf\{t \geq 0 : X \in t\mathcal{C}\} \tag{5.13}$$

where $t\mathcal{C} = \{tX : X \in \mathcal{C}\}$ for any real number t. The norm $\|\cdot\|_{W_1}$ will be referred to as the W_1 *norm*. To prove that this is a norm on $\mathcal{O}$ of course requires some work, which will be done in a more general context in the sequel.

5.4 W_1 between channels

Given this norm, we can define a metric W_1 on $K(A, B)$ via

$$W_1(E_\delta, E_\varepsilon) = \|\delta - \varepsilon\|_{W_1}$$

with E_δ denoting the channel E corresponding to $\delta \in C(\mathcal{BA})$ in (5.10). This metric W_1 is the generalization of the Wasserstein distance of order 1 between states to the case of channels. We consequently refer to it as the *Wasserstein distance of order 1* on $K(A, B)$.

The rough intuition behind this metric follows from δ and ε in the definition of $\mathcal{N}_j$ above being neighbouring channels. This condition tells us that ε and δ coincide when reduced to $A_{\widehat{j}} \to B_{\widehat{j}}$ for some j, i.e., with one system, $\mathcal{A}_j$ and $\mathcal{B}_j$ respectively, removed from each of the composite systems $\mathcal{A}$ and $\mathcal{B}$. In this way the "local differences" between two channels E_δ and E_ε are picked up by W_1, where "local" here is simply in relation to the systems $\mathcal{A}_1, ..., \mathcal{A}_n$ and $\mathcal{B}_1, ..., \mathcal{B}_n$ composing $\mathcal{A}$ and $\mathcal{B}$. A typical case is the dynamics of an open composite system $\mathcal{A}$, where we take $\mathcal{B}_j = \mathcal{A}_j$ for all j. We then expect W_1 to naturally take into account the differences between two dynamical processes in the individual systems $\mathcal{A}_j$.

A basic property of W_1 is additivity (see Chapter 8 for a general discussion) with respect to tensor products. If we partition the set $[n] = \{1, ..., n\}$ into m non-empty and sequential parts, i.e.,

$$\begin{aligned} P(1) &= \{1, ..., n_1\} \\ P(2) &= \{n_1 + 1, ..., n_2\} \\ &\vdots \\ P(m) &= \{n_{m-1} + 1, ..., n\}, \end{aligned}$$

then we can consider the subsystems of $\mathcal{A}$ and $\mathcal{B}$ with observable algebras

$$A_{P(k)} = \bigotimes_{j\in P(k)} A_j \quad \text{and} \quad B_{P(k)} = \bigotimes_{j\in P(k)} B_j$$

respectively. For any channels $D_k, E_k : A_{P(k)} \to B_{P(k)}$ we then have

$$W_1(D_1 \otimes ... \otimes D_m, E_1 \otimes ... \otimes E_m) = \sum_{k=1}^{m} W_1(D_k, E_k),$$

with a resulting stability property when $D_j = E_j$ for some of the j's. This can be refined by dropping the assumption that the partition is sequential, but the form above is for the moment notationally clearer. Keep in mind that W_1 on $K(A_{P(k)}, B_{P(k)})$ is of course defined by the same procedure as for $K(A, B)$.

When $\mathcal{B}_j = \mathbb{C}$, we recover the case of states on $\mathcal{A}$, and indeed our W_1 above then specialises to the Wasserstein distance of order 1 in the DMTL approach, with $\delta, \varepsilon \in C(\mathcal{B}\mathcal{A})$ becoming density matrices ρ and σ of the composite system $\mathcal{A}$. They specifically considered the case $q_1 = ... = q_n = d$.

In the subsequent general theory, the formulation will be in the Heisenberg picture $E : A \to B$ from the outset. The Choi-Jamiołkowski duality will also be side-stepped, with the formulation expressed directly in terms of the channels themselves. In connection to this, note that the basic condition $\mathbf{Tr}_{\mathcal{A}} X = 0$ in (5.11) can equivalently be expressed as $\lambda(1_A) = 0$, with $\lambda : A \to B$ defined by

$$\lambda(a) = \mathbf{Tr}_{\mathcal{A}}[X(1_B \otimes a)]$$

in terms of the given X. The latter extends the formula in (5.10). The condition $\lambda(1_A) = 0$ relates to the unitality of channels, namely $E(1_A) = 1_B$, as will be seen in abstract form in Chapters 6 and 7.

The reader may now turn directly to Section 7.3 and 8.4 to see the general C*-algebraic version of this section. However, the proofs that the Wasserstein distance of order 1 treated there is indeed a metric and satisfies additivity, rely on Chapter 6 and 7 as well as the rest of Chapter 8.

Chapter 6

Pointed spaces and a composite setup

This chapter is motivated by the fact that channels are unital maps, i.e., we have that channels map the identity element of one algebra to the identity element of another algebra. This chapter is aimed at showing the role that this property of channels plays when introduced in a more abstract form between vector spaces, when setting up W_1 seminorms. In the next chapter, these in turn lead to W_1 norms and ultimately to W_1 distances between channels.

6.1 Pointed spaces and maps

The absorbing property as defined in Definition 2.2.1 is needed for the gauge function $\|\cdot\|_{\mathcal{C}}$ on $\mathcal{X}$ in Proposition 2.3.7 to be well defined and finite at every point of $\mathcal{X}$, and it is in that sense the most basic property of a unit ball. In this section we show how it arises under general assumptions for the set to be used as a unit ball for the W_1 norm. We are ultimately interested in the situation where $\mathcal{X}$ is a space of maps between two composite systems. In this section $\mathcal{X}$ is more generally taken as a vector space $\mathcal{O}$ of linear maps between two vector spaces, each with a distinguished point. The maps will be required to map the one distinguished point to the other, as an abstraction of unitality (this refers to unital maps). Actual unitality, in the case of unital algebras and C*-algebras as the pointed spaces, will be treated in the next chapter.

The following definition is motivated by [25, p. 118], but in this case we present this definition for vector spaces.

Definition 6.1.1 (Pointed space). A *pointed space* is a pair (A, u_A) consisting of a vector space A over some field $\mathbb{F}$ and a distinguished point $u_A \in A$ such that $u_A \neq 0$. For simplicity of notation and unless stated otherwise, a pointed space (A, u_A) will be denoted as A, with u_A assumed as the notation for the distinguished point.

One should keep in mind that eventually (Section 7.3) the pointed spaces will be taken to be unital C*-algebras with $\mathbb{F} = \mathbb{C}$, and the units serving as the distinguished points. These C*-algebras will generalise the matrix algebras from Chapter 4 and 5.

In the remainder, all pointed spaces involved are assumed to be over the same field of scalars $\mathbb{F}$ (which is either real or complex). However, in either case certain constructions will involve the span of a subset over real scalars, leading to a real vector space, and such spans will again be indicated by $\text{span}_{\mathbb{R}}$.

We want to use a special case of Definition 3.1.12, where a specified point of some vector space is mapped to a specific point of another space. Hence we give the following definition.

Definition 6.1.2 (Pointed maps). If A and B are pointed spaces, we define a *pointed map* as $\eta \in L(A, B)$ such that $\eta(u_A) = u_B$ for $u_A \in A$ and $u_B \in B$ the distinguised points of A and B respectively. We also define

$$L_u(A, B) = \{\eta \in L(A, B) : \eta(u_A) = u_B\},$$

as the set of *pointed maps* from A to B.

Definition 6.1.2 is of course an abstract version of unital maps in the case where A and B are unital algebras.

The following result will shortly be used in tandem with Lemma 4.2.3 to prove the absorbing property of certain sets.

Lemma 6.1.3. *Let A and B be pointed spaces and consider any subset $\mathcal{L}$ of $L_u(A, B)$. Set*

$$\mathcal{O} := \{\lambda \in \text{span}_{\mathbb{R}} \mathcal{L} : \lambda(u_A) = 0\} \quad \textit{and} \quad \mathcal{V} := \{\eta - \theta : \eta, \theta \in \mathcal{L}\}.$$

Then it follows that $\mathcal{O} = \text{span}_{\mathbb{R}} \mathcal{V}$.

Proof. Clearly $\text{span}_{\mathbb{R}}\,\mathcal{V} \subset \mathcal{O}$. Conversely, take any non-zero $\lambda \in \mathcal{O}$ and write $\lambda_1 := \lambda$. Since $\lambda_1(u_A) = 0$ and $u_B \neq 0$, the coefficients in λ_1's expansion as a linear combination of $\mathcal{L}$'s elements, can not all be strictly positive or all strictly negative. Without loss of generality we can therefore rewrite the expansion in the form

$$\lambda = r_1^{(1)}\eta_1^{(1)} + ... + r_{k^{(1)}}^{(1)}\eta_{k^{(1)}}^{(1)} - s_1^{(1)}\theta_1^{(1)} - ... - s_{k^{(1)}}^{(1)}\theta_{k^{(1)}}^{(1)}$$

where $\eta_j^{(1)}, \theta_j^{(1)} \in \mathcal{L}$ and $r_j^{(1)}, s_j^{(1)} > 0$ for some (finite) $k^{(1)}$, where the superscript (1) is merely an index, and with the proviso that $\lambda(u_A) = 0$. Note that $k^{(1)}$ is simply the biggest of the number of positive coefficients and the number of negative coefficients in λ_1's initial expansion, while the coefficients in smaller number are split to increase their number to $k^{(1)}$. Writing $\lambda_1' = r_1^{(1)}(\eta_1^{(1)} - \theta_1^{(1)}) + ... + r_{k^{(1)}}^{(1)}(\eta_{k^{(1)}}^{(1)} - \theta_{k^{(1)}}^{(1)}) \in \text{span}_{\mathbb{R}}\,\mathcal{V}$ and $\lambda_2 = (r_1^{(1)} - s_1^{(1)})\theta_1^{(1)} + ... + (r_{k^{(1)}}^{(1)} - s_{k^{(1)}}^{(1)})\theta_{k^{(1)}}^{(1)}$, we have $\lambda_1 = \lambda_1' + \lambda_2$, implying that $\lambda_2 \in \mathcal{O}$. If $\lambda_2 \neq 0$, then repeat this procedure for λ_2 instead of λ_1, noting that we can now analogously write

$$\lambda_2 = r_1^{(2)}\eta_1^{(2)} + ... + r_{k^{(2)}}^{(2)}\eta_{k^{(2)}}^{(2)} - s_1^{(2)}\theta_1^{(2)} - ... - s_{k^{(2)}}^{(2)}\theta_{k^{(2)}}^{(2)},$$

but with $k^{(2)} < k^{(1)}$. This delivers $\lambda_2' \in \text{span}_{\mathbb{R}}\,\mathcal{V}$ and $\lambda_3 \in \mathcal{O}$. If $\lambda_3 \neq 0$, then repeat for λ_3, etc. Since $k^{(j+1)} < k^{(j)}$, this process must stop to deliver $\lambda = \lambda_1' + ... + \lambda_m' \in \text{span}_{\mathbb{R}}\,\mathcal{V}$. □

6.2 The composite setup

In this section we set up a framework which will serve as an abstraction of composite systems in terms of pointed spaces. We also define the set for which the absorbing property will be proven in the next section, leading to it being a semi unit ball giving us a W_1 seminorm.

Consider pointed spaces $A_1, ..., A_n$ and $B_1, ..., B_n$. Let

$$A = A_1 \odot ... \odot A_n$$

be the algebraic tensor product of the vector spaces $A_1, ..., A_n$, which is itself a pointed space with distinguished point

$$u_A := u_{A_1} \otimes ... \otimes u_{A_n}.$$

However, as we want to allow for completions of A, in particular in the C*-algebraic framework of Section 7.3 (where $A_1, ..., A_n$ will be unital C*-algebras), we need to allow for completions of the tensor product of $B_1, ..., B_n$

from the outset. To emphasise this, we write

$$B = B_1 \otimes ... \otimes B_n,$$

which may be the algebraic tensor product, or some completion of it with respect to a specified norm. In all cases B is a pointed space with $u_B = u_{B_1} \otimes ... \otimes u_{B_n}$. The particular tensor product B remains fixed throughout this section, however. The point of this is that in the theory developed here, linear maps $A \to B$ are then not constrained to have their image contained in an uncompleted tensor product.

Along the lines of Chapter 4 and 5, we define

$$A_{\widehat{j}} = A_1 \odot ... \widehat{A}_j ... \odot A_n \quad \text{and} \quad B_{\widehat{j}} = B_1 \otimes ... \widehat{B}_j ... \otimes B_n,$$

$$A_{\leq j} = A_1 \odot ... \odot A_j \quad \text{and} \quad A_{\geq j} = A_j \odot ... \odot A_n,$$

and

$$B_{\leq j} = B_1 \otimes ... \otimes B_j \quad \text{and} \quad B_{\geq j} = B_j \otimes ... \otimes B_n,$$

for $j = 1, ..., n$, using the same completion (if relevant) for tensor products of the B_j's as for B.

Let ν_j be a linear functional on B_j such that

$$\nu_j(u_{B_j}) = 1$$

for $j = 1, ..., n$. These functionals will act as reference functionals relative to which linear maps $A \to B$ will be reduced to $A_{\widehat{j}} \to B_{\widehat{j}}$. Later on (Section 7.3), in the case of unital C*-algebras, the ν_j's will be taken as states, generalizing the normalised traces used in Section 5.1.

If B is indeed a completion with respect to some norm, rather than just an algebraic tensor product, we also assume that algebraic tensor products of the ν_j's with themselves as well as with the identity maps $\mathrm{id}_{B_{\leq j-1}}$ and $\mathrm{id}_{B_{\geq j+1}}$ are continuous with respect to this norm and therefore uniquely extendible to the completed tensor products. The relevant reductions of a linear map $\eta \in L(A, B)$ from A to B can then be defined as

$$\eta_{\widehat{j}} = (\mathrm{id}_{B_{\leq j-1}} \otimes \nu_j \otimes \mathrm{id}_{B_{\geq j+1}}) \circ \eta|_{A_{\widehat{j}}} : A_{\widehat{j}} \to B_{\widehat{j}} \tag{6.1}$$

$$\eta_{\leq j} = (\mathrm{id}_{B_{\leq j}} \otimes \nu_{>j}) \circ \eta|_{A_{\leq j}} : A_{\leq j} \to B_{\leq j} \tag{6.2}$$

$$\eta_{\geq j} = (\nu_{<j} \otimes \mathrm{id}_{B_{\geq j}}) \circ \eta|_{A_{\geq j}} : A_{\geq j} \to B_{\geq j} \tag{6.3}$$

for $j = 1, ..., n$, with $\nu_{<j} = \nu_1 \otimes ... \otimes \nu_{j-1}$ (empty when $j = 1$) and $\nu_{>j} = \nu_{j+1} \otimes ... \otimes \nu_n$ (empty when $j = n$), where all the indicated tensor product maps are algebraic when the tensor product B is, or (uniquely) continuously extended to B as assumed possible above, when B is completed. Here the restriction $\eta|_{A_{\widehat{j}}} : A_{\widehat{j}} \to B$ is defined via

$$\eta|_{A_{\widehat{j}}}(a_1 \otimes ...\widehat{a}_j... \otimes a_n) = \eta(a_1 \otimes ... \otimes a_{j-1} \otimes u_{A_j} \otimes a_{j+1} \otimes ... \otimes a_n)$$

for $a_1 \in A_1, ..., a_n \in A_n$ (using the universal property), replacing a_j by u_{A_j}, and analogously for $\eta|_{A_{\leq j}} : A_{\leq j} \to B$ and $\eta|_{A_{\geq j}} : A_{\geq j} \to B$, where the notation $\widehat{a}_j$ indicates the absence of a_j in the elementary tensor.

Note in particular that when $\eta \in L_u(A, B)$, it follows directly from these definitions that

$$\eta_{\widehat{j}} \in L_u(A_{\widehat{j}}, B_{\widehat{j}}), \quad \eta_{\leq j} \in L_u(A_{\leq j}, B_{\leq j}) \quad \text{and} \quad \eta_{\geq j} \in L_u(A_{\geq j}, B_{\geq j}),$$

where the property $\nu_j(u_{B_j}) = 1$ has been used. Note that for $\eta, \theta \in L_u(A, B)$ this implies that

$$\theta_{\leq j} \odot \eta_{\geq j+1} \in L_u(A, B) \tag{6.4}$$

for $j = 1, ..., n-1$, which will implicitly play a role in Lemma 6.3.1 below.

We also fix any subset

$$\mathcal{L} \subset L_u(A, B)$$

and let

$$\mathcal{O} := \{\lambda \in \operatorname{span}_{\mathbb{R}} \mathcal{L} : \lambda(u_A) = 0\} \tag{6.5}$$

as in Lemma 6.1.3. Our ultimate goal, given sufficient additional structure and assumptions, is to define a metric, namely the Wasserstein distance of order 1, on $\mathcal{L}$. This will be done by first obtaining a seminorm, and in the next section under further assumptions a norm, on $\mathcal{O}$.

Let

$$\mathcal{N}_j := \{\eta - \theta : \eta, \theta \in \mathcal{L} \text{ such that } \eta_{\widehat{j}} = \theta_{\widehat{j}}\}, \tag{6.6}$$

generalizing (5.12), though now in the Heisenberg picture, and set

$$\mathcal{N} := \bigcup_{j=1}^{n} \mathcal{N}_j \quad \text{and} \quad \mathcal{C} := \operatorname{conv} \mathcal{N}, \tag{6.7}$$

where $\mathcal{C}$ is an abstract version of the set we ultimately want to use as a unit ball defining the W_1 norm. In this chapter we settle for a seminorm.

6.3 Absorption in the composite setup and W_1 seminorms

In terms of the setup of the previous section, we now show that $\mathcal{C}$ is absorbing for $\mathcal{O}$. The abstract assumption (6.8) below, is made in lieu of complete positivity. Keep in mind that because of (6.4), $\theta_{\leq j} \odot \eta_{\geq j+1}$ in (6.8) is already a pointed map (which is an abstraction of a unital map). The main technical step is the following lemma.

Lemma 6.3.1. *Assume that*

$$\theta_{\leq j} \odot \eta_{\geq j+1} \in \mathcal{L} \tag{6.8}$$

for all $\eta, \theta \in \mathcal{L}$ and $j = 1, ..., n-1$. For any $\eta, \theta \in \mathcal{L}$ it then follows that $\eta - \theta = \lambda_1 + ... + \lambda_n$ for some $\lambda_j \in \mathcal{N}_j$. In particular,

$$\mathcal{O} = \operatorname{span}_{\mathbb{R}} \mathcal{N}.$$

Proof. In line with Lemma 6.1.3, we set $\mathcal{V} := \{\eta - \theta : \eta, \theta \in \mathcal{L}\}$. The case $n = 1$ indeed follows immediately from Lemma 6.1.3, as we then have $\mathcal{N} = \mathcal{V}$. We can therefore assume $n > 1$. For any $\eta, \theta \in \mathcal{L}$, set

$$\begin{aligned}
\lambda_1 &= \eta - \theta_{\leq 1} \odot \eta_{\geq 2} \\
\lambda_2 &= \theta_{\leq 1} \odot \eta_{\geq 2} - \theta_{\leq 2} \odot \eta_{\geq 3} \\
&\vdots \\
\lambda_{n-1} &= \theta_{\leq n-2} \odot \eta_{\geq n-1} - \theta_{\leq n-1} \odot \eta_{\geq n} \\
\lambda_n &= \theta_{\leq n-1} \odot \eta_{\geq n} - \theta.
\end{aligned}$$

Then $\eta - \theta = \lambda_1 + ... + \lambda_n$. If indeed $\lambda_j \in \mathcal{N}_j$, it follows that $\eta - \theta \in \operatorname{span}_{\mathbb{R}} \mathcal{N}$, hence $\mathcal{V} \subset \operatorname{span}_{\mathbb{R}} \mathcal{N}$. From $\mathcal{N} \subset \mathcal{V}$ and Lemma 6.1.3 we can conclude that $\mathcal{O} = \operatorname{span}_{\mathbb{R}} \mathcal{N}$.

It remains to show that $\lambda_j \in \mathcal{N}_j$. Because of (6.8), we simply have to check the equality of the reductions as required in (6.6), i.e., that $(\lambda_j)_{\widehat{j}} = 0$ for $j = 1, ..., n$. To handle all cases at once, set $A_{\leq 0} = \mathbb{C}$, $A_{\geq n+1} = \mathbb{C}$, $\theta_{\leq 0} = \mathrm{id}_{\mathbb{C}}$ and $\eta_{\geq n+1} = \mathrm{id}_{\mathbb{C}}$, making $\eta = \theta_{\leq 0} \odot \eta_{\geq 1}$ and $\theta = \theta_{\leq n} \odot \eta_{\geq n+1}$. For $j = 1, ..., n$, and any $a_{<j} \in A_{\leq j-1}$ and $a_{>j} \in A_{\geq j+1}$, one has the following direct calculation:

$$\begin{aligned}
&\left(\theta_{\leq j-1} \odot \eta_{\geq j}\right)_{\widehat{j}} (a_{<j} \otimes a_{>j}) \\
&= (\mathrm{id}_{B_{\leq j-1}} \otimes \nu_j \otimes \mathrm{id}_{B_{\geq j+1}}) \circ \left(\theta_{\leq j-1} \odot \eta_{\geq j}\right) (a_{<j} \otimes u_{A_j} \otimes a_{>j}) \\
&= \theta_{\leq j-1}(a_{<j}) \otimes \Big((\nu_j \otimes \mathrm{id}_{B_{\geq j+1}}) \circ \eta_{\geq j}(u_{A_j} \otimes a_{>j}) \Big)
\end{aligned}$$

and

$$\begin{aligned}&(\nu_j \otimes \mathrm{id}_{B_{\geq j+1}}) \circ \eta_{\geq j}(u_{A_j} \otimes (\cdot)) \\ &= (\nu_j \otimes \mathrm{id}_{B_{\geq j+1}}) \circ (\nu_{<j} \otimes \mathrm{id}_{B_{\geq j}}) \circ \eta|_{A_{\geq j}}(u_{A_j} \otimes (\cdot)) \\ &= (\nu_{<j+1} \otimes \mathrm{id}_{B_{\geq j+1}}) \circ \eta|_{A_{\geq j+1}} \\ &= \eta_{\geq j+1},\end{aligned}$$

hence $\left(\theta_{\leq j-1} \odot \eta_{\geq j}\right)_{\widehat{j}} = \theta_{\leq j-1} \odot \eta_{\geq j+1}$.

Similarly $\left(\theta_{\leq j} \odot \eta_{\geq j+1}\right)_{\widehat{j}} = \theta_{\leq j-1} \odot \eta_{\geq j+1}$, thus $\left(\theta_{\leq j-1} \odot \eta_{\geq j}\right)_{\widehat{j}} = \left(\theta_{\leq j} \odot \eta_{\geq j+1}\right)_{\widehat{j}}$, as required. □

Using this lemma we can now show that $\mathcal{C}$ is indeed absorbing.

Proposition 6.3.2. *Assuming (6.8), the set $\mathcal{C}$ in (6.7) is a semi unit ball (Definition 2.2.5) for $\mathcal{O}$ given by (6.5).*

Proof. From (6.6) and (6.7) it is clear that $-\mathcal{N} = \mathcal{N}$, i.e., $\mathcal{N}$ is symmetric, hence so is $\mathcal{C}$, which is also convex by definition. In addition, because of $\mathcal{N}$'s symmetry, Lemmas 6.3.1 and 4.2.3 imply that $\mathcal{C}$ is absorbing for $\mathcal{O}$. □

Because of Proposition 2.3.7, this is sufficient to deliver a seminorm. We summarise this as follows.

Definition 6.3.3. The structure set up in Section 6.2, with $\mathcal{L}$ assumed to satisfy (6.8), is called a *pointed* (W_1, n) *structure* or *pointed* W_1 *structure* (if n is clear from context), and is denoted by the shorthand $(A, B, \nu, \mathcal{L})$, where $\nu := (\nu_1, ...\nu_n)$. The rest of the notation in Section 6.2 is then implied. For clarity the space $\mathcal{O}$ and set $\mathcal{C}$ in (6.5) and (6.7) can respectively be denoted by

$$\mathcal{O}_{\mathcal{L}} \quad \text{and} \quad \mathcal{C}_{\mathcal{L}}$$

in this context.

Corollary 6.3.4. *Given a pointed (W_1, n) structure $(A, B, \nu, \mathcal{L})$, the function $\|\cdot\|_{W_1}$ on $\mathcal{O}_{\mathcal{L}}$ defined by*

$$\|\lambda\|_{W_1} = \inf\{t \geq 0 : \lambda \in t\mathcal{C}_{\mathcal{L}}\},$$

for every $\lambda \in \mathcal{O}_{\mathcal{L}}$, is a seminorm referred to as the W_1 seminorm *associated to $(A, B, \nu, \mathcal{L})$. Consequently, the function $W_1 : \mathcal{L} \times \mathcal{L} \to \mathbb{R}$ defined by*

$$W_1(\eta, \theta) = \|\eta - \theta\|_{W_1},$$

is a pseudometric on $\mathcal{L}$.

Although W_1 is at this stage only a pseudometric, which means we may have $W_1(\eta, \theta) = 0$ for $\eta \neq \theta$, it will nevertheless be called the *Wasserstein distance of order 1* associated to $(A, B, \nu, \mathcal{L})$. We still need ray-wise boundedness to achieve a norm and make W_1 a metric, which is what we turn to next.

Chapter 7

The W_1 norm and unital algebras

In this chapter we achieve a norm and consequently obtain a metric serving as the Wasserstein metric order 1 on the set of unital maps between any two specific unital algebras.

7.1 Ray-wise boundedness

To obtain a norm from the W_1 seminorm in Corollary 6.3.4, we need ray-wise boundedness. In order to achieve this, we make use of a more specialised algebraic framework as well as assumptions complementary to those made in the previous chapter. We are going to work in the context of unital algebras. As in the previous chapter, positivity plays no role here or in Section 7.2, though we again make the abstract assumption (6.8) which will be implied by the complete positivity of channels in the next section. On the other hand, the unitality of maps will be used. For the moment we step away from the tensor product setup of the previous section, and consider a simple algebraic setting. The core structure (6.5) will remain in place, however. This allows us to obtain the remaining building block required by Proposition 2.3.7 in the next result. We return to the tensor product setup in the next section.

Note that in the proposition below, all elements of $\mathcal{L}$ are unital maps.

Proposition 7.1.1. *Let A and B be any unital algebras (both of them real, or both of them complex), with their units 1_A and 1_B respectively serving as the distinguished points making A and B pointed spaces. Fix any subset $\mathcal{L}$ of*

$L_u(A,B)$ and set

$$\mathcal{O} := \{\lambda \in \operatorname{span}_{\mathbb{R}} \mathcal{L} : \lambda(1_A) = 0\} \quad \textit{and} \quad \mathcal{V} := \{\eta - \theta : \eta, \theta \in \mathcal{L}\}.$$

Consider any subset $\mathcal{N}$ of $\mathcal{V}$ and set $\mathcal{C} := \operatorname{conv} \mathcal{N}$. Then $\mathcal{C}$ is ray-wise bounded in $\mathcal{O}$.

Proof. In order to decide whether an element of $\mathcal{O}$ lies outside $\mathcal{C}$, it is going to be convenient to attach a quantitative measure to the element, which when too large, will imply that the element is not in $\mathcal{O}$. To attain this, we follow a construction inspired by the representation theory of C*-algebras (see Chapter 3).

For every $\eta \in \mathcal{L}$, define a bilinear map $\langle \cdot, \cdot \rangle_\eta : A \times A \to B$ by

$$\langle x, y \rangle_\eta = \eta(xy)$$

for all $x, y \in A$. We write G_η for A equipped with this bilinear map (G_η can be referred to as a "bilinear space", which is not a standard term). Define $\pi_\eta : A \to L(G_\eta)$, with $L(G_\eta)$ the space of linear maps from G_η to itself, through

$$\pi_\eta(a)x = ax$$

for all $a \in A$ and $x \in G_\eta$, with ax simply being the product in A (It's clear that π_η is a homomorphism). Then

$$\eta(a) = \left\langle 1_A, \pi_\eta(a) 1_A \right\rangle_\eta$$

in analogy to a cyclic representation obtained from the GNS construction, where 1_A serves as the "cyclic vector" (indeed, $\pi_\eta(A)1_A = G_\eta$). We proceed to consider the direct sum (similar to the difinition of a universal representation)

$$(G, \langle \cdot, \cdot \rangle, \pi) = \bigoplus_{\eta \in \mathcal{L}} (G_\eta, \langle \cdot, \cdot \rangle_\eta, \pi_\eta).$$

I.e., every $x \in G$ is of the form $x = (x_\eta)_{\eta \in \mathcal{L}}$ with $x_\eta \in G_\eta$ and $\{\eta \in \mathcal{L} : x_\eta \neq 0\}$ a finite set. Furthermore, $\pi : A \to L(G)$ is defined by $\pi(a)x = (\pi_\eta(a)x_\eta)_{\eta \in \mathcal{L}}$ for all such x and every $a \in A$. Lastly, $\langle x, y \rangle := \sum_{\eta \in \mathcal{L}} \langle x_\eta, y_\eta \rangle_\eta$ for all $x = (x_\eta)_{\eta \in \mathcal{L}}, y = (y_\eta)_{\eta \in \mathcal{L}} \in G$ defines a bilinear map (additional details of the above are given in Subsection 3.2.1).

For any $\eta \in \mathcal{L}$, define $\hat{\eta} \in G$ by

$$\hat{\eta}_\theta = \begin{cases} 1_A & \text{for } \theta = \eta \\ 0 & \text{for } \theta \neq \eta \end{cases}$$

for all $\theta \in \mathcal{L}$. Note that $\langle \hat{\eta}, \pi(a)\hat{\eta} \rangle = \eta(a)$ for all $\eta \in \mathcal{L}$ and $a \in A$.

Consider any $\lambda \in \mathcal{O}$, which using Lemma 6.1.3, we can write as

$$\lambda = \sum_{j=1}^{l} r_j(\eta_j - \theta_j)$$

where $r_j \geq 0$ and $\eta_j, \theta_j \in \mathcal{L}$. Then

$$\lambda = \sum_{j=1}^{l} r_j \left(\langle \hat{\eta}_j, \pi(\cdot)\hat{\eta}_j \rangle - \left\langle \hat{\theta}_j, \pi(\cdot)\hat{\theta}_j \right\rangle \right).$$

We can use this to lift $\lambda : A \to B$ to a linear map $\bar{\lambda} : L(G) \to B$ defined by

$$\bar{\lambda}(T) = \sum_{j=1}^{l} r_j \left(\langle \hat{\eta}_j, T\hat{\eta}_j \rangle - \left\langle \hat{\theta}_j, T\hat{\theta}_j \right\rangle \right)$$

for all $T \in L(G)$. This lifting may not be unique (it may depend on the choice of η_j's and θ_j's), but for our purposes any such lifting will do. In particular, any $\gamma \in \mathcal{C}$ can be lifted to $\bar{\gamma} : L(G) \to B$ in the form

$$\bar{\gamma}(T) = \sum_{j=1}^{m} p_j \left(\langle \hat{\alpha}_j, T\hat{\alpha}_j \rangle - \left\langle \hat{\beta}_j, T\hat{\beta}_j \right\rangle \right)$$

for some $\alpha_j, \beta_j \in \mathcal{L}$, with $p_j \geq 0$ and $p_1 + ... + p_m = 1$.

According to the Hahn-Banach theorem there is a linear functional f on B such that $f(1_B) \neq 0$ (in the case where $1_B = 0$ and therefore $B = \{0\}$, Proposition 7.1.1 is trivial). Normalise it to obtain

$$\nu = \frac{1}{f(1_B)} f.$$

For any $x, y \in G_\zeta$ we use this to define $x \bowtie y : G_\zeta \to G_\zeta$ by $(x \bowtie y)z = x\nu(\langle y, z \rangle_\zeta)$ for all $z \in G_\zeta$, where the notation $x \bowtie y$ is inspired by the Dirac notation $|x\rangle \langle y|$. For $x, y \in G$ this in turn allows us to define

$$x \bowtie_\oplus y = \bigoplus_{\zeta \in \mathcal{L}} x_\zeta \bowtie y_\zeta \in L(G),$$

i.e., $(x \bowtie_\oplus y)z = (x_\zeta \nu(\langle y_\zeta, z_\zeta \rangle_\zeta))_{\zeta \in \mathcal{L}}$ for $z \in G$. For all $\eta, \theta \in \mathcal{L}$ we then have

$$\left\langle \hat{\theta}, (\hat{\eta} \bowtie_{\oplus} \hat{\eta})\hat{\theta} \right\rangle = \sum_{\zeta \in \mathcal{L}} \left\langle \hat{\theta}_\zeta, \hat{\eta}_\zeta \nu \left(\left\langle \hat{\eta}_\zeta, \hat{\theta}_\zeta \right\rangle_\zeta \right) \right\rangle_\zeta$$
$$= \begin{cases} 1_B & \text{for } \theta = \eta \\ 0 & \text{for } \theta \neq \eta. \end{cases}$$

Thus, in terms of $\bar{\gamma}$ above,

$$\bar{\gamma}(\hat{\eta} \bowtie_{\oplus} \hat{\eta}) = \sum_{j=1}^{m} p_j b_j$$

where $b_j \in \{-1_B, 0, 1_B\}$ for all j, and consequently

$$\left| \nu \left(\bar{\gamma}(\hat{\eta} \bowtie_{\oplus} \hat{\eta}) \right) \right| \leq 1 \tag{7.1}$$

for all $\eta \in \mathcal{L}$. This is therefore a condition satisfied by all $\gamma \in \mathcal{C}$, for any lifting $\bar{\gamma}$ of the form above.

On the other hand, for any non-zero $\lambda \in \mathcal{O}$ and a lifting $\bar{\lambda}$ as above, we see that $\overline{s\lambda} := s\bar{\lambda}$ correspondingly lifts $s\lambda$ for any $s > 0$, using sr_j instead of r_j. To simplify the notation in the remainder of the proof, rewrite $\bar{\lambda}$ above as

$$\bar{\lambda}(T) = \sum_{i=1}^{k} q_i \left\langle \hat{\zeta}_i, T\hat{\zeta}_i \right\rangle,$$

where $k = 2l$, $\zeta_j = \eta_j$, $\zeta_{j+k} = \theta_j$, $q_j = r_j$ and $q_{j+k} = -r_j$ for $j = 1, ..., k$. In this form we may as well assume that $\zeta_i \neq \zeta_{i'}$ when $i \neq i'$ by collecting terms with $\zeta_i = \zeta_{i'}$ if needed. As $\lambda \neq 0$, we have $q_i \neq 0$ for some i. In terms of this we have

$$\overline{s\lambda}(\hat{\zeta}_i \bowtie_{\oplus} \hat{\zeta}_i) = sq_i 1_B,$$

hence

$$\left| \nu \left(\overline{s\lambda}(\hat{\zeta}_i \bowtie_{\oplus} \hat{\zeta}_i) \right) \right| = s|q_i|,$$

for all $s > 0$. It follows that there is an $s_0 > 0$ such that

$$\left| \nu \left(\overline{s\lambda}(\hat{\zeta}_i \bowtie_{\oplus} \hat{\zeta}_i) \right) \right| > 1$$

i.e., $s\lambda \notin \mathcal{C}$ because of (7.1), for all $s > s_0$, proving that $\mathcal{C}$ is ray-wise bounded according to Definition 2.2.4. □

With this proposition we have all the elements of the abstract theory in place, which will now allow us to formulate an abstract version of Wasserstein distance of order 1 as a metric, rather than just a pseudometric.

7.2 The composite algebraic setup and W_1 norms

We return to the setup in Section 6.2, but specialise it as follows.

Definition 7.2.1. An *algebraic* (W_1, n) *structure* $(A, B, \nu, \mathcal{L})$ is a pointed (W_1, n) structure as in Definition 6.3.3, where the pointed spaces $A_1, ..., A_n$ and $B_1, ..., B_n$ are unital algebras (all of them real, or all of them complex), with their units serving as their distinguished points,

$$u_{A_j} = 1_{A_j} \quad \text{and} \quad u_{B_j} = 1_{B_j}$$

for $j = 1, ..., n$.

From the preceding development we immediately conclude the following.

Theorem 7.2.2. *Consider an algebraic* (W_1, n) *structure* $(A, B, \nu, \mathcal{L})$. *Then* $\|\cdot\|_{W_1}$ *defined by*

$$\|\lambda\|_{W_1} = \inf\{t \geq 0 : \lambda \in t\mathcal{C}_{\mathcal{L}}\}$$

for all $\lambda \in \mathcal{O}_{\mathcal{L}}$, *is a norm on* $\mathcal{O}_{\mathcal{L}}$, *called the* W_1 norm *associated to* $(A, B, \nu, \mathcal{L})$.

Proof. From Propositions 6.3.2 and 7.1.1 we know that $\mathcal{C}_{\mathcal{L}}$ is a unit ball for $\mathcal{O}_{\mathcal{L}}$, as defined in Definition 2.2.5. By Proposition 2.3.7 we are done. □

Corollary 7.2.3. *In terms of Theorem 7.2.2, the function* $W_1 : \mathcal{L} \times \mathcal{L} \to \mathbb{R}$ *defined by*

$$W_1(\eta, \theta) = \|\eta - \theta\|_{W_1}$$

is a metric on $\mathcal{L}$, *called the* Wasserstein distance of order 1 *on* $\mathcal{L}$ *associated to the algebraic* (W_1, n) *structure* $(A, B, \nu, \mathcal{L})$.

Proof. The above follows from Theorem 2.1.4. □

This theorem and its corollary are the main results up to this point and completes the development without the presence of any form of positivity assumed of the $A \to B$ maps in $\mathcal{L}$. So far the maps in $\mathcal{L}$ were only assumed to be linear and unital. In the next section we add complete positivity in a C*-algebraic framework.

7.3 C*-algebras and complete positivity

For unital C*-algebras A and B, let

$$K(A, B)$$

be the set of all channels $E : A \to B$ from A to B, where a *channel* is a unital completely positive linear (u.c.p.) map (Definition 3.1.22). The goal of this section is to define a Wasserstein distance of order 1 on $K(A, B)$. Conventionally the term Wasserstein distance applies to states, including (integrals with respect to) probability measures in the classical case, i.e., the case $B = \mathbb{C}$. However, here we use the same terminology for channels as well, as already indicated in Chapter 5.

It is a fairly straightforward matter to apply Theorem 7.2.2 in a C*-algebraic framework, essentially taking $\mathcal{L}$ in the previous two sections to be $K(A, B)$, though there are some technical points regarding this which will be made clear in the proof of Theorem 7.3.2 below. In order to make this section as directly accessible as possible, however, we formulate the definitions and results without reference to Chapter 6 and 7. References to these two chapters will only appear in the proof of Theorem 7.3.2. We start with some notation and conventions (which follow from Chapter 3).

The tensor products of C*-algebras are not merely algebraic, but are completed in some norm. Specifically, the tensor products in this section are either all minimal tensor products or all maximal tensor products. These tensor products will simply be indicated by the symbol $\otimes$. Accordingly for tensor products of maps on C*-algebras. Some standard background regarding complete positivity and tensor products of C*-algebras can be reviewed in Chapter 3. In particular we note that tensor products of channels are again channels, for both the minimal and maximal tensor products.

For easy reference, we highlight the main structure which will be used:

Definition 7.3.1. Consider unital C*-algebras $A_1, ..., A_n$ and $B_1, ..., B_n$, as well as a state ν_j on B_j for $j = 1, ..., n$. This will be referred to as an *n-composite C*-system*, denoted $(A_j, B_j, \nu_j : j = 1, ..., n)$, with the notation

$$A = A_1 \otimes ... \otimes A_n \quad \text{and} \quad B = B_1 \otimes ... \otimes B_n$$

being implied.

Given such an n-composite C*-system, the following notation, along the lines of Section 5.1, will be used to set up the Wasserstein norm and subsequent Wasserstein distance of order 1:

$$A_{\widehat{j}} := A_1 \otimes ... \widehat{A}_j ... \otimes A_n \quad \text{and} \quad B_{\widehat{j}} := B_1 \otimes ... \widehat{B}_j ... \otimes B_n,$$

and

$$B_{\leq j} := B_1 \otimes ... \otimes B_j \quad \text{and} \quad B_{\geq j} := B_j \otimes ... \otimes B_n,$$

for $j = 1, ..., n$. Keep in mind that as mentioned above, either all of these tensor products are minimal, or all are maximal. A channel $E : A \to B$ can be reduced to a channel $E_{\widehat{j}} : A_{\widehat{j}} \to B_{\widehat{j}}$ by

$$E_{\widehat{j}} := (\mathrm{id}_{B_{\leq j-1}} \otimes \nu_j \otimes \mathrm{id}_{B_{\geq j+1}}) \circ E|_{A_{\widehat{j}}}$$

for $j = 1, ..., n$. Here the restriction $E|_{A_{\widehat{j}}} : A_{\widehat{j}} \to B$ is defined via

$$E|_{A_{\widehat{j}}}(a_1 \otimes ... \widehat{a}_j ... \otimes a_n) = E(a_1 \otimes ... \otimes a_{j-1} \otimes 1_{A_j} \otimes a_{j+1} \otimes ... \otimes a_n)$$

for $a_1 \in A_1, ..., a_n \in A_n$, where 1_{A_j} denotes the unit of A_j.

Set

$$\mathcal{O}_{A,B} := \{\lambda \in \mathrm{span}_{\mathbb{R}}\, K(A,B) : \lambda(1_A) = 0\},$$

$$\mathcal{N}_j := \{D - E : D, E \in K(A,B) \text{ such that } D_{\widehat{j}} = E_{\widehat{j}}\},$$

and

$$\mathcal{N} := \bigcup_{j=1}^{n} \mathcal{N}_j \quad \text{and} \quad \mathcal{C} := \mathrm{conv}\, \mathcal{N}.$$

Here of course two channels $D, E \in K(A,B)$ are said to be *neighbouring* if $D_{\widehat{j}} = E_{\widehat{j}}$ for some $j \in \{1, ..., n\}$.

This allows us to state the main result of this chapter, which is one of the main results of the dissertation.

Theorem 7.3.2. *Let $(A_j, B_j, \nu_j : j = 1, ..., n)$ be an n-composite C*-system. Then in both the minimal and maximal tensor product setup, $\|\cdot\|_{W_1}$ defined by*

$$\|\lambda\|_{W_1} = \inf\{t \geq 0 : \lambda \in t\mathcal{C}\}$$

for all $\lambda \in \mathcal{O}_{A,B}$, is a norm on $\mathcal{O}_{A,B}$, called the W_1 norm associated to $(A_j, B_j, \nu_j : j = 1, ..., n)$.

Proof. We are going to obtain this from the algebraic (rather than C*-algebraic) setup of the previous section. Therefore we have to convert between algebraic and completed tensor products as needed. Write $A_\odot = A_1 \odot ... \odot A_n$ and define

$$K(A,B)|_\odot = \{E|_{A_\odot} : E \in K(A,B)\}.$$

Since a channel $E \in K(A,B)$ is necessarily continuous (by Remark 3.1.23), it is uniquely determined by its restriction $E|_{A_\odot}$, the latter being the usual restriction of the map E to the subset $A_\odot$ of A. Hence $K(A,B)|_\odot$ and $K(A,B)$ are in one-to-one correspondence.

Also define

$$A_{\odot,\leq j} = A_1 \odot ... \odot A_j \quad \text{and} \quad A_{\odot,\geq j} = A_j \odot ... \odot A_n,$$

and

$$A_{\leq j} = A_1 \otimes ... \otimes A_j \quad \text{and} \quad A_{\geq j} = A_j \otimes ... \otimes A_n,$$

for $j = 1, ..., n$, as well as

$$E_{\leq j} = (\mathrm{id}_{B_{\leq j}} \otimes \nu_{>j}) \circ E|_{A_{\leq j}} : A_{\leq j} \to B_{\leq j}$$
$$E_{\geq j} = (\nu_{<j} \otimes \mathrm{id}_{B_{\geq j}}) \circ E|_{A_{\geq j}} : A_{\geq j} \to B_{\geq j}$$

for any $E \in K(A,B)$.

Note that $(A_j, B_j, \nu_j : j = 1, ..., n)$ gives an algebraic (W_1, n) structure $(A_\odot, B, \nu, K(A,B)|_\odot)$ as in Definition 7.2.1. This follows from the automatic continuity of tensor products of states and identity maps in the C*-algebraic framework, along with the fact that condition (6.8) is satisfied. The latter, in terms of (6.2) and (6.3), being

$$\left(E|_{A_\odot}\right)_{\leq j} \odot \left(D|_{A_\odot}\right)_{\geq j+1} = E_{\leq j}|_{A_{\odot,\leq j}} \odot D_{\geq j+1}|_{A_{\odot,\geq j+1}} \in K(A,B)|_\odot$$

for all $D, E \in K(A,B)$, where $|_{A_\odot}$, $|_{A_{\odot,\leq j}}$ and $|_{A_{\odot,\geq j+1}}$ are the usual restrictions to the indicated algebraic tensor products. This fact in turn is true, since $E_{\leq j} \odot D_{\geq j+1}$ uniquely extends to an element $E_{\leq j} \otimes D_{\geq j+1}$ of $K(A,B)$, as $E_{\leq j}$ and $D_{\geq j+1}$ themselves are channels (being the composition of u.c.p. maps), hence indeed

$$E_{\leq j}|_{A_{\odot,\leq j}} \odot D_{\geq j+1}|_{A_{\odot,\geq j+1}} = (E_{\leq j} \otimes D_{\geq j+1})|_{A_\odot} \in K(A,B)|_\odot.$$

By Theorem 7.2.2 and the one-to-one correspondence between $K(A,B)|_\odot$ and $K(A,B)$ mentioned above, we are done. □

This leads to the following key conclusion.

Corollary 7.3.3. *Given an n-composite C*-system $(A_j, B_j, \nu_j : j = 1, ..., n)$, then in both the minimal and the maximal tensor product setup we obtain a metric W_1 on $K(A, B)$ defined by*

$$W_1(D, E) = \|D - E\|_{W_1}$$

for all $D, E \in K(A, B)$, called the Wasserstein distance of order 1 associated to $(A_j, B_j, \nu_j : j = 1, ..., n)$.

Note that Section 5.1 emerges as a special case of this section, albeit directly in the Heisenberg picture, by simply setting

$$A_j = M_{q_j} \quad \text{and} \quad B_j = M_{r_j}$$

and letting ν_j be the normalised trace on B_j.

Another special case is $B_1 = ... = B_n = \mathbb{C}$, with $\nu_1, ..., \nu_n$ becoming trivial and irrelevant, but with general unital C*-algebras $A_1, ..., A_n$. In this case $K(A, B)$ is the set of all states on $A = A_1 \otimes ... \otimes A_n$, hence W_1 is now the Wasserstein distance of order 1 between states on A. For $A_1 = ... = A_n = M_d$ this reduces to the (quantum) Wasserstein distance of order 1 studied in [29], as can be seen from Chapter 4, keeping in mind that a state on A is exactly a normalised positive linear functional μ, which in this finite dimensional case can be uniquely represented as $\mu(a) = \mathbf{Tr}(\rho a)$ for all $a \in A$ in terms of some density matrix ρ.

We have focussed on the composite system aspect of the framework. From a single system point of view, note that in the finite dimensional setup for states, and setting $n = 1$, for any states ψ and ω on $A_1 = M_d$, we have

$$W_1(\psi, \omega) = \frac{1}{2}\mathbf{Tr}|\rho_\psi - \rho_\omega|,$$

with ρ_ψ and ρ_ω being the density matrices representing ψ and ω respectively, according to [29, Proposition 2]. In the general C*-algebraic case for states (i.e., $B = \mathbb{C}$) with $n = 1$, we can therefore view W_1 as an abstract version of the trace distance between states, despite the fact that no canonical trace is specified on A_1 in this setup. Refer to Chapter 9 for further remarks related to this.

This ends our development of the Wasserstein distance of order 1. Next we study its behaviour in relation to subsystems.

Chapter 8

Subsystems and Additivity

A core idea behind W_1 is that it is built to reflect the composite structure of systems. It is therefore natural to study its basic properties in relation to subsystems of the composite systems, i.e., smaller tensor products. This is what is done in the current chapter, first in terms of the W_1 seminorms obtained for pointed (W_1, n)-structures in Chapter 6, and subsequently for the C*-algebraic framework of Section 7.3. We focus on the additivity of W_1 with respect to tensor products (see Theorems 8.3.1 and 8.4.2) and the resulting stability of W_1 (see Corollary 8.4.3). The additivity results of this chapter generalise those of [29, Section IV.C], though the techniques to achieve them are necessarily different, as [29] makes use of trace norms, which are not available in our context. The reader who wants to see the main results in the C*-algebraic context, can turn directly to Section 8.4, but the proofs and some notation rely on Sections 8.1, 8.2 and 8.3.

8.1 Pointed W_1 substructures and their W_1 seminorms

Let $(A, B, \nu, \mathcal{L})$ be a pointed (W_1, n) structure as defined in Definition 6.3.3, again writing $\mathcal{C}_{\mathcal{L}}$ for the semi unit ball in $\mathcal{O}_{\mathcal{L}}$, as given by Proposition 6.3.2. For simplicity of notation, particularly in the following sections, however, we continue to write $\mathcal{N}_j$ and $\mathcal{N}$ as in Chapter 6, rather than, say, $\mathcal{N}_{\mathcal{L},j}$ and $\mathcal{N}_{\mathcal{L}}$. To describe the related subsystems, we need some further notation.

Write

$$[n] := \{1, ..., n\},$$

and $\mathcal{P}_n$ for the collection of non-empty proper subsets J of $[n]$; by "proper"

we mean that $J \neq [n]$. The complement of $J \in \mathcal{P}_n$ will be written as

$$J' := [n] \backslash J = \{j \in [n] : j \notin J\} \in \mathcal{P}_n.$$

We define

$$A_J = \bigodot_{j \in J} A_j \quad \text{and} \quad B_J = \bigotimes_{j \in J} B_j$$

for any $J \in \mathcal{P}_n$, with the same tensor product conventions as in Section 6.2. Here the order of the A_j's in A_J is taken to be the same as in $A = A_1 \odot ... \odot A_n$, for example $A_{\{2,5\}} = A_2 \odot A_5$ rather than $A_5 \odot A_2$. Similarly for B_J and correspondingly for $\nu_J := (\nu_j)_{j \in J}$. Elementary tensors in A_J can be denoted as

$$\otimes_{j \in J} a_j$$

for $a_j \in A_j$, and similarly for B_J, for any $J \in \mathcal{P}_n$. In particular A_J is a pointed space with distinguished point

$$u_{A_J} := \otimes_{j \in J} u_{A_j}.$$

Similarly for B_J.

We need to define corresponding reductions of maps. Given $\eta \in L(A, B)$, its *reduction*

$$\eta^J := \left(\bigotimes_{j=1}^{n} \varphi_j \right) \circ \eta|_{A_J} \in L(A_J, B_J)$$

to $J \in \mathcal{P}_n$ (or over J') is defined as an obvious generalization of the reductions considered in Section 6.2, where

$$\varphi_j = \begin{cases} \mathrm{id}_{B_j} & \text{for } j \in J \\ \nu_j & \text{for } j \in J' \end{cases}$$

and with $\eta|_{A_J} : A_J \to B$ given via

$$\eta|_{A_J}(\otimes_{j \in J} a_j) = \eta(a_1 \otimes ... \otimes a_n)$$

for $\otimes_{j \in J} a_j$ an elementary tensor in A_J, by setting $a_j = u_{A_j}$ for $j \in J'$. For any subset $\mathcal{S}$ of $L(A, B)$, let

$$\mathcal{S}^J := \{\eta^J : \eta \in \mathcal{S}\}. \tag{8.1}$$

Note that one can of course reduce any $\theta \in \mathcal{S}^J$ to θ^I for any non-empty proper $I \subset J$, by the obvious adjustment of the method above to this case.

Then we have the next basic fact in terms of Definition 6.3.3, where $|J|$ denotes the cardinality of J.

Proposition 8.1.1. *Let $(A, B, \nu, \mathcal{L})$ be a pointed (W_1, n) structure and consider any $J \in \mathcal{P}_n$. The pointed spaces A_j and B_j for $j \in J$, along with $\nu_J = (\nu_j)_{j\in J}$ and $\mathcal{L}^J$ (as in (8.1)), then form a pointed $(W_1, |J|)$ structure $(A_J, B_J, \nu_J, \mathcal{L}^J)$, called a* pointed W_1 substructure *of $(A, B, \nu, \mathcal{L})$.*

Proof. Clearly $\eta^J(u_{A_J}) = u_{B_J}$ for every $\eta \in \mathcal{L}$ by η^J's definition, hence $\mathcal{L}^J \subset L(A_J, B_J)$, the tensor products of ν_j's and identity maps restrict those of $(A, B, \nu, \mathcal{L})$ and are therefore still continuous, while the analogue of (6.8) is easily seen to hold in this context by simply reducing it to $L(A_J, B_J)$. The latter is confirmed by a direct calculation similar to that in Lemma 6.3.1's proof. □

By Corollary 6.3.4, the pointed $(W_1, |J|)$ structure $(A_J, B_J, \nu_J, \mathcal{L}^J)$ provides us with a W_1 seminorm on

$$\mathcal{O}_{\mathcal{L}^J} := \{\lambda \in \operatorname{span}_{\mathbb{R}} \mathcal{L}^J : \lambda(u_{A_J}) = 0\} \tag{8.2}$$

for every $J \in \mathcal{P}_n$, still denoted as

$$\|\cdot\|_{W_1},$$

as well as the resulting pseudometric W_1 on $\mathcal{L}^J$. This is simply a case of Section 6.2, but now of course using

$$\mathcal{N}_{J,j} := \{\eta - \theta : \eta, \theta \in \mathcal{L}^J \text{ such that } \eta_{\widehat{j}} = \theta_{\widehat{j}}\} \quad \text{for} \quad j \in J \tag{8.3}$$

instead of $\mathcal{N}_1, ..., \mathcal{N}_n$, where the latter led to the W_1 seminorm on $\mathcal{O}_{\mathcal{L}}$. Here

$$\eta_{\widehat{j}} := \eta^{J\setminus\{j\}}$$

for $\eta \in \mathcal{L}^J$, which can equivalently be defined by (6.1), but using A_i, B_i and ν_i only for $i \in J$ when setting up Chapter 6, rather than for the entire range $i = 1, ..., n$. The semi unit ball leading to this W_1 seminorm is

$$\mathcal{C}_{\mathcal{L}^J} := \operatorname{conv} \mathcal{N}_J, \tag{8.4}$$

where

$$\mathcal{N}_J := \bigcup_{j\in J} \mathcal{N}_{J,j}. \tag{8.5}$$

8.2 Reducible pointed W_1 structures

A natural question is whether the semi unit ball $\mathcal{C}_{\mathcal{L}^J}$ above can be obtained as the reduction $\mathcal{C}_{\mathcal{L}}^J$ of the original semi unit ball $\mathcal{C}_{\mathcal{L}}$. Similarly for $\mathcal{O}_{\mathcal{L}^J}$ and the sets $\mathcal{N}_{J,j}$. These questions will in fact become relevant in the next section, when we reach the main goal of this chapter, namely to prove additivity properties.

In order to answer these questions positively, assumptions beyond those in Definition 6.3.3 need to be made. We note that these assumptions will automatically be satisfied in the C*-algebraic framework.

As in the previous section we consider a pointed (W_1, n) structure $(A, B, \nu, \mathcal{L})$ and any $J \in \mathcal{P}_n$.

To avoid any mismatches and ambiguities, we always need to preserve the ordering of the A_j's in any tensor product of them. Similarly for the B_j's. Therefore the notation

$$A_I \vec{\odot} A_J := A_{I \cup J} \quad \text{and} \quad B_I \vec{\otimes} B_J := B_{I \cup J}$$

will be used for any $I, J \in \mathcal{P}_n$ with no points in common, i.e., $I \cap J = \varnothing$. But then the tensor product of maps $\eta \in L(A_I, B_I)$ and $\theta \in L(A_J, B_J)$ for $I, J \in \mathcal{P}_n$ with $I \cap J = \varnothing$ need to be defined correspondingly. This is indeed possible. Note that transpositions of adjacent A_j's in any tensor product of A_j's are linear bijections, compositions of which in particular give us a natural unique well defined *ordering map*

$$\alpha_{IJ} : A_I \odot A_J \to A_I \vec{\odot} A_J,$$

such that

$$\alpha_{IJ}((\otimes_{i \in I} a_i) \otimes (\otimes_{j \in J} a_j)) = \otimes_{j \in I \cup J} a_j$$

for arbitrary $a_j \in A_j$. As a simple example to clarify the meaning of this, suppose $I = \{1, 3, 5\}$ and $J = \{2, 4\}$, then for $a_j \in A_j$, we have $\alpha_{IJ}(a_1 \otimes a_3 \otimes a_5 \otimes a_2 \otimes a_4) = a_1 \otimes a_2 \otimes a_3 \otimes a_4 \otimes a_5$. Note that α_{IJ} can be viewed as a *pointed space isomorphism*, i.e., a bijection $\alpha_{IJ} \in L_u(A_I \odot A_J, A_I \vec{\odot} A_J)$.

Similarly we have the ordering map

$$\beta^0_{IJ} : B_I \odot B_J \to B_I \vec{\odot} B_J.$$

Clearly $\beta^0_{IJ} \in L_u(B_I \odot B_J, B_I \vec{\odot} B_J)$.

However, if the tensor products B_J for $J \subset [n]$ are indeed completed in some norm, we need to *assume* that β^0_{IJ} is continuous in this norm, and therefore extends uniquely to a continuous bijection

$$\beta_{IJ} \in L_u(B_I \otimes B_J, B_I \vec{\otimes} B_J),$$

which is the uniquely defined pointed space isomorphism serving as the ordering map on $B_I \otimes B_J$.

Given this, we can define

$$\eta \vec{\odot} \theta : A_I \vec{\odot} A_J \to B_I \vec{\otimes} B_J$$

as

$$\eta \vec{\odot} \theta := \beta_{IJ} \circ (\eta \odot \theta) \circ \alpha^{-1}_{IJ}.$$

It has the following expected property.

Proposition 8.2.1. *In terms of the notation and assumptions so far in this section,*

$$\eta \vec{\odot} \theta = \theta \vec{\odot} \eta,$$

for all $\eta \in L(A_I, B_I)$ and $\theta \in L(A_J, B_J)$, where $I, J \in \mathcal{P}_n$ with $I \cap J = \varnothing$.

Proof. Note that by the definition of the ordering maps, we have $\beta_{IJ}(c \otimes d) = \beta_{JI}(d \otimes c)$ for elements in the algebraic tensor products, $c \in \odot_{i \in I} B_i$ and $d \in \odot_{j \in J} B_j$. Now, for arbitrary $a_j \in A_j$,

$$\begin{aligned}(\eta \odot \theta) \circ \alpha^{-1}_{IJ}(\otimes_{j \in I \cup J} a_j) &= \eta(\otimes_{i \in I} a_i) \otimes \theta(\otimes_{j \in J} a_j) \\ (\theta \odot \eta) \circ \alpha^{-1}_{JI}(\otimes_{j \in I \cup J} a_j) &= \theta(\otimes_{j \in J} a_j) \otimes \eta(\otimes_{i \in I} a_i).\end{aligned}$$

Approximate $\eta(\otimes_{i \in I} a_i)$ and $\theta(\otimes_{j \in J} a_j)$ by sequences (c_l) and (d_l) in the algebraic tensor products $\odot_{i \in I} B_i$ and $\odot_{j \in J} B_j$ respectively. Since $\beta_{IJ}(c_l \otimes d_l) = \beta_{JI}(d_l \otimes c_l)$ and β_{IJ} and β_{JI} are assumed to be continuous, it follows that

$$\beta_{IJ} \circ (\eta \odot \theta) \circ \alpha^{-1}_{IJ}(\otimes_{j \in I \cup J} a_j) = \beta_{JI} \circ (\theta \odot \eta) \circ \alpha^{-1}_{JI}(\otimes_{j \in I \cup J} a_j),$$

as required. □

We also need to strengthen (6.8) in Lemma 6.3.1 to the following: *Assume* that

$$\eta^J \vec{\odot} \theta^{J'} \in \mathcal{L} \tag{8.6}$$

for all $\eta, \theta \in \mathcal{L}$ and $J \in \mathcal{P}_n$.

Definition 8.2.2. The pointed (W_1, n) structure $(A, B, \nu, \mathcal{L})$ is called *reducible* if both the above mentioned assumptions are indeed satisfied, namely (8.6) and the existence of the continuous ordering maps β_{IJ} in the case of completed B.

Note that for a non-empty $I \subset J'$, it follows from (8.6) and the definition of reduction in the previous section, that

$$\eta^J \vec{\odot} \theta^I = (\eta^J \vec{\odot} \theta^{J'})^{J \cup I} \in \mathcal{L}^{J \cup I}. \tag{8.7}$$

In particular, this gives the following simple result.

Proposition 8.2.3. *If the pointed (W_1, n) structure $(A, B, \nu, \mathcal{L})$ is reducible, then so are its pointed W_1 substructures.*

The reason for the terminology "reducible" in Definition 8.2.2, is that the semi unit ball of $(A_J, B_J, \nu_J, \mathcal{L}^J)$ is then obtained from that of $(A, B, \nu, \mathcal{L})$ by reduction. This and related facts are shown below.

In terms of the setup and notation of this chapter (specifically the notation given Section 8.1), we have the following.

Lemma 8.2.4. *Assume that $(A, B, \nu, \mathcal{L})$ is a reducible pointed (W_1, n) structure and consider any $J \in \mathcal{P}_n$. For all $\lambda \in \mathcal{N}_{J,j}$ with $j \in J$, and $\zeta \in \mathcal{L}^{J'}$, it follows that $\lambda \vec{\odot} \zeta \in \mathcal{N}_j$.*

Proof. Given that $\lambda \in \mathcal{N}_{J,j}$, (8.3) implies that $\lambda = \eta - \theta$ for some $\eta, \theta \in \mathcal{L}^J$ such that $\eta_{\hat{j}} = \theta_{\hat{j}}$. Hence $\lambda_{\hat{j}} = 0$, which impiles that $(\lambda \vec{\odot} \zeta)_{\hat{j}} = \lambda_{\hat{j}} \vec{\odot} \zeta = 0$. This means that $\lambda \vec{\odot} \zeta = \eta \vec{\odot} \zeta - \theta \vec{\odot} \zeta$, where $\eta \vec{\odot} \zeta, \theta \vec{\odot} \zeta \in \mathcal{L}$ by (8.6) and $(\eta \vec{\odot} \zeta)_{\hat{j}} = (\theta \vec{\odot} \zeta)_{\hat{j}}$, as needed in (6.6). □

Proposition 8.2.5. *Assume that $(A, B, \nu, \mathcal{L})$ is a reducible pointed (W_1, n) structure and consider any $J \in \mathcal{P}_n$. Then $\mathcal{N}_j^J = \{0\}$ when $j \in J'$, while*

$$\mathcal{N}_j^J = \mathcal{N}_{J,j}$$

for $j \in J$.

Proof. By (8.1), $\mathcal{N}_j^J := \{\eta^J : \eta \in \mathcal{N}_j\}$. For $j \in J'$ the reduction over j is included in the reduction over J', hence $\lambda_{\hat{j}} = 0$ for $\lambda \in L(A, B)$ implies that $\lambda^J = 0$, directly from the definitions of $\lambda_{\hat{j}}$ and λ^J. Thus $\mathcal{N}_j^J = \{0\}$ when $j \in J'$. Now assume that $j \in J$. Consider any $\lambda^J \in \mathcal{N}_j^J$, i.e., we take

$\lambda = \eta - \theta$ with $\eta, \theta \in \mathcal{L}$ and $\eta_{\widehat{j}} = \theta_{\widehat{j}}$. Since $j \in J$, both sides of $(\eta^J)_{\widehat{j}} = (\eta_{\widehat{j}})^J$ are well defined, and indeed equal by the definitions of these reductions. Similarly for θ, which means that $(\eta^J)_{\widehat{j}} = (\theta^J)_{\widehat{j}}$, hence $\lambda^J \in \mathcal{N}_{J,j}$ by (8.3). This shows that $\mathcal{N}_j^J \subset \mathcal{N}_{J,j}$, even if $(A, B, \nu, \mathcal{L})$ is not assumed reducible. Conversely, consider any $\lambda \in \mathcal{N}_{J,j}$. For any $\zeta \in \mathcal{L}^{J'}$ it then follows from Lemma 8.2.4 that $\lambda \vec{\odot} \zeta \in \mathcal{N}_j$. Consequently, $\lambda = (\lambda \vec{\odot} \zeta)^J \in \mathcal{N}_j^J$, proving that $\mathcal{N}_{J,j} \subset \mathcal{N}_j^J$. □

In particular this tells us that the reductions $\mathcal{N}_j^J$ of $\mathcal{N}_j$ for $j \in J$, play the same role for $(A_J, B_J, \nu_J, \mathcal{L}^J)$ as $\mathcal{N}_1, ..., \mathcal{N}_n$ play for the reducible pointed W_1 structure $(A, B, \nu, \mathcal{L})$.

Corollary 8.2.6. *In Proposition 8.2.5 we have*

$$\mathcal{O}_{\mathcal{L}^J} = \mathcal{O}_{\mathcal{L}}^J, \quad \mathcal{N}_J = \mathcal{N}^J \quad \text{and} \quad \mathcal{C}_{\mathcal{L}^J} = \mathcal{C}_{\mathcal{L}}^J$$

for (8.2), (8.5) and (8.4).

Proof. From Proposition 8.2.5, (6.7) and (8.5) one has $\mathcal{N}_J = \mathcal{N}^J$, hence

$$\mathcal{C}_{\mathcal{L}^J} = \operatorname{conv} \mathcal{N}^J = (\operatorname{conv} \mathcal{N})^J = \mathcal{C}_{\mathcal{L}}^J$$

and

$$\mathcal{O}_{\mathcal{L}^J} = \operatorname{span}_{\mathbb{R}} \mathcal{N}^J = (\operatorname{span}_{\mathbb{R}} \mathcal{N})^J = \mathcal{O}_{\mathcal{L}}^J,$$

because of Lemma 6.3.1. □

As one may expect, $\mathcal{O}_{\mathcal{L}^J} = \mathcal{O}_{\mathcal{L}}^J$ can alternatively be proved along the lines of the proof of Proposition 8.2.5.

We can also use Lemma 8.2.4 along with Corollary 8.2.6 to obtain the following property of $\|\cdot\|_{W_1}$.

Proposition 8.2.7. *Assume that $(A, B, \nu, \mathcal{L})$ is a reducible pointed (W_1, n) structure and consider any $I, J \in \mathcal{P}_n$ with $I \cap J = \varnothing$. Then*

$$\left\| (\eta - \theta) \vec{\odot} \zeta \right\|_{W_1} = \|\eta - \theta\|_{W_1}$$

for all $\eta, \theta \in \mathcal{L}^I$ and $\zeta \in \mathcal{L}^J$.

Proof. For any $\eta, \theta \in \mathcal{L}^I$ and $\zeta \in \mathcal{L}^J$, set $\lambda := (\eta - \theta) \vec{\odot} \zeta \in \mathcal{L}^{I \cup J}$, according to (8.7), then $\lambda^I = \eta - \theta$.

For any $\gamma \in \mathcal{C}_{\mathcal{L}^{I \cup J}}$ and $t \geq 0$ such that $\lambda = t\gamma$, one has that $\lambda^I = t\gamma^I$. By Corollary 8.2.6, but applied to the reducible (Proposition 8.2.3) pointed

W_1 structure $(A_{I\cup J}, B_{I\cup J}, \nu_{I\cup J}, \mathcal{L}^{I\cup J})$, we know that $\gamma^I \in \mathcal{C}_{\mathcal{L}^I}$, from which it follows that $\left\|\lambda^I\right\|_{W_1} \leq t$, hence $\left\|\lambda^I\right\|_{W_1} \leq \|\lambda\|_{W_1}$ by the definition of $\|\cdot\|_{W_1}$ in Corollary 6.3.4 (via Proposition 2.3.7).

Conversely, consider any $\gamma \in \mathcal{C}_{\mathcal{L}^I}$ and $t \geq 0$ such that $\lambda^I = t\gamma$. By (8.4) we have

$$\gamma = \sum_{i=1}^{l} p_i \gamma_i$$

for some $p_1, ..., p_l > 0$ with $p_1 + ... + p_l = 1$, and $\gamma_i \in \mathcal{N}_{I,j_i}$ for some $j_i \in I$. Because of Lemma 8.2.4 applied to $(A_{I\cup J}, B_{I\cup J}, \nu_{I\cup J}, \mathcal{L}^{I\cup J})$, it follows that

$$\gamma \vec{\odot} \zeta = \sum_{i=1}^{l} p_i \gamma_i \vec{\odot} \zeta \in \mathcal{C}_{\mathcal{L}^{I\cup J}}.$$

Since $\lambda = t\gamma\vec{\odot}\zeta$, we conclude that $\|\lambda\|_{W_1} \leq t$, thus $\|\lambda\|_{W_1} \leq \left\|\lambda^I\right\|_{W_1}$. □

These results will be applied in the next section to prove the additivity of W_1 with respect to tensor products.

8.3 Additivity

We now arrive at this chapter's main results in the abstract pointed space setup, which will be applied to the C*-algebras in the next section. Consider a reducible pointed (W_1, n) structure $(A, B, \nu, \mathcal{L})$ and any *m-partition* P of $[n]$, by which we mean a function $P : [m] \to \mathcal{P}_n$ such that $P(1) \cup ... \cup P(m) = [n]$ and $P(k) \cap P(l) = \varnothing$ for $k \neq l$. Our goal is to determine how $\|\cdot\|_{W_1}$ for $(A, B, \nu, \mathcal{L})$ relates to $\|\cdot\|_{W_1}$ for the $(A_{P(k)}, B_{P(k)}, \nu_{P(k)}, \mathcal{L}^{P(k)})$'s via reduction. Similarly for W_1, but specifically for product maps. Using the results of the previous section, these relationships can be stated as a form of "reductive superadditivity" of $\|\cdot\|_{W_1}$ and an additivity property of W_1.

Theorem 8.3.1. *Let $(A, B, \nu, \mathcal{L})$ be a reducible pointed (W_1, n) structure and P any m-partition of $[n]$. Then*

$$\|\lambda\|_{W_1} \geq \sum_{k=1}^{m} \left\|\lambda^{P(k)}\right\|_{W_1} \tag{8.8}$$

for all $\lambda \in \mathcal{O}_{\mathcal{L}}$, while

$$W_1(\eta_1 \vec{\odot} ... \vec{\odot} \eta_m, \theta_1 \vec{\odot} ... \vec{\odot} \theta_m) = \sum_{k=1}^{m} W_1(\eta_k, \theta_k) \tag{8.9}$$

for all $\eta_k, \theta_k \in \mathcal{L}^{P(k)}$, *for* $k = 1, ..., m$.

Proof. Let $\lambda \in \mathcal{O}_{\mathcal{L}}$. Consider any $\gamma \in \mathcal{C}_{\mathcal{L}}$ and $t \geq 0$ such that $\lambda = t\gamma$. By $\mathcal{C}_{\mathcal{L}} = \operatorname{conv} \mathcal{N}$,

$$\gamma = \sum_{i=1}^{l} p_i \alpha_i$$

for some $p_1, ..., p_l > 0$ with $p_1 + ... + p_l = 1$, and $\alpha_i \in \mathcal{N} = \cup_{j \in [n]} \mathcal{N}_j$. Set

$$q_k := \sum_{i \in R_k} p_i \quad \text{and} \quad \gamma_k := \sum_{i \in R_k} \frac{p_i}{q_k} \alpha_i \in \mathcal{C}_{\mathcal{L}},$$

where $R_k := \{i \in [l] : \alpha_i \in \cup_{j \in P(k)} \mathcal{N}_j\} \backslash (R_1 \cup ... \cup R_{k-1})$ for $k = 1, ..., m$, with $R_1 \cup ... \cup R_{k-1} = \varnothing$ for $k = 1$. Then by Proposition 8.2.5,

$$\gamma = \sum_{k=1}^{m} q_k \gamma_k \quad \text{and} \quad \gamma^{P(k)} = q_k \gamma_k^{P(k)}$$

with $\gamma_k^{P(k)} \in \mathcal{C}_{\mathcal{L}}^{P(k)} = \mathcal{C}_{\mathcal{L}^{P(k)}}$ by Corollary 8.2.6. This tells us that $\lambda^{P(k)} = t\gamma^{P(k)} = tq_k \gamma_k^{P(k)} \in tq_k \mathcal{C}_{\mathcal{L}^{P(k)}}$, hence $\left\| \lambda^{P(k)} \right\|_{W_1} \leq tq_k$, thus

$$\sum_{k=1}^{m} \left\| \lambda^{P(k)} \right\|_{W_1} \leq t,$$

implying (8.8), by $\|\cdot\|_{W_1}$'s definition in Corollary 6.3.4.

In particular, for $\eta, \theta \in \mathcal{L}$ we have

$$\|\eta - \theta\|_{W_1} \geq \sum_{k=1}^{m} \left\| \eta^{P(k)} - \theta^{P(k)} \right\|_{W_1}.$$

On the other hand, given $\eta_k, \theta_k \in \mathcal{L}^{P(k)}$, by Proposition 8.2.7 we have

$$\begin{aligned}
&\left\| \eta_1 \vec{\odot} ... \vec{\odot} \eta_m - \theta_1 \vec{\odot} ... \vec{\odot} \theta_m \right\|_{W_1} \\
&\leq \left\| (\eta_1 - \theta_1) \vec{\odot} \eta_2 \vec{\odot} ... \vec{\odot} \eta_m \right\|_{W_1} + \left\| \theta_1 \vec{\odot} (\eta_2 \vec{\odot} ... \vec{\odot} \eta_m - \theta_2 \vec{\odot} ... \vec{\odot} \theta_m) \right\|_{W_1} \\
&= \|\eta_1 - \theta_1\|_{W_1} + \left\| \eta_2 \vec{\odot} ... \vec{\odot} \eta_m - \theta_2 \vec{\odot} ... \vec{\odot} \theta_m \right\|_{W_1} \\
&\vdots \\
&\leq \sum_{k=1}^{m} \|\eta_k - \theta_k\|_{W_1}.
\end{aligned}$$

These two inequalities prove (8.9). □

With this result in hand, we can turn to the final aim of this chapter, namely additivity in the C*-algebraic framework.

8.4 C*-algebras

The question being studied at the moment, is how W_1 between product channels from one compound system to another, relate to W_1 between the channels composing the product channels. This was answered in a more abstract form in Theorem 8.3.1 of the previous section, namely they are related in a simple additive way. Now we essentially just translate this additivity of W_1 to the C*-algebraic framework of Section 7.3.

Given an n-composite C*-system $(A_j, B_j, \nu_j : j = 1, ..., n)$ as in Definition 7.3.1, let $[n] := \{1, ..., n\}$, let $\mathcal{P}_n$ be the collection of non-empty proper subsets of $[n]$, and define an *m-partition* P of $[n]$, as a function $P : [m] \to \mathcal{P}_n$ such that $P(1) \cup ... \cup P(m) = [n]$ and $P(k) \cap P(l) = \varnothing$ for $k \neq l$. Furthermore, we set

$$A_J := \bigotimes_{j \in J} A_j \quad \text{and} \quad B_J := \bigotimes_{j \in J} B_j$$

for any $J \in \mathcal{P}_n$, both being minimal tensor products or both maximal tensor products. As in Section 7.3, either all tensor products here are minimal, or all are maximal. For any $I, J \in \mathcal{P}_n$ with no points in common, i.e., $I \cap J = \varnothing$, consider the ordered tensor products

$$A_I \vec{\otimes} A_J := A_{I \cup J} \quad \text{and} \quad B_I \vec{\otimes} B_J := B_{I \cup J}$$

and the corresponding ordered tensor product

$$\eta \vec{\otimes} \theta : A_I \vec{\otimes} A_J \to B_I \vec{\otimes} B_J$$

of bounded linear maps $\eta : A_I \to B_I$ and $\theta : A_J \to B_J$, defined as in Section 8.2, but via the minimal or maximal tensor product $\eta \otimes \theta$, instead of the algebraic tensor product, and where the ordering map α_{IJ} is of course extended to the completed tensor product, like β_{IJ} is. In the C*-algebraic case, this extension of the ordering maps is automatically possible (see the proof of Proposition 8.4.1 below). Note that the reduction η^J of $\eta \in L(A, B)$ is defined analogously to Section 8.1, the only difference being that A and A_J are now completed tensor products.

The next result is the key to convert Theorem 8.3.1 to the C*-algebraic setup.

Proposition 8.4.1. *Consider the setup of this section. For any $J \in \mathcal{P}_n$, it follows that*

$$K(A,B)^J = K(A_J, B_J).$$

In addition, for any $D \in K(A_I, B_I)$ and $E \in K(A_J, B_J)$ with $I, J \in \mathcal{P}_n$ such that $I \cap J = \varnothing$, one has

$$D\vec{\otimes}E \in K(A_I\vec{\otimes}A_J, B_I\vec{\otimes}B_J).$$

Proof. Starting with the latter statement, note that in this C*-algebraic setup the ordering maps from Section 8.2 are $*$-isomorphisms $\alpha_{IJ} : A_I \otimes A_J \to A_I\vec{\otimes}A_J$ and $\beta_{IJ} : B_I \otimes B_J \to B_I\vec{\otimes}B_J$ (and therefore extended to the completions and u.c.p.) from the outset, since they are compositions of the transposition maps mentioned in Section 8.2, which are indeed $*$-isomorphisms (see for example [6, II.9.2.6]). Hence $D\vec{\otimes}E := \beta_{IJ} \circ (D \otimes E) \circ \alpha_{IJ}^{-1}$ is a composition of u.c.p. maps, since $D\vec{\otimes}E$ is a channel. Consequently $D\vec{\otimes}E \in K(A_I\vec{\otimes}A_J, B_I\vec{\otimes}B_J)$, as required.

In particular, for $I = J'$, it follows that $E = (D\vec{\otimes}E)^J \in K(A,B)^J$. Hence $K(A_J, B_J) \subset K(A,B)^J$. Conversely, since the reduction of a channel is again a channel, we have $K(A,B)^J \subset K(A_J, B_J)$. □

Now we can answer the above mentioned question as follows, along with the "reductive superadditivity" of $\|\cdot\|_{W_1}$.

Theorem 8.4.2. *Let $(A_j, B_j, \nu_j : j = 1, ..., n)$ be an n-composite C*-system and P any m-partition of $[n]$. Then in both the minimal and the maximal tensor product setup,*

$$W_1(D_1\vec{\otimes}...\vec{\otimes}D_m, E_1\vec{\otimes}...\vec{\otimes}E_m) = \sum_{k=1}^{m} W_1(D_k, E_k)$$

for all $D_k, E_k \in K(A_{P(k)}, B_{P(k)})$, for $k = 1, ..., m$. In addition,

$$\|\lambda\|_{W_1} \geq \sum_{k=1}^{m} \left\|\lambda^{P(k)}\right\|_{W_1}$$

for all $\lambda \in \mathcal{O}_{A,B}$ as defined in Section 7.3.

Proof. If the algebraic (W_1, n) structure $(A_\odot, B, \nu, K(A,B)|_\odot)$, obtained as in the proof of Theorem 7.3.2 from $(A_j, B_j, \nu_j : j = 1, ..., n)$, is a reducible pointed (W_1, n) structure, and

$$\left(K(A,B)|_\odot\right)^J = K(A_J, B_J)|_\odot, \tag{8.10}$$

for $J \in \mathcal{P}_n$, then the theorem follows directly from Theorem 8.3.1. This is because of the continuity of all the maps involved and the resulting one-to-one correspondence between $K(A,B)|_\odot$ and $K(A,B)$ explained in the proof of Theorem 7.3.2, and similarly for $K(A_J,B_J)|_\odot$ and $K(A_J,B_J)$, which allows us to translate directly between the algebraic and completed tensor products. Condition (8.10) is needed to ensure that for $\mathcal{L} = K(A,B)|_\odot$ in Theorem 8.3.1, we have $\mathcal{L}^{P(k)} = K(A_{P(k)},B_{P(k)})|_\odot$. Note that (8.10) indeed holds because of Proposition 8.4.1 and $\big(K(A,B)|_\odot\big)^J = K(A,B)^J|_\odot$, the latter being true since the only difference between the reductions on the two sides is that they respectively involve restrictions to the algebraic and a completed tensor product of A_j's. We are simply left to verify reducibility.

As mentioned in the proof of Proposition 8.4.1, the β_{IJ}'s are $*$-isomorphisms and therefore continuous, verifying the one condition for reducibility in Definition 8.2.2. The other condition is guaranteed by the special case of Proposition 8.4.1 with $I \cup J = [n]$. □

Recall that the special case in finite dimensions of the first part of this theorem was already mentioned in Section 5.1, in a special form where the ordering of the tensor products was unnecessary.

As an immediate consequence of this theorem, we obtain the following stability result for W_1.

Corollary 8.4.3. *In Theorem 8.4.2, let P a 2-partition of $[n]$. Then*

$$W_1(D_1\vec{\otimes}F_2, E_1\vec{\otimes}F_2) = W_1(D_1,E_1)$$

for all $D_1, E_1 \in K(A_{P(1)},B_{P(1)})$ and any $F_2 \in K(A_{P(2)},B_{P(2)})$. Similarly,

$$W_1(F_1\vec{\otimes}D_2, F_1\vec{\otimes}E_2) = W_1(D_2,E_2)$$

for all $D_2, E_2 \in K(A_{P(2)},B_{P(2)})$ and any $F_1 \in K(A_{P(1)},B_{P(1)})$.

Proof. This is simply because $W_1(F_2,F_2) = 0$ in the first case. Similarly for the second. □

Stability for other distances between channels have been presented and discussed in for example [3] and [37]. The latter in particular emphasises the utility of stability. These references only treated the case where the F_k's in the corollary above were identity maps.

Of course, stability for pointed W_1 structures similarly follow from Theorem 8.3.1.

Chapter 9

Conclusion and outlook

In this dissertation our focus has been the mathematical development of a Wasserstein distance of order 1 between channels from one composite system to another, denoted by W_1. We have not yet investigated any relations or comparisons of W_1 with the diamond norm (mentioned in the introduction) or other distances between channels. See [37] and [11] for a critical assessment of various such distances. A logical first step is to do this in finite dimensions on matrix algebras M_d, where one has a simple canonical trace which plays an important role in the diamond (aka, completely bounded trace) norm. Indeed, [29] extensively explored the relation between W_1 for states and the trace norm, including the characterization of W_1 for the case of $n = 1$, i.e., for single systems rather than composed systems. This can analogously be explored for the case of channels.

We have also not yet explored applications of this distance. Applications of Wasserstein distance of order 1 between states in the finite dimensional case developed in [29], have already been treated and proposed in [29] itself, as well as in a number of papers [28, 31, 42, 47, 52, 61] in various contexts. We expect that the approach of this dissertation to the case of channels should be similarly applicable. In addition, our abstract approach has the potential to allow for applications in other contexts than quantum channels and for further mathematical development.

At the end of Section 8.4 we pointed out that W_1 satisfies stability. However, there are other properties that one may also want a distance between channels to satisfy, depending on the application. See for example [11, 37]. Since W_1 is a metric, a number of basic properties are already satisfied. One property that we have not discussed in this dissertation is chaining or bounds on the overall difference between composed channels; see [37] and [3, Subsec-

tion 5.4]. Whether or not this or similar and other properties hold for W_1, certainly warrants further investigation.

Bibliography

[1] S. Aaronson, G. N. Rothblum, Gentle measurement of quantum states and differential privacy, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, pages 322–333, Association for Computing Machinery, New York, NY, USA, 2019.

[2] J. Agredo, A Wasserstein-type distance to measure deviation from equilibrium of quantum Markov semigroups, *Open Syst. Inf. Dyn.* **20** (2013), 1350009.

[3] D. Aharonov, A. Kitaev, N. Nisan, Quantum circuits with mixed states, in *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, STOC '98, pages 20–30, Association for Computing Machinery, New York, NY, USA, 1998.

[4] J.-D. Benamou, Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.* **84** (2000), 375–393.

[5] P. Biane, D. Voiculescu, A free probability analogue of the Wasserstein metric on the trace-state space, *Geom. Funct. Anal.* **11** (2001), 1125–1138.

[6] B. Blackadar, *Operator Algebras: Theory of C*-Algebras and von Neumann Algebras*, Encyclopaedia of Mathematical Sciences Volume 122, Springer-Verlag, Berlin, Heidelberg, 2006.

[7] F. F. Bonsall, J. Duncan, *Complete normed algebras*, Ergebnisse der Mathematik und ihrer Grenzgebiete Band 80, Springer -Verlag, Berlin, Heidelberg, New York, 1973.

[8] S. P. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge university press, 2004.

[9] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics 1*, 2nd ed. (Springer-Verlag, New York, 1987).

[10] N. Brown and N. Ozawa, *C*-Algebras and Finite-Dimensional Approximations*, Graduate Studies in Mathematics Volume 88, AMS, 2008.

[11] D. G. Bussandri, P. W. Lamberti and K. Życzkowski, Transmission distance in the space of quantum channels. arXiv preprint arXiv:2304.08403.

[12] E. A. Carlen, J. Maas, An analog of the 2-Wasserstein metric in non-commutative probability under which the fermionic Fokker-Planck equation is gradient flow for the entropy, *Comm. Math. Phys.* **331** (2014), 887–926.

[13] E. A. Carlen, J. Maas, Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance, *J. Funct. Anal.* **273** (2017), 1810–1869.

[14] E. A. Carlen, J. Maas, Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems, *J. Stat. Phys.* **178** (2020), 319–378.

[15] F. Caruso, V. Giovannetti, C. Lupo, S. Mancini, Quantum channels and memory effects, *Rev. Mod. Phys.* **86** (2014), 1203–1259.

[16] Y. Chen, W. Gangbo, T. T. Georgiou, A. Tannenbaum, On the matrix Monge-Kantorovich problem, *European J. Appl. Math.* **31** (2020), 574–600.

[17] Y. Chen, T. T. Georgiou, L. Ning, A. Tannenbaum, Matricial Wasserstein-1 distance, *IEEE Control Syst. Lett.* **1** (2017), 14–19.

[18] Y. Chen, T. T. Georgiou, A. Tannenbaum, Matrix optimal mass transport: a quantum mechanical approach, *IEEE Trans. Automat. Control* **63** (2018), 2612–2619.

[19] M.-D. Choi, Completely positive linear maps on complex matrices, *Linear Algebra Appl.* **10** (1975), 285–290.

[20] S. Cole, M. Eckstein, S. Friedland, K. Życzkowski, Quantum optimal transport, arXiv:2105.06922.

[21] A. Connes, Compact metric spaces, Fredholm modules, and hyperfiniteness, *Ergodic Theory Dynam. Systems* **9** (1989), 207–220.

[22] A. Connes, *Noncommutative geometry*, Academic Press, Inc., San Diego, CA, 1994.

[23] A. Connes, Gravity coupled with matter and the foundation of non-commutative geometry, *Comm. Math. Phys.* **182** (1996), 155–176.

[24] J. B. Conway, *A course in functional analysis*, 2nd ed., Graduate texts in mathematics volume 96, Springer, 2007.

[25] T. W. Gamelin and R. E. Greene. *Introduction to topology*, Courier Corporation, 1999.

[26] M. J. Goddard, *The Choi-Jamiołkowski duality*, BSc Physics essay, University of Pretoria, 2021.

[27] F. D'Andrea, P. Martinetti, A view on optimal transport from noncommutative geometry, *SIGMA Symmetry Integrability Geom. Methods Appl.* **6** (2010), 057.

[28] G. De Palma, M. Marvian, C. Rouzé, D. Stilck França, Limitations of variational quantum algorithms: a quantum optimal transport approach, arXiv:2204.03455.

[29] G. De Palma, M. Marvian, D. Trevisan, S. Lloyd, The quantumWasserstein distance of order 1, *IEEE Trans. Inf. Theory* **67** (2021) 6627–6643.

[30] G. De Palma, D. Trevisan, Quantum optimal transport with quantum channels, *Ann. Henri Poincaré* **22**(2021), 3199–3234.

[31] G. De Palma, C. Rouzé, Quantum Concentration Inequalities, *Ann. Henri Poincaré* **23** (2022), 3391–3429.

[32] J. de Pillis, Linear transformations which preserve hermitian and positive semidefinite operators, Pacific J. Math. 23 (1967), 129–137.

[33] R. Duvenhage, Quadratic Wasserstein metrics for von Neumann algebras via transport plans, *J. Operator Theor*y, to appear, arXiv:2012.03564.

[34] R. Duvenhage, M. Snyman, Balance between quantum Markov semigroups, *Ann. Henri Poincaré* **19** (2018), 1747–1786.

[35] J. B. Fraleigh, R. A. Beauregard, *Linear Algebra*, 3rd ed., Addison-Wesley, 1995.

[36] L. Gao, M. Junge, N. LaRacuente, Fisher information and logarithmic Sobolev inequality for matrix-valued functions. *Ann. Henri Poincaré* **21** (2020), no. 11, 3409–3478.

[37] A. Gilchrist, N. K. Langford, M. A. Nielsen, Distance measures to compare real and ideal quantum processes, *Phys. Rev. A* **71** (2005), 062310.

[38] F. Golse, C. Mouhot, T. Paul, On the mean field and classical limits of quantum mechanics, *Comm. Math. Phys.* **343** (2016), 165–205.

[39] F. Golse, T. Paul, The Schrödinger equation in the mean-field and semiclassical regime, *Arch. Ration. Mech. Anal.* **223** (2017), 57–94.

[40] F. Golse, T. Paul, Wave packets and the quadratic Monge-Kantorovich distance in quantum mechanics, *C. R. Math. Acad. Sci. Paris* **356** (2018), 177–197.

[41] R. Haggarty, *Fundamentals of mathematical analysis*, Prentice Hall, 1993.

[42] C. Hirche, C. Rouzé, D. Stilck França, Quantum Differential Privacy: An Information Theory Perspective, arXiv:2202.10717.

[43] A. S. Holevo, V. Giovannetti, Quantum channels and their entropic characteristics, *Rep. Progr. Phys.* **75** (2012), 046001.

[44] D. F. Hornshaw, L^2-Wasserstein distances of tracial W*-algebras and their disintegration problem, arXiv:1806.01073.

[45] A. Jamiołkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, *Rep. Mathematical Phys.* **3** (1972), 275–278.

[46] R. V. Kadison and J. R, Ringrose. *Fundamentals of the Theory of Operator Algebras. Volume 1, Elementary Theory,* New York: Academic Press, 1983.

[47] B. T. Kiani, G. De Palma, M. Marvian, Z.-W. Liu, S. Lloyd, Learning quantum data with the quantum earth mover's distance, *Quantum Sci. Technol.* **7** (2022), 045002.

[48] A. Kitaev, Quantum computations; algorithms and error correction, *Russian Math. Surveys* **52** (1997), 1191–1249.

[49] A. W. Knapp, *Basic algebra*, Germany, Birkhäuser Boston, 2006.

[50] S. Koudia, A. S. Cacciapuoti, K. Simonov, M. Caleffi, How Deep the Theory of Quantum Communications Goes: Superadditivity, Superactivation and Causal Activation, *IEEE Communications Surveys & Tutorials*, to appear, https://doi.org/10.1109/COMST.2022.3196449.

[51] E. Kreyszig, *Introductory functional analysis with applications*, John Wiley and Sons Inc, 1978.

[52] L. Li, K. Bu, D. E. Koh, A. Jaffe, S. Lloyd, Wasserstein Complexity of Quantum Circuits, arXiv:2208.06306.

[53] E. Merzbacher, *Quantum mechanics*, 3rd ed. John Wiley and Sons, 1998.

[54] G.J Murphy, *C*-Algebras and Operator Theory*, Academic Press, San Diego, 1990.

[55] L. Narici, E. Beckenstein, *Topological vector spaces*, Second edition, Pure and Applied Mathematics (Boca Raton), 296, CRC Press, Boca Raton, FL, 2011.

[56] M. A. Nielson, I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary ed.*, Cambridge University Press, Cambridge, 2010.

[57] L. Ning, T. T. Georgiou, Metrics between matrix-valued measures via test functions, in *53rd IEEE Conference on Decision and Control*, pages 2642–2647, IEEE, 2014.

[58] D. Ornstein, An application of ergodic theory to probability theory, *Ann. Probability* **1** (1973), 43–65.

[59] M. A. Rieffel, Metrics on state spaces, *Doc. Math.* **4** (1999), 559–600.

[60] C. Rouzé, N. Datta, Concentration of quantum states from quantum functional and transportation cost inequalities, *J. Math. Phys.* **60** (2019), 012202.

[61] C. Rouzé, D. Stilck França, Learning quantum many-body systems from a few copies, arXiv:2107.03333.

[62] W. Rudin, *Functional analysis*, 2nd ed., McGraw-Hill. Inc., New York, 1991.

[63] E. K. Ryu, Y. Chen, W. Li, S. Osher, Vector and matrix optimal mass transport: theory, algorithm, and applications, *SIAM J. Sci. Comput.* **40** (2018), A3675–A3698.

[64] G. Simmons, *Introduction to Topology and Modern Analysis*, McGraw-Hill. Inc., New York, 1963.

[65] M. Takesaki, On the cross-norm of the direct product of C*-algebras, *Tôhoku Math. J.* (2) **16** (1964), 111–122.

[66] D. W. Van Wyk, *Quantum Probability*, BScHons mathematics essay. University of Pretoria, Pretoria, 2008.

[67] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, 58, American Mathematical Society, Providence, RI, 2003.

[68] T. V. Vu, K. Saito, Thermodynamic unification of optimal transport: thermodynamic uncertainty relation, minimum dissipation, and thermodynamic speed limits, arXiv:2206.02684.

[69] J. Watrous, *The theory of quantum information*, Cambridge University Press, Cambridge, 2018.

[70] M. Wirth, A noncommutative transport metric and symmetric quantum Markov semigroups as gradient flows of the entropy, arXiv:1808.05419.