Coalgebraic Semantics for Quantum Computation

Frank Roumen

Master's thesis in Mathematics

July 2012

Supervisor: Prof. Dr. Bart Jacobs

Second reader: Prof. Dr. Ieke Moerdijk

Radboud University Nijmegen

Summary

Coalgebras generalize various kinds of dynamical systems occuring in mathematics and computer science. Examples of systems that can be modeled as coalgebras include automata and Markov chains. In this thesis we will present a coalgebraic representation of systems occuring in the field of quantum computation. This will allow us to derive a method to convert quantum mechanical systems into simpler probabilistic systems with the same behaviour. Furthermore we will formulate the equivalence of the Schrödinger and the Heisenberg pictures of quantum mechanics in a coalgebraic fashion.

Contents

Introduction

In the early 20th century, it was discovered that the laws of physics that govern large-scale phenomena fail to explain the behaviour of particles near the size of atoms. Several physicists have developed the theory of quantum mechanics to account for this small-scale behaviour. Quantum mechanics contains a number of counter-intuitive features that have no counterpart in classical physics. An example is the principle of superposition, which states, roughly speaking, that a particle need not have a definite position, but it can be in more than one position at the same time.

If we use these phenomena to perform computations on a quantum version of data, we obtain a type of computation that is fundamentally different from ordinary computation. This new type of computation is called quantum computation. The first theoretical model of quantum computation was the quantum Turing machine proposed in [8]. In addition to their importance in theoretical computer science, quantum computers are also interesting for technological applications. This is because there are quantum algorithms that are more efficient than classical algorithms that solve the same problem. One of the first examples of this was Deutsch's algorithm that checks whether a given function is constant [9].

For studying complex computational systems, it is often helpful to use an abstract description of the systems. This helps to focus on the most important parts of the system under consideration and see similarities and differences between distinct kinds of systems. The notion of a coalgebra, which originated in category theory, gives such an abstract view on dynamical systems. There is a large class of systems that can be described using coalgebras, and one can reason about these system using a unified theory. An overview can be found in [30].

The aim of this thesis is to describe systems occuring in quantum computation in the coalgebraic framework. This will facilitate comparison between quantum systems and, for example, deterministic and probabilistic systems. It will also enable us to apply facts from the general theory of coalgebras to quantum mechanical systems. We will see what minimization amounts to for quantum coalgebras, and obtain duality results for quantum systems.

The use of categories in the foundations of quantum physics was initiated in [2], via tensor categories, and in [6], via topoi. Representation of quantum systems with coalgebras is also considered in [1]. However, there are several differences between [1] and our work. We focus on the dynamics of systems via unitary operators, whereas [1] models the dynamics of iterated measurements.

The outline of this thesis is as follows. Chapter 1 is an introduction to the theory of coalgebras. It presents some well-known results that will be useful in this thesis. In particular, we will define categories of coalgebras and discuss the role of final objects in such categories. Then we will explain how coalgebras can be used to minimize systems, and explore the duality between algebras and coalgebras. We illustrate the theory with several examples of deterministic and probabilistic systems.

In Chapter 2 we will describe the basics of the mathematical formulation of quantum mechanics and quantum computation. The description uses density matrices and effects on a Hilbert space, since this will turn out to be the most convenient level of generality for the coalgebraic view. The material in this Chapter is also well-known in the literature.

The main new contributions are in Chapter 3. We define a class of coalgebras that represent quantum mechanical systems, and give several examples to show how to fit specific systems in the framework. We will also apply the minimization procedure from Chapter 1 to quantum systems. Finally we will relate the Schrödinger and Heisenberg pictures of quantum mechanics using the duality results from Chapter 1.

Throughout the text, we assume that the reader is familiar with the language of category theory. An introduction to category theory is for example [24].

Acknowledgements

I would like to thank my supervisor Bart Jacobs for his guidance during this master thesis project. He was always enthusiastic and I have learned a lot from his clear explanations. Thanks are also due to Ieke Moerdijk for being the second reader of this thesis.

Chapter 1

Coalgebras

In mathematics and computer science, one encounters many different kinds of state-based systems. Examples include finite automata, Turing machines, Markov chains, and differential equations. There are many similarities between these kinds of systems, so instead of studying each of these systems in isolation, it is better to develop a general framework to study all kinds of systems in a unified way. This abstract view often leads to a clearer understanding of the theory.

The goal of this chapter is to show that the category-theoretic notion of a coalgebra is a generalization of various kinds of systems. We will define coalgebras and give several examples to show that this notion is indeed suited as a general framework for studying systems. We will also present some well-known results about coalgebras that generalize theorems about specific systems.

This chapter only contains basic results about coalgebras that will be used in this thesis. For a more complete introduction, see e.g. [19, 30].

1.1 Definition and examples

The essential feature of a system is that it is determined by its states and transitions between these states. We will model this via the concept of a coalgebra.

Definition 1.1. Let F be an endofunctor on a category $\mathbf{C}$. An *F-coalgebra* consists of an object $X \in \mathbf{C}$ and a morphism $c : X \to F(X)$ in $\mathbf{C}$. The object X is called the *state space* of the coalgebra, and the morphism c is called the *dynamics*.

The functor F in the definition is a parameter that determines the structure of the dynamics, and hence the kind of system. Coalgebras for a fixed endofunctor constitute a category with the following morphisms.

Definition 1.2. Let $F : \mathbf{C} \to \mathbf{C}$ be an endofunctor. For two F-coalgebras $c : X \to F(X)$ and $d : Y \to F(Y)$, a *homomorphism* or *coalgebra morphism* is a morphism $f : X \to Y$ in $\mathbf{C}$ making the following diagram commute.

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "c"'] & Y \arrow[d, "d"] \\ F(X) \arrow[r, "F(f)"'] & F(Y) \end{tikzcd}$$

The category of F-coalgebras and homomorphisms is denoted $\mathbf{CoAlg}(F)$.

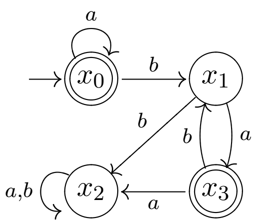
We will now look at some examples to show how to model several kinds of systems as coalgebras.

Example 1.3. Consider the endofunctor $F(X) = A \times X$ on the category **Sets**, where A is a fixed set. An F-coalgebra is a function $c : X \to A \times X$, which can also be viewed as a pair of functions $c = \langle f, g \rangle$, where $f : X \to A$ and $g : X \to X$. This coalgebra is a model of a machine with two buttons. The internal state of the machine is an element of the set X. The first button, corresponding to the function f, yields an output in the set A, for each state in X. When someone presses the second button, represented by g, the machine will take a transition to a new state. Hence f is an observation map and g is a transition map.

Example 1.4. In theoretical computer science, finite automata provide a useful, mechanical way to describe languages. Here we will first discuss the basics of finite automata, and then we will show how to define these automata coalgebraically.

Before we can define automata, we have to define formal languages. An *alphabet* is a finite set, whose elements we call *letters* or *symbols*. The set of finite sequences, or *words*, with entries in an alphabet A is written as A^*. This set of words forms a monoid, where the monoid operation is concatenation and the empty word ε acts as an identity element. The monoid A^* is the free monoid over A. A *language* over A is a subset of A^*.

A *deterministic automaton* over an alphabet A is a quadruple (X, δ, x_0, U), where X is a set of states, $\delta : X \times A \to X$ is the transition function, $x_0 \in X$ is the initial state, and $U \subseteq X$ is the set of accepting states. Automata are often graphically represented by their state diagrams. In a state diagram, each state of the automaton is drawn as a circle. A transition $\delta(x, a) = y$ is indicated by an arrow, labeled with the letter a, from the circle x to the circle y. There is an additional arrow pointing into the initial state, and accepting states are drawn as double circles. As an example, consider the automaton $\mathbb{A} = (X, \delta, x_0, U)$ over $A = \{a, b\}$ with the following diagram.



Here, the set of states is $X = \{x_0, x_1, x_2, x_3\}$, the transition function is completely determined by the arrows in the diagram, and the subset of accepting states is $\{x_0, x_3\}$.

The possible inputs of an automaton over A are the words in A^*. The automaton processes a word by starting in the initial state, reading the letters of the word from left to right, and taking the transition corresponding to the read letter to a new state. If the automaton finishes in an accepting state, then the word is said to be accepted, otherwise it is rejected. In the example above, the word $aaba$ is accepted, whereas the word $aabaa$ is rejected.

This procedure can be made precise by extending the transition function $\delta : X \times A \to X$ to a function from $X \times A^*$ to X. We will denote this extended transition function by δ^*, or, if no confusion is possible, by δ. The function δ^* is defined by induction on $u \in A^*$.

$$\begin{aligned} \delta^*(x)(\varepsilon) &= x, \\ \delta^*(x)(au) &= \delta^*(\delta(x)(a))(u). \end{aligned}$$

Each automaton has an associated language consisting of all words $u \in A^*$ for which $\delta^*(x_0)(u) \in U$. Words in this language are said to be *accepted* by the automaton. The language recognized by the example automaton is

$$\{a^m(ba)^n \mid m, n \in \mathbb{N}\}.$$

We would like to find a functor F whose coalgebras are deterministic automata. Let (X, δ, x_0, U) be an automaton. We will ignore the initial state x_0. As we shall see in the discussion on final coalgebras in Section 1.2, it works better not to have a fixed initial state. The transition function $\delta : X \times A \to X$ can be curried to obtain a function $\tilde{\delta} : X \to X^A$, and the set of accepting states $U \subseteq X$ can be represented as its characteristic function $\chi_U : X \to 2$. Combining the functions $\tilde{\delta}$ and χ_U, we obtain a single function $\langle \tilde{\delta}, \chi_U \rangle : X \to X^A \times 2$, which encodes all transition and observation structure of the automaton. This function has the shape of a coalgebra, so we see that automata (without initial state) are exactly the coalgebras for the functor $F(X) = X^A \times 2$.

1.2 Final coalgebras

When we describe a system as a coalgebra, this is always an internal description, because it specifies all states and transitions inside the system. Not all of this information is visible for a user of the system. By working with the system, the user can only obtain an external description of the system, which contains information about the observable behaviour of the system.

We will elaborate these ideas by continuing Example 1.3. We saw that a coalgebra for the functor $F(X) = A \times X$ consists of an output function $f : X \to A$ and a transition function $g : X \to X$. These two functions provide the internal description of the system. We think of this system as a black box, which means that it is impossible to observe the current state of the system. However, the output in A is observable and therefore it should belong to an external description. Suppose that the system is in an unknown state $x \in X$. If we want to extract as much information as possible from the system, then we could look at the values $f(x)$, $f(g(x))$, $f(g^2(x))$, $\dots$, which lie in A and are therefore observable. Thus the observable behaviour of this system is an infinite list of elements of A, which is called a *stream* over A.

The collection $A^{\mathbb{N}}$ of streams also carries an F-coalgebra structure, given by

$$\begin{aligned}
&\langle \mathrm{head}, \mathrm{tail} \rangle : A^{\mathbb{N}} \to A \times A^{\mathbb{N}} \\
&\mathrm{head}(a_0, a_1, a_2, \dots) = a_0 \\
&\mathrm{tail}(a_0, a_1, a_2, \dots) = (a_1, a_2, \dots).
\end{aligned}$$

This coalgebra of streams plays a special role among all F-coalgebras, because of the following fact.

Proposition 1.5. *For every F-coalgebra $\langle f, g \rangle : X \to A \times X$ there exists a unique coalgebra morphism from $\langle f, g \rangle$ to $\langle \mathrm{head}, \mathrm{tail} \rangle$. In other words, the coalgebra $\langle \mathrm{head}, \mathrm{tail} \rangle$ is a final object in the category $\mathbf{CoAlg}(F)$.*

Proof. We have to find a function $\varphi : X \to A^{\mathbb{N}}$ such that the diagram

$$\begin{tikzcd}
X \arrow[r, "\varphi"] \arrow[d, "\langle f,g \rangle"'] & A^{\mathbb{N}} \arrow[d, "\langle \mathrm{head},\mathrm{tail} \rangle"] \\
A \times X \arrow[r, "\mathrm{id}\times\varphi"'] & A \times A^{\mathbb{N}}
\end{tikzcd}$$

commutes. For the function φ we take

$$\varphi(x) = (f(x), f(g(x)), f(g^2(x)), \cdots),$$

i.e. the behaviour of the coalgebra when x is considered as initial state. This is a coalgebra morphism because

$$\begin{aligned}\langle \mathrm{head}, \mathrm{tail}\rangle(\varphi(x)) &= \langle \mathrm{head}, \mathrm{tail}\rangle(f(x), f(g(x)), f(g^2(x)), \ldots) \\ &= \langle f(x), (f(g(x)), f(g^2(x)), \ldots)\rangle \\ &= \langle f(x), \varphi(g(x))\rangle \\ &= (\mathrm{id} \times \varphi)(\langle f, g\rangle(x)).\end{aligned}$$

To prove uniqueness, suppose that $\psi : X \to A^{\mathbb{N}}$ is a function for which

$$\langle \mathrm{head}, \mathrm{tail}\rangle(\psi(x)) = (\mathrm{id} \times \psi)(\langle f, g\rangle(x))$$

for all $x \in X$. Denoting the nth entry of a stream σ by σ_n, we have to prove that $\varphi(x)_n = \psi(x)_n$ for all $x \in X$ and $n \in \mathbb{N}$. We will apply induction on n. For $n = 0$ the statement holds since

$$\varphi(x)_0 = \mathrm{head}(\varphi(x)) = f(x) = \mathrm{head}(\psi(x)) = \psi(x)_0.$$

Assume that the assertion holds for a number n. Then:

$$\varphi(x)_{n+1} = \mathrm{tail}(\varphi(x))_n = \varphi(g(x))_n = \psi(g(x))_n = \mathrm{tail}(\psi(x))_n = \psi(x)_{n+1},$$

where the third equality sign holds because of the induction hypothesis. Therefore the streams $\varphi(x)$ and $\psi(x)$ are equal. $\square$

This example suggests that a final object in a category of coalgebras is the collection of all possible behaviours, and the unique morphism into the final coalgebra assigns to a coalgebra its behaviour. It turns out that this is not only true for coalgebras for the functor $X \mapsto A \times X$, but for a large variety of coalgebras. This justifies the following terminology.

Definition 1.6. Let $F : \mathbf{C} \to \mathbf{C}$ be an endofunctor for which the category $\mathbf{CoAlg}(F)$ has a final object $\omega : \Omega \to F(\Omega)$. This coalgebra ω is called the *final coalgebra* for F. If $c : X \to F(X)$ is an arbitrary F-coalgebra, then the unique coalgebra morphism from c to ω is called the *behaviour morphism* and is denoted $\mathrm{beh}_c : X \to \Omega$.

In every category, final objects are unique up to isomorphism. Therefore we can speak of *the* final coalgebra instead of *a* final coalgebra.

We have already seen the role of the final coalgebra for the endofunctor $X \mapsto A \times X$. The next example shows that also in the case of automata we can think of the final coalgebra as the collection of possible behaviours.

Example 1.7. Automata over an alphabet A are coalgebras of the form $X \to F(X) = X^A \times 2$. The language recognized by an automaton is a reasonable candidate for its behaviour, so we will prove that the set $\mathcal{P}(A^*)$ of languages over A is the state space of the final F-coalgebra. First we have to endow $\mathcal{P}(A^*)$ with the structure of an F-coalgebra, that is, define a function $\omega = \langle \omega_1, \omega_2 \rangle : \mathcal{P}(A^*) \to \mathcal{P}(A^*)^A \times 2$. For the transition function $\omega_1 : \mathcal{P}(A^*) \to \mathcal{P}(A^*)^A$ we take the so-called Brzozowski derivative

$$\omega_1(L)(a) = \{u \in A^* \mid au \in L\}.$$

The output function $\omega_2 : \mathcal{P}(A^*) \to 2$ is given by

$$\omega_2(L) = 1 \text{ if and only if } \varepsilon \in L.$$

We claim that this coalgebra ω is final. To prove this, let $c = \langle \delta, \beta \rangle : X \to X^A \times 2$ be any F-coalgebra. Define a morphism $\mathrm{beh}_c : X \to \mathcal{P}(A^*)$ by

$$\mathrm{beh}_c(x) = \{u \in A^* \mid \beta(\delta^*(x)(u)) = 1\}.$$

This behaviour morphism assigns to a state $x \in X$ the language recognized by the automaton with x as initial state. To prove that beh_c is a coalgebra morphism, we have to show that $\omega_1(\mathrm{beh}_c(x))(a) = \mathrm{beh}_c(\delta(x)(a))$ and $\omega_2(\mathrm{beh}_c(x)) = \beta(x)$. The first claim holds because

$$\begin{aligned}
\omega_1(\mathrm{beh}_c(x))(a) &= \omega_1(\{u \in A^* \mid \beta(\delta^*(x)(u)) = 1\})(a) \\
&= \{u \in A^* \mid \beta(\delta^*(x)(au) = 1\} \\
&= \{u \in A^* \mid \beta(\delta^*(\delta(x)(a))(u)) = 1\} \\
&= \mathrm{beh}_c(\delta(x)(a)).
\end{aligned}$$

For the second claim, note that

$$\begin{aligned}
\omega_2(\mathrm{beh}_c(x)) = 1 &\iff \beta(\delta^*(x)(\varepsilon)) = 1 \\
&\iff \beta(x) = 1.
\end{aligned}$$

Now we will prove uniqueness of beh_c. Suppose that $\varphi : X \to \mathcal{P}(A^*)$ is any coalgebra morphism. We will prove by induction on $u \in A^*$ that $u \in \varphi(x)$ if and only if $u \in \mathrm{beh}_c(x)$ for all $x \in X$. For $u = \varepsilon$, this follows because

$$\begin{aligned}
\varepsilon \in \varphi(x) &\iff \omega_2(\varphi(x)) = 1 \\
&\iff \beta(x) = 1 \\
&\iff \omega_2(\mathrm{beh}_c(x)) = 1 \\
&\iff \varepsilon \in \mathrm{beh}_c(x).
\end{aligned}$$

Suppose that for each $x \in X$, the conditions $u \in \varphi(x)$ and $u \in \mathrm{beh}_c(x)$ are equivalent. Then, since φ is a coalgebra morphism, $au \in \varphi(x)$ if and only if $u \in \varphi(\delta(x)(a))$. By the induction hypothesis, this is true precisely if $u \in \mathrm{beh}_c(\delta(x)(a))$, which is equivalent to $au \in \mathrm{beh}_c(x)$. This proves that $\varphi = \mathrm{beh}_c$, and hence that ω is the final coalgebra.

The next result gives a general property of final coalgebras that is often useful. It can be used, for instance, to show that certain functors do not possess a final coalgebra.

Proposition 1.8 (Lambek). *Let $\omega : \Omega \to F(\Omega)$ be the final F-coalgebra for a functor $F : \mathbf{C} \to \mathbf{C}$. Then ω is an isomorphism in $\mathbf{C}$.*

Proof. Apply the functor F to the final coalgebra ω to obtain a new coalgebra $F(\omega) : F(\Omega) \to F(F(\Omega))$. By finality of ω there exists a unique coalgebra morphism $\varphi = \mathrm{beh}_{F(\omega)}$ from $F(\omega)$ to ω, which means that the following diagram commutes.

$$\begin{tikzcd}
F(\Omega) \arrow[r, "\varphi"] \arrow[d, "F(\omega)"'] & \Omega \arrow[d, "\omega"] \\
F(F(\Omega)) \arrow[r, "F(\varphi)"'] & F(\Omega)
\end{tikzcd}$$

We claim that φ is the inverse of ω. Since $\omega \circ \varphi \circ \omega = F(\varphi) \circ F(\omega) \circ \omega = F(\varphi \circ \omega) \circ \omega$, the diagram

$$\begin{tikzcd}
\Omega \arrow[r, "\varphi\circ\omega"] \arrow[d, "\omega"'] & \Omega \arrow[d, "\omega"] \\
F(\Omega) \arrow[r, "F(\varphi\circ\omega)"'] & F(\Omega)
\end{tikzcd}$$

also commutes. In other words, $\varphi \circ \omega$ is a coalgebra morphism from the final coalgebra to itself. But the only endomorphism of a final coalgebra is the identity, because of the uniqueness of behaviour maps. Therefore $\varphi \circ \omega = \mathrm{id}_\Omega$. From this result and the first commuting diagram it follows that the composition $\omega \circ \varphi$ also equals the identity:

$$\omega \circ \varphi = F(\varphi \circ \omega) = F(\mathrm{id}) = \mathrm{id}.$$

Thus ω is invertible and hence an isomorphism. □

Example 1.9. The power set functor $\mathcal{P}$ on **Sets** lacks a final coalgebra. If $\omega : \Omega \to \mathcal{P}(\Omega)$ would be a final coalgebra, then ω would be an isomorphism from a set onto its power set, which is impossible by Cantor's theorem.

1.3 Algebras and duality

Reversing the arrows in a category-theoretic concept leads to the dual concept, which is often equally interesting. This section discusses algebras, which are the categorical duals of coalgebras. In the literature, coalgebras are usually introduced as the duals of algebras, but here we take the reverse road, since our focus is on coalgebras. We will not develop the theory of algebras in much depth, but mainly concentrate on the connection between algebras and coalgebras via adjunctions.

Definition 1.10. An *algebra* for an endofunctor $F : \mathbf{C} \to \mathbf{C}$ consists of an object $X \in \mathbf{C}$ and a morphism $F(X) \to X$. A *homomorphism* from an algebra $a : F(X) \to X$ to $b : F(Y) \to Y$ is a morphism $f : X \to Y$ in $\mathbf{C}$ making the following diagram commute.

$$\begin{tikzcd} F(X) \arrow[r, "F(f)"] \arrow[d, "a"'] & F(Y) \arrow[d, "b"] \\ X \arrow[r, "f"'] & Y \end{tikzcd}$$

The category of F-algebras and homomorphisms in denoted $\mathbf{Alg}(F)$.

Example 1.11. Consider the functor $F(X) = 1 + A \times X$ on **Sets**, where A is a fixed alphabet. We write $*$ for the only element of the set 1. An F-algebra $b : 1 + A \times X \to X$ can be split into a function $1 \to X$, i.e. an element of X, and a function $A \times X \to X$. Algebras for the functor F can be viewed as an alternative way for describing automata. In this view, the set X is the state space, the element of X is the initial state, and the function $A \times X \to X$ is the dynamics. In the coalgebraic view we ignored the initial state. Dually, the algebraic view leaves out the subset of accepting states.

We have seen that a final object in a category of coalgebras is often interesting. Dually, an initial object in a category of algebras usually plays a special role. We claim that the initial algebra for the functor F is the set of words A^*, with algebra structure defined as follows.

$$\begin{aligned} \alpha : 1 + A \times A^* &\to A^* \\ * \qquad\qquad &\mapsto \varepsilon \\ (a, u) &\mapsto au. \end{aligned}$$

To establish initiality, let $b : 1 + A \times X \to X$ be an arbitrary F-algebra. Define a homomorphism $f : A^* \to X$ using induction by $f(\varepsilon) = b(*)$ and $f(au) = b(a, f(u))$. It is easy to see that f is a homomorphism of algebras and that it is unique. In the view of F-algebras as automata, the map $f : A^* \to X$ sends $u \in A^*$ to the state reached if we run the automaton with input u.

In Example 1.14 we will see how to phrase the connection between both perspectives on automata as an adjunction between F-algebras and coalgebras for the functor $X \mapsto X^A \times 2$.

The cornerstone of the results connecting algebras and coalgebras is the following property, which states that an adjunction between two categories can be lifted to an adjunction between categories of algebras.

Proposition 1.12. *Consider the situation*

$$\begin{tikzcd} \mathbf{C} \arrow[loop left, "F"] \arrow[r, bend left, "L"] \arrow[r, phantom, "\perp"] & \mathbf{D} \arrow[loop right, "G"] \arrow[l, bend left, "R"] \end{tikzcd}$$

Suppose that the functors LF and GL are naturally isomorphic. Then the adjunction $L \dashv R$ lifts to an adjunction

$$\begin{tikzcd} \mathbf{Alg}(F) \arrow[r, bend left, "\overline{L}"] \arrow[r, phantom, "\perp"] \arrow[d] & \mathbf{Alg}(G) \arrow[l, bend left, "\overline{R}"] \arrow[d] \\ \mathbf{C} \arrow[loop left, "F"] \arrow[r, bend left, "L"] \arrow[r, phantom, "\perp"] & \mathbf{D} \arrow[loop right, "G"] \arrow[l, bend left, "R"] \end{tikzcd}$$

Proof. Denote the natural isomorphism $GL \Rightarrow LF$ by τ. The functor $\overline{L}$: $\mathbf{Alg}(F) \to \mathbf{Alg}(G)$ is maps an algebra $a : F(X) \to X$ to

$$(\, GL(X) \xrightarrow[\cong]{\tau} LF(X) \xrightarrow{L(a)} L(X)\,).$$

On morphisms, $\overline{L}$ acts the same as L. To prove that this is well-defined, we have to show that $L(f)$ is a homomorphism whenever f is. If f is a homomorphism from a to a', then $a' \circ F(f) = f \circ a$, so $L(f) \circ L(a) \circ \tau = L(a') \circ LF(f) \circ \tau$. By naturality of τ, this equals $L(a') \circ \tau \circ GL(f)$. This proves that $L(f)$ is a homomorphism from $L(a)$ to $L(a')$.

To define the functor $\overline{R}$, we first have to define a natural transformation $\sigma : FR \Rightarrow RG$. Consider the natural transformation

$$(LFR \overset{\cong}{\Longrightarrow} GLR \overset{G\varepsilon}{\Longrightarrow} G),$$

where ε is the counit of the adjunction $L \dashv R$. Take the adjoint transpose to obtain the desired transformation σ. This enables us to define $\overline{R}$ on an algebra $b : G(Y) \to Y$ as $(FR(Y) \xrightarrow{\sigma} RG(Y) \xrightarrow{R(b)} R(Y))$. The functor $\overline{R}$ maps a homomorphism g to $R(g)$.

Proving that $\overline{L}$ is left adjoint to $\overline{R}$ amounts to constituting a bijective correspondence between diagrams of the following two forms.

$$\begin{tikzcd} LF(X) \arrow[r, "G(f)\circ\tau^{-1}"] \arrow[d, "L(a)"'] & G(Y) \arrow[d, "b"] \\ L(X) \arrow[r, "f"'] & Y \end{tikzcd} \qquad \begin{tikzcd} F(X) \arrow[r, "\sigma\circ F(g)"] \arrow[d, "a"'] & RG(Y) \arrow[d, "R(b)"] \\ X \arrow[r, "g"'] & R(Y) \end{tikzcd}$$

This correspondence is already contained in the adjunction between L and R. It remains to be verified that g is a homomorphism whenever f is and vice versa. As the proofs of both assertions are similar, we will only show that if f is a homomorphism, then so is g. Suppose that $b \circ G(f) \circ \tau^{-1} = f \circ L(a)$. Then

$$\begin{aligned} R(b) \circ \sigma \circ F(g) &= R(b \circ G\varepsilon \circ \tau^{-1}) \circ \eta_{FR} \circ FR(f) \circ F(\eta) \\ &= R(b \circ G\varepsilon \circ \tau^{-1}) \circ RLFR(f) \circ RLF(\eta) \circ \eta_F \\ &= R(b \circ G\varepsilon \circ \tau^{-1} \circ LF(R(f) \circ \eta)) \circ \eta_F \\ &= R(b \circ G\varepsilon \circ GL(R(f) \circ \eta) \circ \tau^{-1}) \circ \eta_F \\ &= R(b \circ G(f \circ \varepsilon \circ L(\eta)) \circ \tau^{-1}) \circ \eta_F \\ &= R(b \circ G(f) \circ \tau^{-1}) \circ \eta_F \\ &= R(f \circ L(a)) \circ \eta_F \\ &= R(f) \circ \eta \circ a \\ &= g \circ a. \end{aligned}$$

□

By duality, it is also possible to lift adjunctions to categories of coalgebras. The lifting results are often applied to dual adjunctions. In that case we obtain two dual adjunctions between algebras and coalgebras, using that $\mathbf{Alg}(F^{\mathrm{op}}) \cong \mathbf{CoAlg}(F)^{\mathrm{op}}$.

Corollary 1.13. *Assume a dual adjunction*

$$\begin{tikzcd} \mathbf{C} \arrow[loop left, "F"] \arrow[r, bend left, "L"] \arrow[r, phantom, "\perp"] & \mathbf{D}^{\mathrm{op}} \arrow[l, bend left, "R"] \arrow[loop right, "G^{\mathrm{op}}"] \end{tikzcd}$$

This dual adjunction gives rise to two adjunctions between algebras and coalgebras:

1. *If $LF \cong G^{\mathrm{op}}L$, then there is an adjunction*

$$\begin{tikzcd} \mathbf{Alg}(F) \arrow[r, bend left] \arrow[r, phantom, "\perp"] & \mathbf{CoAlg}(G)^{\mathrm{op}} \arrow[l, bend left] \end{tikzcd}$$

2. *If $RG^{\mathrm{op}} \cong RF$, then there is an adjunction*

$$\begin{tikzcd} \mathbf{CoAlg}(F) \arrow[r, bend left] \arrow[r, phantom, "\perp"] & \mathbf{Alg}(G)^{\mathrm{op}} \arrow[l, bend left] \end{tikzcd}$$

In particular the initial algebra and final coalgebra are related. Consider the adjunction between $\mathbf{Alg}(F)$ and $\mathbf{CoAlg}(G)^{\mathrm{op}}$. Since left adjoints preserve initial objects, an initial F-algebra yields an initial object in $\mathbf{CoAlg}(G)^{\mathrm{op}}$, which is a final G-coalgebra. Analogously, an initial G-algebra gives a final F-coalgebra, because right adjoints preserve final objects.

Example 1.14. We will apply the duality between algebras and coalgebras to connect words and automata. It will only work for a special case of automata. Our base adjunction will be a special case of Stone duality, namely the duality between sets and Boolean algebras. Denote the category of Boolean algebras by **BA**. Then the contravariant power set functor $\mathcal{P}$: **Sets** $\to$ $\mathbf{BA}^{\mathrm{op}}$ and the functor $\mathcal{UF} : \mathbf{BA}^{\mathrm{op}} \to \mathbf{Sets}$, assigning to a Boolean algebra the set of its ultrafilters, constitute a dual adjunction:

$$\begin{tikzcd} \mathbf{Sets} \arrow[r, bend left, "\mathcal{P}"] \arrow[r, phantom, "\perp"] & \mathbf{BA}^{\mathrm{op}} \arrow[l, bend left, "\mathcal{UF}"] \end{tikzcd}$$

We will now define the functors F and G to establish the situation in Corollary 1.13. Fix an alphabet A. Pick the functor $F(X) = 1 + A \times X$ on **Sets**, whose algebras were discussed in Example 1.11. On the other side of the adjunction we define the functor $G(X) = 2 \times X^A$ on **BA**. We will call coalgebras for this functor *Boolean automata*. Spelling out the definition, a Boolean automaton consists of a Boolean algebra X equipped with dynamics $\delta : X \to X^A$ and a map $\beta : X \to 2$ indicating the accepting states. The maps δ and β are both homomorphisms of Boolean algebras. The condition $\mathcal{P}F \cong G\mathcal{P}$ is satisfied, since $\mathcal{P}(1 + A \times X) \cong 2 \times \mathcal{P}(X)^A$.

By Corollary 1.13, the dual adjunction lifts to an adjunction between $\mathbf{Alg}(F)$ and $\mathbf{CoAlg}(G)$. For instance, we describe how to translate an F-algebra $b : 1 + A \times X \to X$ into a G-coalgebra. Write $x_0 = b(*)$ and $b(a, x) = b_a(x)$, so that b_a is a function from X to X for each $a \in A$. The corresponding G-coalgebra is obtained by applying the power set functor to this algebra. The resulting dynamics map $\delta : \mathcal{P}(X) \to \mathcal{P}(X)^A$ maps U to $(b_a^{-1}(U))_{a \in A}$. The output map $\beta : \mathcal{P}(X) \to 2$ maps U to 1 if and only if $x_0 \in U$. Thus an automaton considered as F-algebra b gives a Boolean automaton with output function obtained from the initial state of b and whose dynamics is a "reversed" version of the dynamics of b.

Since the initial F-algebra is A^* and the functor $\mathcal{P}$ maps the initial F-algebra to the final G-coalgebra, we see that $\mathcal{P}(A^*)$ is the state space of the final G-coalgebra. Thus the behaviour of ordinary automata and Boolean automata can be represented using the same final coalgebra. For the F-algebra b we obtain a map $f : A^* \to X$ by initiality. The behaviour map of the corresponding coalgebra is $\mathcal{P}(f) : \mathcal{P}(X) \to \mathcal{P}(A^*)$ given by $\mathcal{P}(f)(U) = \{u \in A^* \mid f(u) \in U\}$.

1.4 Factorization

In the category **Sets**, every morphism can be factored as an epimorphism followed by a monomorphism. In this section, we will consider the more general situation of categories equipped with two classes of morphisms enjoying properties similar to the factorization structure on **Sets**. Factorization of coalgebra morphisms will be particularly interesting, since this is fundamental for the minimization of coalgebras discussed in Section 1.5.

Definition 1.15. Let $\mathbf{C}$ be a category and let E and M be classes of morphisms in $\mathbf{C}$. In diagrams, we will write morphisms in E as arrows $\twoheadrightarrow$, and morphisms in M as arrows $\rightarrowtail$. The pair (E, M) is called a *factorization system* on $\mathbf{C}$ if the following conditions hold.

- Both E and M are closed under composition and contain all isomorphisms.
- Each morphism $f : X \to Z$ can be factored as

  $$\begin{tikzcd} X \arrow[r, two heads, "e"'] \arrow[rr, bend left, "f"] & Y \arrow[r, tail, "m"'] & Z \end{tikzcd},$$

  i.e. $f = m \circ e$ with $m \in M$ and $e \in E$.
- The diagonalization property: for each commutative square

  $$\begin{tikzcd} W \arrow[r, two heads] \arrow[d] & X \arrow[d] \\ Y \arrow[r, tail] & Z \end{tikzcd}$$

  there exists a unique diagonalization morphism $X \to Y$ making both triangles in the following diagram commute.

  $$\begin{tikzcd} W \arrow[r, two heads] \arrow[d] & X \arrow[d] \arrow[dl, dotted] \\ Y \arrow[r, tail] & Z \end{tikzcd}$$

The most common example of a factorization system is given by epis and monos in **Sets**. A function $f : X \to Y$ between two sets can be factored through its image as $X \twoheadrightarrow \mathrm{im}(f) \rightarrowtail Y$.

For the minimization procedure in Section 1.5 it will be useful to have a method to produce factorization systems on categories of coalgebras. We will give such a method by showing that coalgebra factorizations can be obtained by lifting the factorization system on the underlying category.

Proposition 1.16. *Let (E, M) be a factorization system on a category $\mathbf{C}$. Furthermore assume an endofunctor F on $\mathbf{C}$ preserving morphisms in M. Then the coalgebra morphisms in E and M form a factorization system on the category of F-coalgebras.*

Proof. The categories $\mathbf{C}$ and $\mathbf{CoAlg}(F)$ have the same isomorphisms, so also in $\mathbf{CoAlg}(F)$ both classes of morphisms contain all isos. Furthermore the composition is the same in $\mathbf{C}$ and $\mathbf{CoAlg}(F)$, hence E and M are closed under composition in $\mathbf{CoAlg}(F)$.

Take a coalgebra morphism $f : X \to Z$. Factorize f as $f = m \circ e$ in the underlying category $\mathbf{C}$. We get a diagram of the form

$$\begin{tikzcd}
X \arrow[r, "e"] \arrow[d] & Y \arrow[r, tail, "m"] \arrow[d, dotted, "d"] & Z \arrow[d] \\
F(X) \arrow[r, "F(e)"'] & F(Y) \arrow[r, tail, "F(m)"'] & F(Z)
\end{tikzcd}$$

The morphism $F(m)$ is in M by hypothesis. The morphism d has been obtained by diagonalization, and provides a coalgebra structure on Y. This gives a factorization of f as coalgebra morphism.

To prove the diagonalization property, take a commutative square in the category $\mathbf{CoAlg}(F)$:

$$\begin{tikzcd}
FW \arrow[rr] \arrow[dd] & & FX \arrow[dd] \arrow[ddll, dotted, "F(d)"] & \\
& W \arrow[ul, "c_W"] \arrow[rr, "e"] \arrow[dd, "f"] & & X \arrow[ul, "c_X"'] \arrow[dd, "g"] \arrow[ddll, dotted, "d"] \\
FY \arrow[rr, tail] & & FZ & \\
& Y \arrow[ul, "c_Y"] \arrow[rr, tail, "m"'] & & Z \arrow[ul, "c_Z"']
\end{tikzcd}$$

The morphism d arises from the diagonalization property in the underlying category. To prove that everything in the cube commutes, it remains to be shown that d is a coalgebra morphism, i.e. $c_Y \circ d = F(d) \circ c_X$. Consider the following square in $\mathbf{C}$.

$$\begin{tikzcd}
W \arrow[r, "e"] \arrow[d, "c_Y \circ f"'] & X \arrow[d, "c_Z \circ g"] \\
FY \arrow[r, tail, "F(m)"'] & FZ
\end{tikzcd}$$

We will show that $c_Y \circ d$ and $F(d) \circ c_X$ are both diagonalizations for this square. This follows from four calculations:

$$c_Y \circ d \circ e = c_Y \circ f$$

$$F(m) \circ c_Y \circ d = c_Z \circ m \circ d = c_Z \circ g$$

$$F(d) \circ c_X \circ e = F(d) \circ F(e) \circ c_W = F(d \circ e) \circ c_W = F(f) \circ c_W = c_Y \circ f$$

$$F(m) \circ F(d) \circ c_X = F(m \circ d) \circ c_X = F(g) \circ c_X$$

From uniqueness of diagonalizations it follows that $c_Y \circ d = F(d) \circ c_X$. $\square$

Remark. As proven in [4], every endofunctor on **Sets** can be replaced by a mono-preserving endofunctor with the same values on non-empty sets and the same final coalgebra. Thus the condition in Proposition 1.16 that F preserves monos imposes no restrictions when working in the category **Sets**, and consequently there is a factorization system on all categories of coalgebras over **Sets**.

1.5 Minimization

In Section 1.2 we discussed how to obtain an external description of a system, given an internal description. We will now turn our attention to the reverse problem: if we know the behaviour of a system, how do we find a coalgebra having that behaviour? Of course, in practice there are many coalgebras with the same behaviour. We are often interested in the most efficient one, i.e. the coalgebra with the smallest state space among those with the same behaviour. Finding the coalgebra with a minimal state space is known as the problem of minimization.

Minimal realizations of automata were first constructed in [26]. This was generalized to categorical settings in [13], see also [3, 29] for the coalgebraic version.

Throughout this section, we fix a category $\mathbf{C}$ with a factorization system (E, M). Furthermore we stipulate that an endofunctor F on $\mathbf{C}$ is given that preserves morphisms in M. By Proposition 1.16 this yields a factorization system on the category of F-coalgebras.

We will first define the reduction of a coalgebra. For this we follow [3], where it is called minimization. Reduction amounts to merging states which are externally indistinguishable. Thus it gives a coalgebra which is smaller, but has the same behaviour.

Definition 1.17. The *reduction* of a coalgebra $c : X \to F(X)$ in $\mathbf{C}$ is a coalgebra $Z \to F(Z)$ together with a coalgebra morphism $e : X \twoheadrightarrow Z$ such that for every $e' : X \twoheadrightarrow Y$ there is a necessarily unique $f : Y \to Z$ such that $e = f \circ e'$. The coalgebra c is called *reduced* if e is an isomorphism from c onto its reduction.

The term "reduced" comes from automata theory, see for example [11]. The above definition is a generalization to arbitrary coalgebras. Note that it depends on the choice of factorization system.

Theorem 1.18. *Suppose that the functor F has a final coalgebra Ω. Then the reduction of any coalgebra $c : X \to F(X)$ is obtained by factorizing the* *behaviour morphism* $\mathrm{beh}_c : X \to \Omega$*:*

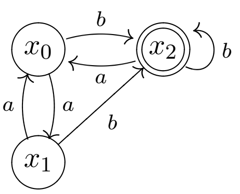
$$\begin{tikzcd}
X \arrow[r, "e"'] \arrow[rr, bend left, "\mathrm{beh}_c"] \arrow[d, "c"'] & Z \arrow[r, "m"'] \arrow[d, "c_{\mathrm{red}}"] & \Omega \arrow[d, "\omega"] \\
F(X) \arrow[r, "F(e)"] \arrow[rr, bend right, "F(\mathrm{beh}_c)"'] & F(Z) \arrow[r, "F(e)"] & F(\Omega)
\end{tikzcd}$$

Proof. We will prove the defining property of the reduction for c_{red}. Assume that we have a coalgebra morphism $e' : X \twoheadrightarrow Y$ from c to a coalgebra $c' : Y \to F(Y)$. Form the square

$$\begin{tikzcd}
X \arrow[r, "e'"] \arrow[d, "e"'] & Y \arrow[d, "\mathrm{beh}_{c'}"] \\
Z \arrow[r, "m"'] & \Omega
\end{tikzcd}$$

of coalgebra morphisms, which commutes by finality of ω. By diagonalization we obtain the desired map $f : Y \to Z$. □

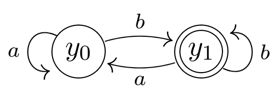
Example 1.19. We illustrate this theorem by applying it to automata. We use the ordinary factorization system consisting of epis and monos on **Sets**. Since the functor $X \mapsto X^A \times 2$ preserves monos, the factorization system lifts to the category of automata. We wish to compute the reduction of the following automaton c.



The behaviour morphism is given by

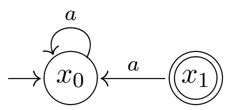
$$\begin{aligned}
\mathrm{beh}_c : x_0 &\mapsto \{ub \mid u \in \{a,b\}^*\} \\
x_1 &\mapsto \{ub \mid u \in \{a,b\}^*\} \\
x_2 &\mapsto \{ub \mid u \in \{a,b\}^*\} \cup \{\varepsilon\}
\end{aligned}$$

The state space of the reduction is the image of beh_c, which consists of two states. The quotient map from $\{x_0, x_1, x_2\}$ onto the reduction identifies the states x_0 and x_1. Thus the reduction is



When we ignore initial states, as is often done in the theory of coalgebras, reduction is a good notion of minimization. However, when looking for minimal realisation of behaviour, one sometimes wants to take the initial state into account. In that case, a reduced coalgebra need not be the one with a minimal state space, as the following example shows.

Example 1.20. Consider the following automaton over the singleton alphabet $\{a\}$ with an initial state.



This automaton is reduced, but since the state x_1 cannot be reached from the initial state, it can be removed to get a smaller automaton that recognizes the same language. Thus the above automaton is not minimal.

To define minimality, we need an additional condition that ensures that all states can be reached from the initial state. First we will need a notion of initial state in an arbitrary category. In **Sets**, an initial state of a coalgebra $c : X \to F(X)$ is an element of its state space X. Equivalently, it is a morphism $1 \to X$. For a coalgebra on an arbitrary category in $\mathbf{C}$, an initial state is a generalized element of the state space X, i.e. a morphism $b : B \to X$ for some fixed object $B \in \mathbf{C}$. In particular, if the functor F has a final coalgebra $\omega : \Omega \to F(\Omega)$, then a behaviour is a morphism $b : B \to \Omega$.

Definition 1.21. Suppose that the functor F has a final coalgebra $\omega : \Omega \to F(\Omega)$. We say that a coalgebra $c : X \to F(X)$ with initial state $b : B \to X$ *has behaviour* $b' : B \to \Omega$ if the following diagram commutes.

$$\begin{tikzcd} B \arrow[r, "b"] \arrow[dr, "b'"'] & X \arrow[d, "\mathrm{beh}_c"] \\ & \Omega \end{tikzcd}$$

Definition 1.22. A *subcoalgebra* of a coalgebra $c : X \to F(X)$ is a coalgebra $s : S \to F(S)$ together with a coalgebra morphism $S \rightarrowtail X$ in the class M.

Definition 1.23. Let $c : X \to F(X)$ be a coalgebra and $b : B \to X$ a morphism. The subcoalgebra of c *generated* by b is a subcoalgebra $\langle b \rangle : S \to F(S)$ of c together with a morphism $b' : B \to S$, such that the following two conditions hold.

- The following diagram commutes.

$$\begin{tikzcd} B \arrow[r, "b'"] \arrow[dr, "b"'] & S \arrow[d, tail] \\ & X \end{tikzcd}$$

- For every subcoalgebra $d : T \to F(T)$ of c together with $b'' : B \to T$ there exists a unique coalgebra morphism $S \rightarrowtail T$ making the following diagram commute.

$$\begin{tikzcd}
B \arrow[r, "b'"] \arrow[dr, "b''"'] & S \arrow[d, tail] \arrow[dr, tail] & \\
& T \arrow[r, tail] & X
\end{tikzcd}$$

The pair (c, b) is said to be *accessible* if the inclusion morphism from the subcoalgebra of c generated by b into c is an isomorphism.

Example 1.24. Consider an automaton $c = \langle \delta, \beta \rangle : X \to X^A \times 2$. A subset $S \subseteq X$ can be identified with a monomorphism $b : S \to X$. The subcoalgebra $\langle b \rangle$ of c generated by b is obtained by closing S under the transitions of c. Thus the coalgebra $\langle b \rangle$ has state space

$$\{\delta(s)(u) \mid s \in S, u \in A^*\}.$$

The observations and transitions are inherited from c.

Consider the special case where S is the singleton set consisting of the initial state x_0 of the automaton. Then the pair $(c, \{x_0\})$ is accessible if and only if the map $\delta(x_0) : A^* \to X$ is surjective. This means that every state of the automaton can be reached from the initial state by running the automaton with a suitable input.

Definition 1.25. A coalgebra $c : X \to F(X)$ together with an initial state $b : B \to X$ is called *minimal* if c is reduced and (c, b) is accessible.

With this concept of minimality, we can answer the question posed at the beginning of this section. The question can be phrased as follows: given a behaviour $b : B \to \Omega$, what is the minimal coalgebra having this behaviour?

Theorem 1.26. *Suppose that the functor F has a final coalgebra $\omega : \Omega \to F(\Omega)$. Let $b : B \to \Omega$ be a behaviour. The minimal coalgebra with behaviour b is the subcoalgebra of ω generated by b.*

Proof. We start by proving that the subcoalgebra $\langle b \rangle : S \to F(S)$ with initial state $b' : B \to S$ has behaviour b. The inclusion $m : S \rightarrowtail \Omega$ is a coalgebra morphism and satisfies $m \circ b' = b$. It is unique by finality of ω, so it must be the behaviour morphism.

The subcoalgebra of $\langle b \rangle$ generated by b' is $\langle b \rangle$ itself, so $\langle b \rangle$ is accessible. It remains to compute the reduction of $\langle b \rangle$. According to Theorem 1.18 it is the factorization of $\mathrm{beh}_{\langle b \rangle} : S \to \Omega$. But this behaviour map is already in M, so the coalgebra $\langle b \rangle$ is isomorphic to its own reduction. This establishes minimality. $\square$

Example 1.27. We will use the above result to calculate the minimal automaton recognizing a language. If L is a language over an alphabet A, then it can be seen as a state in the final coalgebra $\mathcal{P}(A^*)$ of the functor $X \mapsto X^A \times 2$. The minimal automaton is the subcoalgebra of $\mathcal{P}(A^*)$ generated by the language L. Since the coalgebra structure of $\mathcal{P}(A^*)$ is given by Brzozowski derivatives, the generated subcoalgebra assumes the form

$$\langle L \rangle = \{L_u \mid u \in A^*\}.$$

As an example, let L be the language $\{b\} \cup \{ab^n \mid n \geq 0\}$ over $\{a, b\}$. The Brzozowski derivatives are

$$L_\varepsilon = L$$

$$L_a = \{b^n \mid n \geq 0\}$$

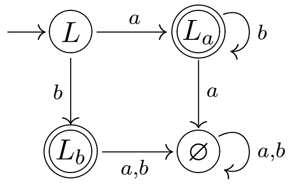
$$L_b = \{\varepsilon\}$$

$$L_{aa} = \varnothing$$

$$L_{ab} = L_a$$

$$L_{ba} = L_{bb} = \varnothing$$

With the transitions $\delta(K)(c) = K_c$ and $\beta(K) = 1$ if and only if $\varepsilon \in K$ this gives the following minimal automaton for L.



1.6 Probabilistic systems

All systems encountered in the previous sections were deterministic: given an input, there is only one possible time evolution of the system. We will now study probabilistic systems, in which each state has several successors, and one of these is chosen according to a probability distribution. This section shows how to model these systems as coalgebras. Some of the results for probabilistic coalgebras will also be useful for quantum systems, as these also behave probabilistically.

Transitions to a probability distribution over successor states will be modeled using functors involving the so-called *distribution monad*. Define a functor $\mathcal{D} : \mathbf{Sets} \to \mathbf{Sets}$ sending a set X to the set of finite convex combinations, or probability distributions, on X:

$$\mathcal{D}(X) = \{\varphi : X \to [0,1] \mid \mathrm{supp}(\varphi) \text{ is finite and } \sum_{x \in X} \varphi(x) = 1\}.$$

Here $\mathrm{supp}(\varphi) = \{x \in X \mid \varphi(x) \neq 0\}$ is the *support* of φ. An element φ of $\mathcal{D}(X)$ can also be written as a formal sum $r_1x_1 + \cdots + r_nx_n$, where $\{x_1, \ldots, x_n\} = \mathrm{supp}(\varphi)$ and $r_i = \varphi(x_i)$. On a morphism $f : X \to Y$, the functor $\mathcal{D}$ is defined as

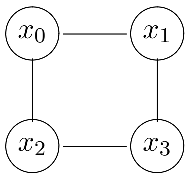
$$\mathcal{D}(f)(r_1x_1 + \cdots + r_nx_n) = r_1f(x_1) + \cdots + r_nf(x_n).$$

The functor $\mathcal{D}$ is a monad with unit and multiplication

$$\begin{aligned} \eta : X &\to \mathcal{D}(X) & \qquad \mu : \mathcal{D}(\mathcal{D}(X)) &\to \mathcal{D}(X) \\ x &\mapsto 1x & \sum_i r_i(\sum_j s_{ij}x_{ij}) &\mapsto \sum_{i,j} r_i s_{ij} x_{ij} \end{aligned}$$

We will show how the distribution monad is used in the representation of probabilistic systems by an example.

Example 1.28. Imagine a particle moving on the following graph.



Let $X = \{x_0, x_1, x_2, x_3\}$ be the set of vertices. The particle starts at one of the vertices of the graph. In each time step, the particle can move to one of the two adjacent points. Each of these points is chosen with probability $\frac{1}{2}$. This system can be written as a coalgebra for the distribution monad:

$$\begin{aligned} c : X &\to \mathcal{D}(X) \\ x_0 &\mapsto \tfrac{1}{2}x_1 + \tfrac{1}{2}x_2 \\ x_1 &\mapsto \tfrac{1}{2}x_0 + \tfrac{1}{2}x_3 \\ x_2 &\mapsto \tfrac{1}{2}x_0 + \tfrac{1}{2}x_3 \\ x_3 &\mapsto \tfrac{1}{2}x_1 + \tfrac{1}{2}x_2 \end{aligned}$$

A coalgebra for $\mathcal{D}$ is called a *Markov chain.*

We consider the trajectory of the particle when it starts in the vertex x_0. Let $\varphi_n \in \mathcal{D}(X)$ denote the probability distribution over the vertices of the graph after n steps. Then the first few values of φ_n are:

$$\begin{aligned} \varphi_0 &= x_0 \\ \varphi_1 &= \tfrac{1}{2}x_1 + \tfrac{1}{2}x_2 \\ \varphi_2 &= \tfrac{1}{2}x_0 + \tfrac{1}{2}x_3 \\ \varphi_3 &= \tfrac{1}{2}x_1 + \tfrac{1}{2}x_2 \end{aligned}$$

We obtain a repetition after three steps, so $\varphi_{2n} = \varphi_2$ for $n > 0$ and $\varphi_{2n+1} = \varphi_1$ for all $n \in \mathbb{N}$.

It is reasonable to view the sequence $(\varphi_n)_{n\in\mathbb{N}}$ as the behaviour of the Markov chain. However, if we model the system as a $\mathcal{D}$-coalgebra, the morphism into the final coalgebra does not give the desired behaviour. This can be solved by replacing the underlying category **Sets** of the coalgebra by another category. There are at least two possible replacements known in the literature: [16] proposes to work in the Kleisli category of the monad $\mathcal{D}$, and [31] proposes to work in the category $\mathbf{EM\text{-}Alg}(\mathcal{D})$ of Eilenberg-Moore algebras for $\mathcal{D}$. The solutions work for coalgebras for several monads, not only for the distribution monad. Both approaches are compared in [21]. Here we will briefly describe how to model probabilistic systems in $\mathbf{EM\text{-}Alg}(\mathcal{D})$, since this approach will also be used for quantum systems.

An Eilenberg-Moore algebra for the distribution monad is called a convex set. In a convex set X, we can assign to each convex combination $\sum_{i=1}^n r_i = 1$ a function $X^n \to X$ denoted $(x_1, \ldots, x_n) \mapsto \sum_{i=1}^n r_i x_i$. A homomorphism of convex sets preserves all convex combinations and is called a convex or affine map. The category of convex sets and maps is written as **Conv** instead of $\mathbf{EM\text{-}Alg}(\mathcal{D})$.

The next result ensures that several functors on the category **Conv** have a final coalgebra, which enables us to speak about the behaviour of probabilistic systems.

Lemma 1.29. *Let $\mathbf{C}$ be a category with all products. Fix an object $B \in \mathbf{C}$ and a set A. The functor $F : \mathbf{C} \to \mathbf{C}$ defined by $F(X) = B \times X^A$, where X^A denotes a power, has a final coalgebra.*

Proof. The underlying object of the final coalgebra is the power B^{A^*}. The projection $B^{A^*} \to B$ onto the coordinate with index $u \in A^*$ will be denoted π_u. The dynamics is a map $\omega : B^{A^*} \to B \times (B^{A^*})^A$, whose first component is π_ε and whose second component in the coordinates $a \in A$, $u \in A^*$ is π_{au}. We will now prove the finality of ω. Given a coalgebra $c = \langle f, \langle g_a \rangle_{a\in A} \rangle : X \to B \times X^A$, define $\mathrm{beh}_c : X \to B^{A^*}$ as follows: first extend the family of morphisms $(g_a)_{a\in A})$ for $a \in A$ to a family $(g_u)_{u\in A^*}$ by defining inductively

$$\begin{aligned} g_\varepsilon &= \mathrm{id}, \\ g_{au} &= g_u \circ g_a. \end{aligned}$$

Then let $\mathrm{beh}_c = \langle f \circ g_u \rangle_{u\in A^*}$. We verify that beh_c is a coalgebra morphism from c to ω.

$$\begin{aligned} \omega \circ \mathrm{beh}_c &= \langle \pi_\varepsilon, \langle \langle \pi_{au} \rangle_{u\in A^*} \rangle_{a\in A} \rangle \circ \langle f \circ g_u \rangle_{u\in A^*} \\ &= \langle f \circ g_\varepsilon, \langle \langle f \circ g_{au} \rangle_{u\in A^*} \rangle_{a\in A} \rangle \\ &= \langle f \circ \mathrm{id}, \langle \langle f \circ g_u \rangle_{u\in A^*} \circ g_a \rangle_{a\in A} \rangle \\ &= \langle f, \langle \mathrm{beh}_c \circ g_a \rangle_{a\in A} \rangle \\ &= (\mathrm{id} \times \mathrm{beh}_c^A) \circ c. \end{aligned}$$

It remains to be shown that beh_c is the unique such coalgebra morphism. Let $\varphi : X \to B^{A^*}$ be any coalgebra morphism, and denote the component with coordinate $u \in A^*$ by φ_u. We will prove by induction on the word $u \in A^*$ that $\varphi_u = f \circ g_u$. First suppose that $u = \varepsilon$. Since φ is a coalgebra morphism,

$$(\mathrm{id} \times \varphi^A) \circ c = \omega \circ \varphi. \tag{1.1}$$

Taking the first component of this equation gives

$$f = \pi_1 \circ \omega \circ \varphi = \pi_\varepsilon \circ \varphi = \varphi_\varepsilon.$$

This proves the assertion for the empty word ε. Now assume that $\varphi_u = f \circ g_u$. We have to show that $\varphi_{au} = f \circ g_{au}$. Take the second component of (1.1) to obtain

$$\varphi^A \circ g = \langle \langle \pi_{au} \rangle_{u \in A^*} \rangle_{a \in A} \circ \varphi = \langle \langle \varphi_{au} \rangle_u \rangle_a.$$

Pre-composing both sides with $\pi_u \circ \pi_a$ results in

$$\varphi_{au} = \pi_u \circ \pi_a \circ \varphi^A \circ g = \pi_u \circ \varphi \circ g_a = \varphi_u \circ g_a = f \circ g_u \circ g_a = f \circ g_{au}.$$

Therefore the claim holds for all words u. □

Example 1.30. The coalgebra from Example 1.28 can also be represented as a coalgebra in **Conv**, in such a way that the morphism into the final coalgebra yields the list of probability distributions associated to the Markov chain. Fix a set X and take the functor $F(Y) = \mathcal{D}(X) \times Y$ on the category **Conv**, where X is the set of vertices of the graph. Represent the Markov chain as an F-coalgebra d with state space $\mathcal{D}(X)$, i.e. a convex map $\mathcal{D}(X) \to \mathcal{D}(X) \times \mathcal{D}(X)$. Since $\mathcal{D}(X)$ is the free convex set on X, it suffices to define d on the set X of generators. Let $d(x) = (1x, c(x)) \in \mathcal{D}(X) \times \mathcal{D}(X)$. From the proof of Lemma 1.29 we see that the final F-coalgebra exists and has state space $\mathcal{D}(X)^{\mathbb{N}}$, so we obtain a map $\mathrm{beh}_d : \mathcal{D}(X) \to \mathcal{D}(X)^{\mathbb{N}}$. This behaviour map sends the initial state $x_0 \in \mathcal{D}(X)$ to the list of probability distributions $(\varphi_n)_{n \in \mathbb{N}}$ obtained in Example 1.28.

We can apply this coalgebraic model to minimize probabilistic systems. For this we need a factorization system on the category of convex sets.

Proposition 1.31. *Suppose that the category* $\mathbf{C}$ *carries a factorization system* (E, M) *and* $T : \mathbf{C} \to \mathbf{C}$ *is a monad preserving morphisms in* E*. Then* (E, M) *is also a factorization system on* $\mathbf{EM\text{-}Alg}(T)$*.*

Proof. Analogous to the proof of Proposition 1.16. □

Remark. All epis in **Sets** are split, so they are preserved by all functors. Thus if T is a monad on **Sets**, then morphisms in $\mathbf{EM\text{-}Alg}(T)$ can always be factored through their image. In particular there is a factorization system on the category **Conv**.

Corollary 1.32. *Let A be a set and B a convex set.*

1. *Let $c : X \to B \times X^A$ be a coalgebra in* **Conv**. *The reduction of c exists and equals the image of the behaviour morphism* $\mathrm{beh}_c : X \to B^{A^*}$.
2. *Each behaviour $1 \to B^{A^*}$ has a minimal realization, obtained as subcoalgebra of B^{A^*}.*

Chapter 2

Quantum computation

Quantum computation is a type of computation in which the data and the operations are subject to the laws of quantum mechanics. An overview of this rapidly growing field can be found in [28]. This chapter contains the basics of quantum computation. To pave the way for the coalgebraic description in Chapter 3, we will formulate some of the results using category theory.

Section 2.1 contains some mathematical preliminaries. In Section 2.2 we present the standard formulation of quantum mechanics in terms of states and projections on Hilbert spaces. This is generalized in Sections 2.3 and 2.4 to a formulation in terms of density matrices as states and effects as observables. We will discuss the duality between these two notions in Section 2.5.

2.1 Hilbert spaces

The mathematical formulation of quantum physics is based on Hilbert spaces. This section contains some of the mathematical preliminaries on Hilbert spaces that is needed to understand quantum mechanics and quantum computation. The material in this section can be found in any book on functional analysis, for instance [23, 33]. As the results in this section are well-known, we will not prove most of them and refer to the above books instead.

Definition 2.1. A *inner product space* is a complex vector space H equipped with an inner product, that is, a map $\langle - | - \rangle : H \times H \to \mathbb{C}$ subject to the following conditions.

- Sesquilinearity: $\langle x | \alpha y + \beta z \rangle = \alpha \langle x | y \rangle + \beta \langle x | z \rangle$ for all $x, y, z \in H$ and $\alpha, \beta \in \mathbb{C}$.
- Skew-symmetry: $\langle x | y \rangle = \overline{\langle y | x \rangle}$ for all $x, y \in H$.
- Non-degeneracy: $\langle x | x \rangle \geq 0$ with equality if and only if $x = 0$.

The elements of H are called *points* or *vectors*. The inner product gives rise to a norm $\|-\| : H \to \mathbb{R}$ defined by $\|x\| = \sqrt{\langle x|x\rangle}$. In turn, the norm induces a metric given by $d(x,y) = \|x-y\|$. The space H is a *Hilbert space* if it is complete in the induced metric, that is, every Cauchy sequence in H converges to a point in H.

Hilbert spaces are the objects of a category **Hilb**, in which the morphisms are linear maps that are continuous with respect to the metrics. These morphisms are called *operators*.

In general, Hilbert spaces can also be defined over the field of real numbers instead of complex numbers, but for applications to quantum mechanics, only complex spaces are relevant.

Example 2.2. An easy example of a Hilbert space is the space

$$\mathbb{C}^n = \{(x_1,\ldots,x_n) \mid x_i \in \mathbb{C}\}$$

of finite sequences of complex numbers endowed with the inner product

$$\langle (x_1,\ldots,x_n)|(y_1,\ldots,y_n)\rangle = \overline{x_1}y_1 + \cdots + \overline{x_n}y_n.$$

This inner product induces a metric given by $d((x_1,\ldots,x_n),(y_1,\ldots,y_n)) = \sqrt{|x_1-y_1|^2+\cdots+|x_n-y_n|^2}$, which is the standard metric on $\mathbb{C}^n$. This forms a complete metric, so the inner product indeed turns $\mathbb{C}^n$ into a Hilbert space.

There exists a generalization of this example to infinite sequences. Fix a set X and define $\ell^2(X)$ to be the set of functions $\varphi : X \to \mathbb{C}$ for which the set $\mathrm{supp}(\varphi) := \{x \in X \mid \varphi(x) \neq 0\}$ is at most countable, and $\sum_{x\in X}|\varphi(x)|^2$ is finite. The first condition is needed to ensure that we can speak about convergence of the sum in the second condition. The inner product is defined by

$$\langle \varphi|\psi\rangle = \sum_{x\in X}\overline{\varphi(x)}\psi(x).$$

One can prove that this sum is well-defined and induces a complete metric on $\ell^2(X)$. If we take X to be a finite set, then we obtain the example of finite sequences above.

In a sense, the spaces $\ell^2(X)$ are the only Hilbert spaces, because every Hilbert space is isomorphic to $\ell^2(X)$ for some set X.

In linear algebra, vector spaces are often studied using bases. A useful analogue for Hilbert spaces is an orthonormal basis.

Definition 2.3. A subset A of a Hilbert space H is called an *orthonormal basis* if the following conditions hold.

- For all $a, b \in A$ we have

$$\langle a|b\rangle = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{otherwise} \end{cases}$$

- Every $x \in H$ can be written as a countable linear combination

$$x = \sum_{i=1}^{\infty} c_i a_i,$$

with $c_i \in \mathbb{C}$ and $a_i \in A$.

Proposition 2.4. *Every Hilbert space has an orthonormal basis.*

An essential feature of quantum computations is reversibility. In the Hilbert space formalism this arises from a bijection between operators $H \to K$ and operators $K \to H$, where H and K are Hilbert spaces. We will describe this situation using the categorical framework of dagger categories.

Definition 2.5. A *dagger category* is a category $\mathbf{C}$ together with a contravariant functor $\dagger : \mathbf{C} \to \mathbf{C}$ that acts as the identity on objects and satisfies $\dagger^2 = \mathrm{Id}$. The action of the functor $\dagger$ on morphisms $f : X \to Y$ in $\mathbf{C}$ is written as $\dagger(f) = f^\dagger : Y \to X$. The map $f^\dagger$ is called the *dagger* or *adjoint* of f.

Proposition 2.6. *The category* **Hilb** *is a dagger category.*

Proof. We will only indicate the construction of the dagger, referring to [23] for the details. Given any operator $A : H \to K$, there exists a unique operator $B : K \to H$ for which $\langle B(x)|y\rangle = \langle x|A(y)\rangle$ for each $x \in K$ and $y \in H$. We define $A^\dagger$ to be this map B. This forms a contravariant functor and $A^{\dagger\dagger} = A$. □

We can identify points x in a Hilbert space H with operators $\mathbb{C} \to H$, whose value at $1 \in \mathbb{C}$ is x. In particular, complex numbers are identified with linear maps $\mathbb{C} \to \mathbb{C}$ in this way. Using this identification, there is a connection between daggers and inner products given by $x^\dagger y = \langle x|y\rangle$. This observation lies at the basis of Dirac's bra-ket notation. The bra-ket notation gives an independent meaning to the constituents $\langle x|$ and $|y\rangle$ of an inner product $\langle x|y\rangle$. The so-called *ket* $|y\rangle$ is interpreted as the operator $\mathbb{C} \to H$ corresponding to $y \in H$, and the *bra* $\langle x|$ as the operator $x^\dagger : H \to \mathbb{C}$. Then the bra-ket or bracket $\langle x|y\rangle$ is the inner product of x and y. This notation simplifies formal manipulations involving inner products and daggers. For example, we can define the *outer product* of the points x and y using bra-ket notation as $|x\rangle\langle y|$. Sometimes we want to apply an operator to a ket before taking the inner product. For this we use the notation $\langle x|A|y\rangle := \langle x|A(y)\rangle = \langle A^\dagger(x)|y\rangle$.

Adjoint operators serve to define a few classes of maps on Hilbert spaces that play an important role in the mathematical formulation of quantum mechanics. Actually these classes make sense for any dagger category, not only for Hilbert spaces.

Definition 2.7. An operator $A : H \to K$ on Hilbert spaces is called

- an *isometry* if $A^\dagger A = \mathrm{id}_H$;
- *unitary* if it is an isometry and moreover $AA^\dagger = \mathrm{id}_K$;
- *self-adjoint* if $A^\dagger = A$ (in this case $H = K$);
- a *projection* if it is self-adjoint and $A^2 = A$.

The term 'isometry' is justified by the following result, which characterizes the isometries as the distance-preserving maps.

Proposition 2.8. *The following are equivalent for an operator $A : H \to K$.*

1. *A is an isometry.*
2. *For all $x, y \in H$ we have $\langle A(x)|A(y)\rangle = \langle x|y\rangle$.*
3. *For all $x, y \in H$ we have $d(A(x), A(y)) = d(x, y)$, where d is the metric induced by the inner product.*

There are some functorial constructions that cannot be performed on all morphisms in **Hilb**, but only on isometries, such as the formations of density matrices and effects in Sections 2.3 and 2.4. Therefore it is useful to define a category $\mathbf{Hilb}_{\mathbf{Isomet}}$ of Hilbert spaces with isometries as morphisms.

Operators on finite-dimensional vector spaces always have a trace, defined as the sum of the diagonal entries of any of its matrix representations. This definition is independent of the choice of matrix representation. However, in infinite-dimensional spaces, the sum of the diagonal entries may be infinite or basis-dependent. We will define a class of operators which are well-behaved in the sense that these phenomena do not occur. There are several possible definitions; here we follow the development of [14].

Definition 2.9. An operator A on a Hilbert space H is said to be of *trace class* if there exist two orthonormal bases $(e_i)_{i\in I}$ and $(f_i)_{i\in I}$ of H such that the sum $\sum_{i,j} |\langle e_i|A|f_j\rangle|$ is finite.

Proposition 2.10. *Let A be a trace class operator on H, and let $(e_i)_{i\in I}$ and $(f_i)_{i\in I}$ be arbitrary orthonormal bases. Then:*

1. *The sum $\sum_{i,j} |\langle e_i|A|f_j\rangle|$ is finite.*
2. *$\sum_i \langle e_i|A|e_i\rangle = \sum_i \langle f_i|A|f_i\rangle < \infty$.*

Definition 2.11. For a trace class operator $A : H \to H$, we define the *trace* of f, denoted $\mathrm{tr}(f)$, as follows. Let $(e_i)_{i\in I}$ be an orthonormal basis for H, and set

$$\mathrm{tr}(f) = \sum_i \langle e_i|A|e_i\rangle.$$

This is finite and basis-independent by Proposition 2.10.

The next result lists some properties of trace class operators.

Proposition 2.12.

1. *Spectral theorem: if A is a trace class operator on H, then H possesses an orthonormal basis consisting of all eigenvectors of A.*

2. *Cyclic property: if A is trace class and B is any operator, then AB and BA are trace class, and* $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

3. *Trace class operators on a fixed Hilbert space form an inner product space. Addition and scalar multiplication are defined pointwise, and the inner product is given by* $\langle A|B\rangle = \mathrm{tr}(A^\dagger B)$.

We record a special case of the third point. When H is finite-dimensional, all operators on H are trace class. Hence the vector space $\mathrm{End}(H)$ of all operators from H to H forms an inner product space. Because $\mathrm{End}(H)$ is finite-dimensional, it is automatically complete and thus a Hilbert space.

Given two Hilbert spaces, we can construct a new Hilbert space using the tensor product. Their tensor product as vector spaces need not be a Hilbert space. Therefore we define the tensor product of Hilbert spaces using a completion as follows.

Definition 2.13. Let H and K be Hilbert spaces with inner products $\langle - | - \rangle_H$ and $\langle - | - \rangle_K$, respectively. Consider H and K as vector spaces and call the tensor product $H \otimes K =: L$. Define an inner product on L by letting

$$\langle x_1 \otimes y_1 | x_2 \otimes y_2 \rangle = \langle x_1 | y_1 \rangle_H \cdot \langle x_2 | y_2 \rangle_K$$

and extending this sesquilinearly to a function $L \to \mathbb{C}$. The *tensor product* $H \otimes K$ is defined as the completion of L in the metric induced by this inner product.

2.2 Postulates of quantum mechanics

A physical theory specifies the possible states of the universe, or usually an isolated part of the universe that can be considered as an entity on its own. These states can evolve according to physical laws, and they can be measured in experiments. This section presents the states, evolution of states, and measurements for quantum mechanics. This description of the quantum world is due to John von Neumann [27]. Since our motivation comes from the field of quantum computation, our examples will be from a computational nature. The formulation of the postulates of quantum mechanics presented in this section is well-known in physics, see for example [7, 22]. An introduction to this subject from a perspective of computer science is for example [28]. Here we do not aim for an exhaustive description of quantum mechanics, but we only discuss the fragment that is relevant for our coalgebraic perspective later on.

We postulate that the state space of a quantum system is given by a Hilbert space H. The states of the system are unit vectors in H, that is, elements $|\psi\rangle$ of H with norm 1. The letter ψ for states of a quantum system is conventional in physics. As an example, in quantum computation we wish to view data as a state in some Hilbert space. The fundamental unit of data is a qubit, which is the quantum analogue of the classical notion of a bit. We let $\mathbb{C}^2$ be the state space for a qubit, and we write the standard basis as

$$|0\rangle = (1,0), \qquad |1\rangle = (0,1).$$

Then a qubit is a unit vector in $\mathbb{C}^2$, hence it is of the form $\alpha|0\rangle + \beta|1\rangle$, where $|\alpha|^2 + |\beta|^2 = 1$. We can recover the classical bits $|0\rangle$ and $|1\rangle$ as special instances of a qubit. So a qubit is a linear combination of classical bits, which is called a superposition state. Because quantum data can be in a superposition state, there are more ways to manipulate quantum data than to manipulate classical data.

If we have a quantum system consisting of multiple particles, then each particle can be described with a certain Hilbert space. Often we want to take the interaction between the particles into account. In that case we have to form a composite state space from the individual Hilbert spaces. Let $(H_i)_{i\in I}$ be the family of Hilbert spaces describing the individual particles. Then we stipulate that the Hilbert space for the composite system is the tensor product $\bigotimes_{i\in I} H_i$. For example, a qubyte consists of eight qubits, so a qubyte is represented as a state in the Hilbert space

$$\underbrace{\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{8} \cong \mathbb{C}^{256}.$$

Observers who want to obtain information about a quantum system have to perform a measurement. For our purposes, a measurement or observable is a yes-no question about the system. If the Hilbert space H is our state space, then we represent an observable as a projection operator P on H. The quantum world is probabilistic, which means that if we prepare a system twice in exactly the same state, then measurement might yield a different value both times. We can only specify the probability that a measurement yields 'yes' or 'no'. This probability is given by the Born rule: if the system is in state $|\psi\rangle$ and we measure the observable P, then the probability that the result is 'yes' is $\langle\psi|P|\psi\rangle$.

Finally we have to discuss the dynamics, which tells us how to modify data during a computation. There are two distinct, though equivalent, visions on the dynamics of a quantum system. The first one is the Schrödinger picture, in which the states change in time, while the observables are constant. We require that states can only be modified using unitary operators.

If we apply the unitary operator U to the state $|\psi\rangle$, then the resulting state is simply $U|\psi\rangle$. This is again a unit vector, since

$$\langle U\psi|U|\psi\rangle = \langle\psi|U^\dagger U|\psi\rangle = \langle\psi|\psi\rangle = 1.$$

Because of the unitarity condition, all quantum processes are reversible.

An example of a unitary operator that acts on qubits is the Hadamard operator H. It is defined on the basis vectors of $\mathbb{C}^2$ by

$$H|0\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$
$$H|1\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$$

This operator satisfies $H^\dagger = H$ and $H^2 = \mathrm{id}$, so it is unitary. Suppose that we start with the qubit $|1\rangle$, we apply H to the state, and we want to know the probability that the qubit still has value 1 after the transition. The qubit is in state $\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$ after the transition. The observable corresponding to the value 1 is $|1\rangle\langle 1|$, so according to the Born rule, the probability is

$$\left(\frac{1}{\sqrt{2}}\langle 0| - \frac{1}{\sqrt{2}}\langle 1|\right)|1\rangle\langle 1|\left(\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle\right) = \frac{1}{2}.$$

The second vision on dynamics is the Heisenberg picture, according to which the observables are variable while the states are fixed. The result of applying the unitary operator U to the observable P is $U^\dagger PU$. From the following computations it follows that $U^\dagger PU$ is a projection.

$$(U^\dagger PU)^\dagger = U^\dagger P^\dagger U = U^\dagger PU$$

$$(U^\dagger PU)^2 = U^\dagger PUU^\dagger PU = U^\dagger P^2 U = U^\dagger PU$$

It is not hard to see that both pictures of dynamics are equivalent. Indeed, if we prepare a system in state $|\psi\rangle$, then apply the unitary transformation U, and finally measure the projection P, then the probability that this measurement results in 'yes' is, according to the Schrödinger picture, $\langle U\psi|P|U\psi\rangle$. In the Heisenberg picture we would get the probability $\langle\psi|U^\dagger PU|\psi\rangle$. Both probabilities are equal by the defining property of adjoint operators.

2.3 Density matrices

The probabilistic nature of quantum mechanics is a fundamental physical fact and does not arise from human ignorance. Hence, even if the state of a system is known to an observer, (s)he may not be able to predict the outcome of a measurement. But if the observer does not know the state of the system, this introduces another layer of uncertainty. The standard treatment of this statistical uncertainty uses density matrices, which generalize the states in a Hilbert space. Besides being more general than the formalism of states in Hilbert spaces, the framework of density matrices will also turn out to be more suitable for a coalgebraic viewpoint on quantum systems. Density matrices were proposed by von Neumann in [27].

Before defining density matrices, it is convenient to introduce a partial order on operators.

Definition 2.14. Given two self-adjoint operators A and B on a Hilbert space H, define $A \le B$ if and only if $\langle x|A|x\rangle \le \langle x|B|x\rangle$ for each $x \in H$. As a special case, the operator A is said to be *positive* if $A \ge 0$.

This definition makes sense because $\langle x|A|x\rangle$ is always a real number if A is self-adjoint. We will show that $\le$ is indeed a partial order on self-adjoint operators.

Lemma 2.15. *Let A and B be operators on H.*

1. *If $\langle x|A|x\rangle = 0$ for every $x \in H$, then A is the zero operator.*
2. *If $\langle x|A|x\rangle = \langle x|B|x\rangle$ for every $x \in H$, then $A = B$.*

Proof. This proof comes from [23].

1. The equation $\langle x|A|x\rangle = 0$ holds for all x, so in particular it holds for all x of the form $\alpha Ay + y$ with $\alpha \in \mathbb{C}$ and $y \in H$. Expanding this gives

$$\begin{aligned} 0 &= \langle x|A|x\rangle \\ &= \langle \alpha Ay + y|A|\alpha Ay + y\rangle \\ &= |\alpha|^2\langle Ay|A|Ay\rangle + \overline{\alpha}\langle Ay|A|y\rangle + \alpha\langle y|A|Ay\rangle + \langle y|A|y\rangle \\ &= \overline{\alpha}\langle Ay|A|y\rangle + \alpha\langle y|A|Ay\rangle. \end{aligned}$$

   Taking $\alpha = 1$ and $\alpha = i$ gives two equations:

$$\langle Ay|Ay\rangle + \langle y|A|y\rangle = 0$$

$$\langle Ay|Ay\rangle - \langle y|A|y\rangle = 0$$

   Adding these equations gives $\langle Ay|Ay\rangle = 0$, hence $Ay = 0$ for each y, thus $A = 0$.

2. From $\langle x|A|x\rangle = \langle x|B|x\rangle$ it follows that $\langle x|A - B|x\rangle = 0$. Apply part 1 to obtain $A = B$. □

Proposition 2.16. *The relation $\le$ from Definition 2.14 is a partial order on self-adjoint operators.*

Proof. Reflexivity and transitivity are obvious. For anti-symmetry, suppose that $A \le B$ and $B \le A$. Then $\langle x|A|x\rangle = \langle x|B|x\rangle$ for each x, so $A = B$ by Lemma 2.15. □

Definition 2.17. A *density operator* or *density matrix* on a Hilbert space H is a positive trace-class operator $\rho : H \to H$ for which $\mathrm{tr}(\rho) = 1$.

Unit vectors in a Hilbert space are called *pure states* in physics, and density operators are called *mixed states*. These mixed states represent probability distributions over pure states. We will now investigate how to construct the mixed state corresponding to such a probability distribution.

If $|\psi\rangle$ is a pure state, then it is easy to see that $|\psi\rangle\langle\psi|$ is a density matrix. This represents the situation in which it is certain that the system is in state $|\psi\rangle$. An arbitrary probability distribution over pure states is given by a list of states $|\psi_i\rangle$ for $i = 1, \dots, n$ with corresponding probabilities p_i summing to 1. It gives a mixed state with the shape $\sum_{i=1}^n p_i|\psi_i\rangle\langle\psi_i|$. This is again a density matrix, since a convex combination of density matrices is always a density matrix.

In the previous section, we described evolution and measurements of pure states. We can do the same for mixed states. If a system is in the mixed state ρ, then application of the unitary operator U leaves it in the state $U\rho U^\dagger$. Measurement of the projection P results in the answer 'yes' with probability $\mathrm{tr}(\rho P)$. The following lemma shows that these are well-defined.

Lemma 2.18. *Let ρ be a density matrix, A an isometry, and P a projection. Then:*

1. *The operator $A\rho A^\dagger$ is a density matrix.*
2. *The number $\mathrm{tr}(\rho P)$ lies in the interval $[0, 1]$.*

Proof.

1. The operator $A\rho A^\dagger$ is positive because

$$\langle x|A\rho A^\dagger|x\rangle = \langle A^\dagger x|\rho|A^\dagger x\rangle \ge 0,$$

and it has trace 1 since

$$\mathrm{tr}(A\rho A^\dagger) = \mathrm{tr}(A^\dagger A\rho) = \mathrm{tr}(\rho) = 1,$$

where we used that A is an isometry. For the trace class requirement, let $(e_i)_{i\in I}$ be a basis for $\mathrm{im}(A) \subseteq H$. We claim that $(A^\dagger e_i)_{i\in I}$ is an orthonormal basis for H. The operator A is an isometry, so it is unitary as a map onto its image. Therefore, for all $i, i' \in I$,

$$\langle A^\dagger e_i|A^\dagger e_{i'}\rangle = \langle e_i|AA^\dagger e_{i'}\rangle = \langle e_i|e_{i'}\rangle = \begin{cases} 1 & \text{if } i = i' \\ 0 & \text{otherwise.} \end{cases}$$

Pick a point x in H. We wish to write x as a countable linear combination of the vectors $A^\dagger e_i$. Write $Ax \in \mathrm{im}(A)$ in the basis $(e_i)_{i\in I}$ as $Ax = \sum_i c_i e_i$. Then $x = A^\dagger A x = \sum_i c_i A^\dagger e_i$.

2. The density matrix ρ is trace class, so by Proposition 2.12 the operator ρP is also trace class. Hence $\mathrm{tr}(\rho P)$ is well-defined. Since ρ is trace class, we can choose a basis (e_i) of eigenvectors of ρ, with corresponding eigenvalues (λ_i). Calculate the trace in this basis to obtain

$$\mathrm{tr}(\rho P) = \mathrm{tr}(P\rho) = \sum_i \langle e_i|P\rho|e_i\rangle = \sum_i \lambda_i \langle e_i|P|e_i\rangle.$$

The eigenvalues λ_i have sum 1 because ρ has trace 1, and the values $\langle e_i|P|e_i\rangle$ lie in $[0,1]$. Therefore this expression lies between 0 and 1. □

The above assertion can also be understood from a categorical viewpoint. Density matrices live in the category **Conv** introduced in Section 1.6, and their construction is functorial.

Proposition 2.19. *There is a functor*

$$\mathcal{DM} : \mathbf{Hilb_{Isomet}} \to \mathbf{Conv}$$

defined on objects by

$$\mathcal{DM}(H) = \{\rho : H \to H \mid \rho \textit{ is a density matrix}\}$$

and on morphisms by

$$\mathcal{DM}(A : H \to K)(\rho) = A\rho A^\dagger.$$

Proof. The set of density matrices on a Hilbert space is convex, since convex combinations of positive maps are positive, the trace preserves convex combinations, and trace-class operators form a vector space. The operator $\mathcal{DM}(A)(\rho)$ is a density matrix by Lemma 2.18. It is clear that $\mathcal{DM}(A)$ is a convex map, and that $\mathcal{DM}$ is a functor. □

We see that if U is a unitary operator describing a transition on pure states, then $\mathcal{DM}(U)$ provides the corresponding transition on mixed states.

We finish this section by showing that the frameworks of Hilbert spaces and density matrices yield the same experimental predictions when applied to pure states. Let $|\psi\rangle$ be a state in a Hilbert space H, U a unitary operator on H, and P an observable. We first apply U to the state $|\psi\rangle$ and then measure P. According to the framework in Section 2.2, the probability of obtaining 'yes' is $\langle U\psi|P|U\psi\rangle$. With density matrices, we start in the state $|\psi\rangle\langle\psi|$, so this probability is $\mathrm{tr}(U|\psi\rangle\langle\psi|U^\dagger P)$.

Proposition 2.20. *With $|\psi\rangle$, U, and P as described above, $\langle U\psi|P|U\psi\rangle = \mathrm{tr}(U|\psi\rangle\langle\psi|U^\dagger P)$.*

Proof. By the cyclic property of the trace, $\mathrm{tr}(U|\psi\rangle\langle\psi|U^\dagger P) = \mathrm{tr}(\langle\psi|U^\dagger PU|\psi\rangle)$. The right-hand side is a trace in the one-dimensional space $\mathbb{C}$, so it equals $\langle\psi|U^\dagger PU|\psi\rangle = \langle U\psi|P|U\psi\rangle$. □

2.4 Effects

We have generalized the quantum states to density matrices, but it will turn out also to be useful to generalize the observables to so-called effects. Physically, effects are unsharp measurements, which differ from ordinary measurements in that they need not satisfy the principle of non-contradiction. Effect algebras were introduced in [12]. For an overview of the theory of effects and effect algebras we refer to [10].

Definition 2.21. An *effect* on a Hilbert space H is a operator A for which $0 \leq A \leq \mathrm{id}$.

The outcomes of measurements for effects behave the same as for projections: measurement of an effect A when the system is in state ρ succeeds with probability $\mathrm{tr}(\rho A)$.

We will now take a look at the algebraic and categorical organization of effects. The sum of two effects may or may not be an effect, because the condition that it lies below the identity may be violated. So addition is a partial operation on effects. Furthermore, for every effect A the complement $\mathrm{id} - A$ is again an effect. We capture this structure in the following notion.

Definition 2.22. An *effect algebra* consists of a set E equipped with two distinguished elements $0, 1 \in E$ and a partial binary operation $\boxplus$ satisfying the following conditions.

- Commutativity: if $x \boxplus y$ is defined, then so is $y \boxplus x$, and $x \boxplus y = y \boxplus x$.
- Associativity: if $x \boxplus y$ and $(x \boxplus y) \boxplus z$ are defined, then so are $y \boxplus z$ and $x \boxplus (y \boxplus z)$, and $(x \boxplus y) \boxplus z = x \boxplus (y \boxplus z)$.
- Orthocomplement: for every $x \in E$ there exists a unique $y \in E$ such that $x \boxplus y$ is defined and equals 1.
- Zero-one law: if $x \boxplus 1$ is defined, then $x = 0$.

A homomorphism from E to E' is a map $f : E \to E'$ for which $f(0) = 0$, $f(1) = 1$, and if $x \boxplus y$ is defined, then $f(x) \boxplus f(y)$ is defined and $f(x) \boxplus f(y) = f(x \boxplus y)$. The category of effect algebras and homomorphisms is called **EA**.

We will use a few conventions when talking about effect algebras. Whenever we write $x \boxplus y$, we presuppose that this expression is defined. Furthermore, given an element x in an effect algebra, the unique y with $x \boxplus y = 1$ is denoted $y = x^\perp$.

Examples 2.23.

1. A standard example of an effect algebra is the unit interval $[0,1]$ with addition as binary operation. The complement is given by $x^\perp = 1 - x$.

2. As we observed before Definition 2.22, the set of effects on a Hilbert space forms an effect algebra. For two effects A and B, $A \boxplus B$ is defined if and only if $A + B \leq \mathrm{id}$, and in that case $A \boxplus B = A + B$. The zero and one are the zero operator and the identity, respectively. Furthermore $A^\perp = \mathrm{id} - A$.

3. There is a class of examples coming from lattice theory. A bounded lattice L is called *orthomodular* if there is an operation ${}^\perp : L \to L$ such that:

   - $x \leq y$ implies $y^\perp \leq x^\perp$;
   - $x^{\perp\perp} = x$;
   - $x \vee x^\perp = 1$;
   - Orthomodular law: if $x \leq y$, then $x \vee (x^\perp \wedge y) = y$.

   Every orthomodular lattice can be turned into an effect algebra where $x \boxplus y$ is defined if and only if $x \leq y^\perp$, and in that case $x \boxplus y = x \vee y$. The elements 0 and 1 exist because we assumed a bounded lattice. The orthocomplement is the operation ${}^\perp$. This example lies at the basis of quantum logic, since quantum logic studies projections on Hilbert space, and these projections form an orthomodular lattice.

4. The forgetful functor $\mathbf{EA} \to \mathbf{Sets}$ has a left adjoint that gives free effect algebras. We will give an explicit description of the free effect algebra on a set X, which is denoted $MO(X)$. It consists of the elements 0, 1, x, and $x^\perp$ for $x \in X$. The partial operation $\boxplus$ is defined as follows.

   - $z \boxplus 0 = 0 \boxplus z = z$ for all $z \in MO(X)$;
   - $x \boxplus x^\perp = x^\perp \boxplus x = 1$ for all $x \in X$;
   - For all other cases, $\boxplus$ is undefined.

Our aim in Section 2.5 will be to connect density matrices and effects on Hilbert spaces. For this we need one more piece of structure on the effects, namely the multiplication with scalars in the unit interval $[0,1]$.

Definition 2.24. An *effect module* is an effect algebra E endowed with a scalar product $\cdot : [0,1] \times E \to E$ for which the following conditions hold for all $r, s \in [0,1]$ and $x, y \in E$.

- $r \cdot (s \cdot x) = (rs) \cdot x$.
- If $r + s \leq 1$, then $(r + s) \cdot x = r \cdot x \boxplus s \cdot x$.
- If $x \boxplus y$ is defined, then $r \cdot (x \boxplus y) = r \cdot x \boxplus r \cdot y$.
- $1 \cdot x = x$.

The morphisms between effect modules are effect module homomorphisms, which are effect algebra homomorphisms f that additionally preserve the scalar product, as in $f(r \cdot x) = r \cdot f(x)$. The resulting category is denoted **EMod**.

Effect modules were first considered in [15] under the name 'convex effect algebras', and generalized in [20] to modules over arbitrary effect algebras with a monoid structure, rather than just over the interval $[0,1]$. For our purposes this generality is not needed, so we restrict to effect modules with scalars in the unit interval.

Example 2.25. We briefly review our examples of effect algebras. The unit interval $[0,1]$ is an easy example of an effect module where the scalar product is ordinary multiplication. Effects on a Hilbert space also form an effect module with the scalar multiplication on operators. It is in general not possible to turn orthomodular lattices into effect modules.

Given an effect algebra E, the free effect module on E is $[0,1] \otimes E$, where $\otimes$ denotes the tensor product of effect algebras. The tensor product is characterized by the property that effect algebra homomorphisms $E \otimes E' \to E''$ correspond to bimorphisms $f : E \times E' \to E''$, which are morphisms preserving $\boxplus$ in both components and satisfying $f(1,1) = 1$. Thus the forgetful functor $\mathbf{EMod} \to \mathbf{EA}$ has a left adjoint.

The construction of effects on a Hilbert space can be extended to a contravariant functor from $\mathbf{Hilb_{Isomet}}$ to **EMod**.

Proposition 2.26. *There exists a functor*

$$\mathcal{E}f : \mathbf{Hilb_{Isomet}} \to \mathbf{EMod}^{\mathrm{op}}$$
$$\mathcal{E}f(H) = \{A \mid 0 \leq A \leq \mathrm{id}\}$$
$$\mathcal{E}f(B : H \to K)(A) = B^\dagger AB$$

Proof. From the facts that A is an effect and $\langle x|B^\dagger AB|x\rangle = \langle Bx|A|Bx\rangle$ it follows that $\mathcal{E}f(B)(A)$ is an effect.

To show that $\mathcal{E}f(B)$ is a homomorphism of effect modules, suppose that $A \boxplus A'$ is defined. This means that $A + A' \le \mathrm{id}$. Then for each $x \in K$ we have

$$\langle x|(B^\dagger AB + B^\dagger A'B)|x\rangle = \langle Bx|A + A'|Bx\rangle \le 1,$$

hence $\mathcal{E}f(B)(A) + \mathcal{E}f(B)(A') \le \mathrm{id}$ and $\mathcal{E}f(B)(A) + \mathcal{E}f(B)(A') = \mathcal{E}f(B)(A + A')$. The map $\mathcal{E}f(B)$ preserves scalar multiplication by linearity. Furthermore $\mathcal{E}f(B)$ preserves the constants of the effect modules: for 0 this is obvious, and for id this follows because B is an isometry.

Finally it is clear that the assignment $\mathcal{E}f$ is a functor. □

As effects are generalized observables, we can formulate the Heisenberg picture of quantum mechanics in terms of effects. Effects evolve in the same manner as projections: if A is an effect and U is a unitary operator on a Hilbert space H, then application of U gives the effect $\mathcal{E}f(U)(A) = U^\dagger AU$. The resulting picture of dynamics is once again equivalent to the Schrödinger picture in terms of density matrices: consider an experiment involving an initial state ρ, a unitary transformation U and an effect A. The probability that this experiment succeeds is $\mathrm{tr}(U\rho U^\dagger A)$ according to the Schrödinger picture, and $\mathrm{tr}(\rho U^\dagger AU)$ when the probability is calculated in the Heisenberg picture. Both probabilities are equal by the cyclic property of the trace.

2.5 Relation between density matrices and effects

In [17], it is shown that convex sets and effect algebras are related via a dual adjunction. This is strenghtened to a dual adjunction between convex sets and effect modules in [20], which we will describe here. It leads to a connection between the density matrices and effects on a Hilbert space, in other words, the possible states and possible measurements on a quantum mechanical system.

We will start by defining functors that assign convex sets to effect modules and vice versa. Subsequently we will see that these functors form an adjunction.

Lemma 2.27.

1. *The assignment* $\mathbf{Conv}(-,[0,1])$ *is a functor from* $\mathbf{Conv}$ *to* $\mathbf{EMod}^{\mathrm{op}}$.

2. *The assignment* $\mathbf{EMod}(-,[0,1])$ *is a functor from* $\mathbf{EMod}^{\mathrm{op}}$ *to* $\mathbf{Conv}$.

Proof.

1. For each convex set X, the set of morphism $\mathbf{Conv}(X,[0,1])$ carries the structure of an effect module with pointwise operations. Explicitly, the 0 and 1 of $\mathbf{Conv}(X,[0,1])$ are the constant maps with values 0 and 1. Given convex maps $f, g : X \to [0,1]$, the sum $f \boxplus g$ is defined if

and only if $f(x)+g(x) \leq 1$ for every $x \in X$, and in that case $(f \boxplus g)(x) = f(x) + g(x)$. Then $f \boxplus g$ is again a convex map, since

$$\begin{aligned}(f \boxplus g)(\sum_i r_i x_i) &= f(\sum_i r_i x_i) + g(\sum_i r_i x_i) \\ &= \sum_i r_i f(x_i) + \sum_i r_i g(x_i) \\ &= \sum_i r_i (f \boxplus g)(x_i).\end{aligned}$$

Similarly one checks that, for each $r \in [0,1]$, the map rf defined by $(rf)(x) = rf(x)$ is convex.

Let $f : X \to Y$ be a convex map. Then $\mathbf{Conv}(f,[0,1]) : \mathbf{Conv}(Y,[0,1]) \to \mathbf{Conv}(X,[0,1])$ given by $\varphi \mapsto \varphi \circ f$ is an effect module homomorphism. For example, it preserves addition $\boxplus$, since $(\varphi \boxplus \psi) \circ f = \varphi \circ f \boxplus \psi \circ f$ for all $\varphi, \psi : Y \to [0,1]$.

2. Analogous to 1. $\square$

Theorem 2.28. *There is a dual adjunction between convex sets and effect modules:*

$$\begin{tikzcd} \mathbf{Conv} \arrow[r, bend left, "\mathbf{Conv}(-,{[0,1]})"] \arrow[r, phantom, "\perp"] & \mathbf{EMod}^{\mathrm{op}} \arrow[l, bend left, "\mathbf{EMod}(-,{[0,1]})"] \end{tikzcd}$$

Proof. We have to establish a natural bijection

$$\begin{array}{c l} X \to \mathbf{EMod}(Y,[0,1]) & \text{in } \mathbf{Conv} \\ \hline\hline Y \to \mathbf{Conv}(X,[0,1]) & \text{in } \mathbf{EMod} \end{array}$$

Given a convex map $f : X \to \mathbf{EMod}(Y,[0,1])$, define the map $\bar{f} : Y \to \mathbf{Conv}(X,[0,1])$ by $\bar{f}(y)(x) = f(x)(y)$. The map $\bar{f}$ is a homomorphism because

$$\begin{aligned}\bar{f}(y \boxplus y') &= \lambda x \in X.f(x)(y \boxplus y') \\ &= \lambda x \in X.(f(x)(y) + f(x)(y')) && \text{since } f(x) \text{ is a morphism in } \mathbf{EMod} \\ &= (\lambda x \in X.f(x)(y)) \boxplus (\lambda x \in X.f(x)(y')) \\ &= \bar{f}(y) \boxplus \bar{f}(y'),\end{aligned}$$

and similarly $\bar{f}$ preserves the other operations of effect modules.

Conversely, from a map of effect modules $g : Y \to \mathbf{Conv}(X,[0,1])$ we construct $\bar{g} : X \to \mathbf{EMod}(Y,[0,1])$ as $\bar{g}(x)(y) = g(y)(x)$. It is easy to show that $\bar{g}$ is a convex map, using the fact that each $g(y)$ preserves convex combinations. Obviously both constructions are inverses, which proves the desired. $\square$

We have seen that the construction of density matrices and effects on a Hilbert space is functorial. We can collect this information together with the above result in the following diagram.

$$\begin{tikzcd}
& \mathbf{Hilb}_{\mathbf{Isomet}} \arrow[dl, "\mathcal{DM}"'] \arrow[dr, "\mathcal{E}f"] & \\
\mathbf{Conv} \arrow[rr, bend left, "\mathbf{Conv}(-,[0,1])"] \arrow[rr, phantom, "\perp"] & & \mathbf{EMod}^{\mathrm{op}} \arrow[ll, bend left, "\mathbf{EMod}(-,[0,1])"]
\end{tikzcd}$$

One may ask if this diagram commutes. On objects this means that, for each Hilbert space H, $\mathbf{Conv}(\mathcal{DM}(H),[0,1]) \cong \mathcal{E}f(H)$ and $\mathbf{EMod}(\mathcal{E}f(H),[0,1]) \cong \mathcal{DM}(H)$. We will show that this is the case for finite-dimensional Hilbert spaces, as proven in [5].

Theorem 2.29. *Let H be a finite-dimensional Hilbert space. Then the convex sets $\mathcal{DM}(H)$ and $\mathbf{EMod}(\mathcal{E}f(H),[0,1])$ are isomorphic.*

Proof. Define the convex map

$$\begin{aligned}\Phi : \mathcal{DM}(H) &\to [0,1]^{\mathcal{E}f(H)}\\ \rho &\mapsto \lambda A \in \mathcal{E}f(H).\,\mathrm{tr}(\rho A)\end{aligned}$$

By choosing an orthonormal basis of eigenvectors of ρ, we see that $\mathrm{tr}(\rho A)$ always lies in $[0,1]$, so Φ is well-defined. We will show that Φ is an isomorphism. For injectivity, suppose that $\mathrm{tr}(\rho -) = \mathrm{tr}(\sigma -)$. The operator $|x\rangle\langle x|$ is an effect for each $x \in H$, so $\mathrm{tr}(\rho|x\rangle\langle x|) = \mathrm{tr}(\sigma|x\rangle\langle x|)$. Take an orthonormal basis for H in which x occurs as a basis vector and compute the traces with respect to this basis. Then we obtain $\langle x|\rho|x\rangle = \langle x|\sigma|x\rangle$. From Lemma 2.15 we conclude that $\sigma = \rho$.

Now consider surjectivity. Pick an arbitrary homomorphism $f : \mathcal{E}f(H) \to [0,1]$. We have to find a density matrix ρ such that $f = \mathrm{tr}(\rho -)$. To do this, we will first extend f to an operator from $\mathrm{End}(H)$ to $\mathbb{C}$, also denoted f, in several steps.

We start by defining f on all positive operators. If A is a positive operator, then there exists a natural number n such that $\frac{1}{n}A$ is an effect. Define $f(A) = nf(\frac{1}{n}A)$. This is independent of the choice of n: suppose that $\frac{1}{m}A$ is also an effect, and assume without loss of generality that $n \le m$. Then $\frac{n}{m} \in [0,1]$, so

$$mf\left(\frac{1}{m}A\right) = mf\left(\frac{n}{m}\frac{1}{n}A\right) = m\frac{n}{m}f\left(\frac{1}{n}A\right) = nf\left(\frac{1}{n}A\right),$$

since f preserves scalar multiplication. It is easy to see that this is indeed an extension of the original map $f : \mathcal{E}f(H) \to [0,1]$. Also observe that the extended map preserves addition.

The next step is to define f also on self-adjoint operators. Each self-adjoint A can be written as a difference $A = B - C$ of two positive operators. Define $f(A) = f(B) - f(C)$, which is independent of the choice of operators: If $A = B - C = B' - C'$, then $B + C' = B' + C$. Since f preserves addition this implies $f(B) + f(C') = f(B') + f(C)$, hence $f(B) - f(C) = f(B') - f(C')$.

Finally, to define f on an arbitrary operator A, use the fact that A can be written as $B + iC$, where both B and C are self-adjoint.

Recall that $\mathrm{End}(H)$ has an inner product defined by $\langle A|B\rangle = \mathrm{tr}(A^\dagger B)$. The standard inner product on $\mathbb{C}$ is $\langle x|y\rangle = \overline{x}y$. The map $f : \mathrm{End}(H) \to \mathbb{C}$ has an adjoint $f^\dagger : \mathbb{C} \to \mathrm{End}\, H$ with respect to these inner products for which

$$\overline{x}f(A) = \mathrm{tr}((f^\dagger(x))^\dagger A)$$

by the defining property of the dagger. Let $\rho = (f^\dagger(1))^\dagger$, then the above equation with $x = 1$ becomes

$$f(A) = \mathrm{tr}(\rho A),$$

which is one of the desired properties for ρ.

It remains to be shown that ρ is a density matrix. For positivity, take an arbitrary $x \in H$. Then

$$\langle x|\rho|x\rangle = \mathrm{tr}(\rho|x\rangle\langle x|) = f(|x\rangle\langle x|).$$

Since $|x\rangle\langle x|$ is an effect and f maps effects to elements of $[0,1]$, we get $\langle x|\rho|x\rangle \geq 0$. Because H is finite-dimensional, ρ is automatically trace-class. The trace of ρ can be calculated as follows.

$$\mathrm{tr}(\rho) = \mathrm{tr}(\rho \circ \mathrm{id}) = f(\rho) = 1.$$

The last equality sign holds because f is a map of effect algebras, hence it preserves 1. □

Theorem 2.30. *Let H be a finite-dimensional Hilbert space. Then the effect modules $\mathcal{E}f(H)$ and $\mathbf{Conv}(\mathcal{DM}(H), [0,1])$ are isomorphic.*

Proof. As in the proof of Theorem 2.29, we define the candidate isomorphism using the trace:

$$\begin{aligned} \mathcal{E}f(H) &\to [0,1]^{\mathcal{DM}(H)} \\ A &\mapsto \lambda\rho \in \mathcal{DM}(H).\, \mathrm{tr}(\rho A). \end{aligned}$$

This is clearly an effect module homomorphism.

We will prove injectivity. Suppose that $\mathrm{tr}(-A) = \mathrm{tr}(-B)$. Since $|x\rangle\langle x|$ is a density matrix for each x, we have

$$\langle x|A|x\rangle = \mathrm{tr}(|x\rangle\langle x|A) = \mathrm{tr}(|x\rangle\langle x|B) = \langle x|B|x\rangle.$$

Hence $A = B$ by Lemma 2.15.

For surjectivity we apply a similar strategy as in the proof of Theorem 2.29. Let $f : \mathcal{DM}(H) \to [0,1]$ be a convex map. Extend f to a map defined on all positive operators: for positive A let

$$f(A) = \begin{cases} 0 & \text{if } \mathrm{tr}(A) = 0 \\ f\left(\frac{A}{\mathrm{tr}(A)}\right) & \text{otherwise} \end{cases}$$

The construction in the previous proof shows how to find an operator A such that $f = \mathrm{tr}(-A)$. The operator A is an effect, since

$$\langle x|A|x\rangle = \mathrm{tr}(A|x\rangle\langle x|) = f(|x\rangle\langle x|) \in [0,1]. \qquad \square$$

Chapter 3

Quantum coalgebras

In this chapter we will see a few examples of state-based systems that occur in quantum computation. We will show how to model these faithfully as coalgebras, in such a way that the morphism into the final coalgebra corresponds to the probability distributions over the output of the system.

In Section 3.1 we give several examples of quantum systems, which will be represented as coalgebras in Section 3.2, by abstracting from the Schrödinger picture of quantum mechanics. With this model, we can apply the general theory of coalgebras developed in Chapter 1 to quantum systems. We illustrate this by discussing minimization in Section 3.3 and duality in Section 3.4 and Section 3.5.

3.1 Examples of discrete quantum systems

Before describing quantum systems in general using coalgebra, we will present a few examples. We would like the general theory to include all of these examples.

Example 3.1. This is a version of Example 1.28, but now using quantum probability instead of classical probability. We consider a quantum mechanical particle walking on the square graph with X as set of vertices. It is not possible to say that the particle is always on a well-determined vertex. Instead, the position of the particle is a superposition of the vertices of the graph. In Example 1.28, the particle chose one of the adjacent points in each step randomly with certain probabilities. Here we replace these probabilities by probability amplitudes. Thus after the step the particle is again in a superposition state.

We describe this situation using the formalism of Chapter 2. The state space of this system is the Hilbert space $\mathbb{C}^4$. Write the standard basis as $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$, where $|k\rangle$ corresponds to the vertex labeled x_k. The dynamics is given by the unitary matrix

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix}.$$

For each vertex k, we can measure the probability that the particle is in state k using the projection $|k\rangle\langle k|$.

We consider the trajectory of the particle when it starts in the state $|0\rangle$. After one time step, the particle is in the superposition state $U|0\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)$. Measurement of the particle's position yields 1 or 2, both with probability $\frac{1}{2}$. But after two steps, the particle's state is $U^2|0\rangle = |0\rangle$, so the particle is back in its original state with certainty. This is a quantum mechanical phenomenon that does not occur for the particle's walk governed by classical probability.

A simple classical model of computation is a deterministic automaton. The quantum analogue was defined in [25]. Like a classical automaton, a quantum automaton has states and a transition on these states for each letter in an alphabet, but for a quantum automaton these states and transitions are subject to the postulates of quantum mechanics: the states form a Hilbert space, and the transitions are unitary. The subset of accepting states from a classical automaton is replaced by an effect. This slightly generalizes the situation in [25], where a projection is used for the output of the automaton.

Definition 3.2. A *quantum language* over an alphabet A is a function $A^* \to [0,1]$. One can think of a quantum language as a fuzzy or probabilistic language. The function assigns to a word in A^* the probability that it is in the language.

A *quantum automaton* $\mathbb{A}$ over an alphabet A consists of:

- A Hilbert space H;
- For each letter $a \in A$, a unitary operator $\delta_a : H \to H$;
- An initial state $\psi_0 \in H$ with norm 1;
- An effect ϵ on H.

Define, for each word $u \in A^*$, the extended transition operator $\delta_u : H \to H$ inductively by

$$\delta_\varepsilon(\psi) = \psi,$$
$$\delta_{au}(\psi) = \delta_u(\delta_a(\psi)).$$

The probability that the word u is accepted by the automaton $\mathbb{A}$ is

$$\langle \delta_u \psi_0 | \epsilon | \delta_u \psi_0 \rangle,$$

that is, the probability that measurement of P gives 'yes' in state $\delta_u\psi_0$.

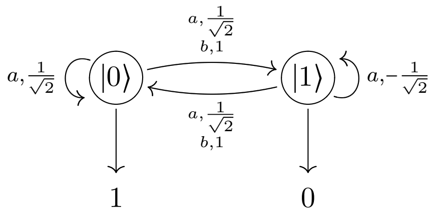
The quantum language *recognized* by $\mathbb{A}$ is the function $A^* \to [0,1]$ that sends a word $u \in A^*$ to the probability that u is accepted by $\mathbb{A}$.

Example 3.3. There is a version of a state diagram for quantum automata. Let $\mathbb{A} = (H, (\delta_a)_{a\in A}, \psi_0, \epsilon)$ be a finite-dimensional quantum automaton over the alphabet A, and let $B = \{|0\rangle, |1\rangle, \dots, |n-1\rangle\}$ be an orthonormal basis for H. In most cases that occur in practice, there is one standard choice for the basis. In the state diagram for $\mathbb{A}$, the basis vectors in B are drawn as circles. A general state is a superposition of these basic states. If $\delta_a(|k\rangle) = \sum_j c_j|j\rangle$, then we draw an arrow from $|k\rangle$ to each $|j\rangle$ for which the coefficient c_j is non-zero. This arrow is labeled with the letter a and the probability amplitude c_j. The effect ϵ is indicated by arrows pointing out of the states $|k\rangle$ to the value $\langle k|\epsilon|k\rangle$.

We illustrate this with an example of a quantum automaton over the alphabet $A = \{a,b\}$. Take the two-dimensional state space $\mathbb{C}^2$ with the standard basis written as $|0\rangle$, $|1\rangle$. The transition matrices are

$$\delta_a = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad \delta_b = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The initial state is $\psi_0 = |0\rangle$ and the effect is $\epsilon = |0\rangle\langle 0|$. This automaton can be drawn with the following state diagram.



The automaton gives a probabilistic language $\{a,b\}^* \to [0,1]$. For instance, the probability that the word ba is accepted is $\langle \delta_a\delta_b 0|0\rangle\langle 0|\delta_a\delta_b 0\rangle = \frac{1}{2}$.

The quantum walks form another class of examples of systems. A quantum walk is a variation on Example 3.1 in which the movement of the particle is controlled by an additional structure, usually the spin of the particle. Quantum walks are discussed extensively in [32]. See also [18] for coalgebraic versions. We will take the example from the latter reference.

Example 3.4. Consider a particle walking on the line $\mathbb{Z}$ of integers. In addition to the position of the particle on $\mathbb{Z}$, we take its spin into account. The spin is a superposition of the basic spin states $|\uparrow\rangle$ and $|\downarrow\rangle$, so the state space for the spin is $\mathbb{C}^2$. The state space for the position is $\ell^2(\mathbb{Z})$, so the Hilbert space modeling the composite system of the particle's spin and position is $\mathbb{C}^2 \otimes \ell^2(\mathbb{Z})$. Write the basis vectors $|\uparrow\rangle \otimes |k\rangle$ and $|\downarrow\rangle \otimes |k\rangle$ as $|\uparrow k\rangle$

and $|\downarrow k\rangle$, respectively. We stipulate that the particle starts in state $|\uparrow 0\rangle$. The dynamics of the walk is given by the unitary operator

$$\begin{aligned} U : |\uparrow k\rangle &\mapsto \frac{1}{\sqrt{2}}|\uparrow k-1\rangle + \frac{1}{\sqrt{2}}|\downarrow k+1\rangle \\ |\downarrow k\rangle &\mapsto \frac{1}{\sqrt{2}}|\uparrow k-1\rangle - \frac{1}{\sqrt{2}}|\downarrow k+1\rangle \end{aligned}$$

Furthermore it is possible to measure the position of the particle, but not the spin. So the admissible observables are $|\uparrow k\rangle\langle\uparrow k| + |\downarrow k\rangle\langle\downarrow k|$ for $k \in \mathbb{Z}$.

If the particle walks during n time steps, then its position can be described using a probability distribution over $\mathbb{Z}$. The probability that we encounter the particle on position k is

$$\langle U^n(\uparrow 0)|\big(|\uparrow k\rangle\langle\uparrow k| + |\downarrow k\rangle\langle\downarrow k|\big)|U^n(\uparrow 0)\rangle.$$

Denote the probability distribution after n steps by $\varphi_n \in \mathcal{D}(\mathbb{Z})$. The unit distribution k is written using Dirac notation as $|k\rangle$. The first few probability distributions are:

$$\begin{aligned} \varphi_0 &= |0\rangle \\ \varphi_1 &= \tfrac{1}{2}|-1\rangle + \tfrac{1}{2}|1\rangle \\ \varphi_2 &= \tfrac{1}{4}|-2\rangle + \tfrac{1}{2}|0\rangle + \tfrac{1}{4}|2\rangle \\ \varphi_3 &= \tfrac{1}{8}|-3\rangle + \tfrac{5}{8}|-1\rangle + \tfrac{1}{8}|1\rangle + \tfrac{1}{8}|3\rangle \end{aligned}$$

3.2 Coalgebraic model

We would like to model the systems in Section 3.1 in such a way that the system is a coalgebra for a certain endofunctor, and the morphism into the final coalgebra gives the quantum language or probability distribution determined by the system. To achieve this, we will work with coalgebras in the category **Conv** of convex algebras. There are two reasons for this. Firstly, we can model computations for both systems in pure states and systems in mixed states. Secondly, the category **Conv** incorporates both quantum probability via density matrices and the classical probability that is needed for the output.

Let H be a Hilbert space underlying a quantum system, and let S be a set of unitary operators that can be applied to the system. Represent the possible measurements on the system by a subset E of $\mathcal{E}f(H)$. We do not use the entire set $\mathcal{E}f(H)$ since usually not all effects are possible or interesting. It is often the case that the sum of the effects in E is id. In this case, the subset E is called a *test*, see [10] for more information. Consider the functor

$$F(X) = [0,1]^E \times X^S \tag{3.1}$$

on the category **Conv**. Form the F-coalgebra

$$\begin{aligned} f : \mathcal{DM}(H) &\to [0,1]^E \times \mathcal{DM}(H)^S \\ \rho &\mapsto ((\mathrm{tr}(\rho\epsilon))_{\epsilon\in E}, (\mathcal{DM}(U)(\rho))_{U\in S}) . \end{aligned} \tag{3.2}$$

The part $\mathrm{tr}(\rho\epsilon)$ represents the observations on the coalgebra, since this is the probability that measurement of the effect ϵ succeeds when the system is in mixed state ρ. The part $\mathcal{DM}(U)(\rho)$ is the evolution of the system according to the density matrix formalism.

We will now show that (3.2) has the desired behaviour. First we have to check that the behaviour exists, i.e. that F has a final coalgebra.

The category **Conv** inherits all products from **Sets**, so Lemma 1.29 applies to the functor defined in (3.1). The functor F has a final coalgebra with the convex set $([0,1]^E)^{S^*}$ as state space. The dynamics is a map $([0,1]^E)^{S^*} \to [0,1]^E \times (([0,1]^E)^{S^*})^S$. The first component of the dynamics is the projection onto the component with the empty word ε as index, and for the second component, the map with index $a \in S$ and $u \in S^*$ is the projection onto component au.

The coalgebra (3.2) gives a behaviour map $\mathrm{beh}_f : \mathcal{DM}(H) \to ([0,1]^E)^{S^*}$. This map can alternatively be seen as a indexed family of maps $\mathcal{DM}(H) \to [0,1]$ for each $\epsilon \in E$ and $u \in S^*$. The map with index ϵ and $u = u_1 \dots u_n \in S^*$ sends the density matrix ρ to $\mathrm{tr}(u_n \dots u_1 \rho u_1^\dagger \dots u_n^\dagger \epsilon)$, as desired.

The construction of the coalgebra (3.2) involves three parameters: the underlying state space H, the set of unitary operators S, and the set of possible measurements E. By choosing these parameters appropriately we can fit the examples of Section 3.1 in this framework.

Example 3.5. We revisit Example 3.1. The Hilbert space is $\mathbb{C}^4$, there is one unitary operator U, and each vertex x_k gives an observable $|k\rangle\langle k|$, so the set E of effects is $\{|k\rangle\langle k| \mid k = 0,1,2,3\}$. The coalgebra (3.2) becomes

$$\begin{aligned} f : \mathcal{DM}(\mathbb{C}^4) &\to [0,1]^4 \times \mathcal{DM}(\mathbb{C}^4) \\ \rho &\mapsto \big((\mathrm{tr}(\rho|k\rangle\langle k|))_{k=0,1,2,3}, U\rho U^\dagger\big) . \end{aligned}$$

The codomain can be restricted to $\mathcal{D}(4) \times \mathcal{DM}(\mathbb{C}^4)$ because the effects in E sum to id. Given an initial mixed state ρ and a number of steps n, the behaviour morphism $\mathrm{beh}_f : \mathcal{DM}(\mathbb{C}^4) \to \mathcal{D}(4)^{\mathbb{N}}$ gives the probability distribution over the vertices after n steps, starting in ρ. In Example 3.1 we looked at the initial state $|0\rangle$. The corresponding mixed state is $|0\rangle\langle 0|$, so the morphism beh_f maps $|0\rangle\langle 0|$ to the list of probability distributions

$$\Big(|0\rangle, \frac{1}{2}|1\rangle + \frac{1}{2}|2\rangle, |0\rangle, \frac{1}{2}|1\rangle + \frac{1}{2}|2\rangle, \dots\Big) .$$

Examples 3.3 and 3.4 can be treated similarly. A quantum automaton $\mathbb{A} = (H, (\delta_a)_{a \in A}, \psi_0, \epsilon)$ gives a coalgebra

$$\begin{aligned} \mathcal{DM}(H) &\to [0,1] \times \mathcal{DM}(H)^A \\ \rho &\mapsto \left(\mathrm{tr}(\rho\epsilon), \left(\delta_a \rho \delta_a^\dagger\right)_{a \in A}\right) \end{aligned}$$

Here the set of unitary operators is $\{\delta_a \mid a \in A\}$ and the only possible measurement is ϵ. As in the classical case the initial state is not part of the coalgebra. The behaviour map $\mathcal{DM}(H) \to [0,1]^{A^*}$ maps $|\psi_0\rangle\langle\psi_0|$ to the quantum language recognized by the automaton.

The quantum walk can be seen as a coalgebra

$$\begin{aligned} \mathcal{DM}(\mathbb{C}^2 \otimes \ell^2(\mathbb{Z})) &\to [0,1]^{\mathbb{Z}} \times \mathcal{DM}(\mathbb{C}^2 \otimes \ell^2(\mathbb{Z})) \\ \rho &\mapsto \left((\mathrm{tr}(\rho|\uparrow k\rangle\langle\uparrow k| + \rho|\downarrow k\rangle\langle\downarrow k|))_{k \in \mathbb{Z}}, U\rho U^\dagger\right). \end{aligned}$$

This time it is impossible to restrict the first component of the codomain to $\mathcal{D}(\mathbb{Z})$, because the output probability distributions need not have finite support. We can replace it by $\mathcal{D}_\omega(\mathbb{Z})$, where $\mathcal{D}_\omega$ is the infinite distribution monad, defined on objects by

$$\mathcal{D}_\omega(X) = \{\varphi : X \to [0,1] \mid \mathrm{supp}(\varphi) \text{ is at most countable and } \sum_{x \in X} \varphi(x) = 1\}.$$

On morphisms, $\mathcal{D}_\omega$ acts the same as the ordinary distribution monad $\mathcal{D}$.

The behaviour map for the walk coalgebra assigns to the initial state $|\uparrow 0\rangle\langle\uparrow 0|$ and a number of steps n the probability distribution after n steps, so we get the list of distributions from Example 3.4.

These examples show that the representation (3.2) captures many quantum systems in a uniform way.

We shall describe the relation between quantum systems and the coalgebraic representation slightly more abstractly for automata. Fix an alphabet A. Define a category **QAut** whose objects are quantum automata without initial state. Thus an object of **QAut** consists of a Hilbert space H, a collection of unitary operators $\delta_a : H \to H$ for $a \in A$, and an effect ϵ on H. The dependence of **QAut** on the alphabet A is left implicit. A morphism in **QAut** from an automaton $(H, (\delta_a)_{a \in A}, \epsilon)$ to $(H', (\delta'_a)_{a \in A}, \epsilon')$ is an isometry $f : H \to H'$ making the following diagrams commute.

$$\begin{tikzcd} H \arrow[r, "f"] \arrow[d, "\delta_a"'] & H' \arrow[d, "\delta'_a"] \\ H \arrow[r, "f"'] & H' \end{tikzcd} \qquad \begin{tikzcd} H \arrow[r, "f"] \arrow[d, "\epsilon"'] & H' \arrow[d, "\epsilon'"] \\ H \arrow[r, "f"'] & H' \end{tikzcd}$$

There is an obvious forgetful functor $U : \mathbf{QAut} \to \mathbf{Hilb}_{\mathbf{Isomet}}$.

We modeled quantum automata as coalgebras for the functor $F(X) = [0,1] \times X^A$ on **Conv**. There is also a forgetful functor $V : \mathbf{CoAlg}(F) \to \mathbf{Conv}$. The construction that assigns to a quantum automaton its associated F-coalgebra is functorial.

Proposition 3.6. *There is a functor* Φ *making the following square in* **Cat** *commute.*

$$\begin{tikzcd} \mathbf{QAut} \arrow[r, "\Phi"] \arrow[d, "U"'] & \mathbf{CoAlg}(F) \arrow[d, "V"] \\ \mathbf{Hilb_{Isomet}} \arrow[r, "\mathcal{DM}"'] & \mathbf{Conv} \arrow[loop below, "F"] \end{tikzcd}$$

Proof. The functor Φ assigns to a quantum automaton $(H, (\delta_a)_{a \in A}, \epsilon)$ the F-coalgebra $\mathcal{DM}(H) \to [0,1] \times \mathcal{DM}(H)^A$ mapping a density matrix ρ to $(\mathrm{tr}(\rho\epsilon), (\mathcal{DM}(\delta_a)(\rho))_{a \in A})$. A morphism $f : H \to H'$ of quantum automata is mapped to the morphism $\mathcal{DM}(f)$. To verify that this is a coalgebra morphism, we have to show that $\mathrm{tr}(\mathcal{DM}(f)(\rho)\epsilon') = \mathrm{tr}(\rho\epsilon)$ for all ρ and $\mathcal{DM}(\delta'_a) \circ \mathcal{DM}(f) = \mathcal{DM}(f) \circ \mathcal{DM}(\delta_a)$. The first assertion holds because

$$\mathrm{tr}(\mathcal{DM}(f)(\rho)\epsilon') = \mathrm{tr}(f\rho f^\dagger \epsilon') = \mathrm{tr}(\rho f^\dagger \epsilon' f) = \mathrm{tr}(\rho f^\dagger f \epsilon) = \mathrm{tr}(\rho\epsilon),$$

where we used that f is an isometry. For the second claim, use that $\mathcal{DM}$ is a functor and $\delta'_a \circ f = f \circ \delta_a$. It is clear that the functor Φ makes the diagram commute. □

The behaviour of quantum automata is captured by the final object in $\mathbf{CoAlg}(F)$. It is not possible to describe the behaviour via finality in the category **QAut**, since **QAut** does not contain a final object.

3.3 Minimization of quantum coalgebras

In Section 1.5 we have described minimization of coalgebras in general. Now that we know how to view quantum systems as coalgebras, we can apply this general theory to obtain minimal realizations of quantum mechanical behaviours.

Consider coalgebras for the functor $F(X) = [0,1]^E \times X^S$, which was defined in (3.1). The final coalgebra for F has underlying state space $([0,1]^E)^{S^*}$. Since elements of a convex set X correspond with convex maps $1 \to X$, a behaviour for F is a convex map $b : 1 \to ([0,1]^E)^{S^*}$. The minimal coalgebra with behaviour b exists by Corollary 1.32 and is a subcoalgebra of $([0,1]^E)^{S^*}$. An interesting feature of this approach is that the minimal coalgebra with behaviour b need not be a quantum coalgebra, even if the behaviour arises from a quantum mechanical system. Since the subcoalgebra of the final coalgebra lies in the category **Conv**, the minimal realization will certainly be a probabilistic system, but its state space is not necessarily of the form $\mathcal{DM}(H)$. Thus minimization gives a procedure for turning quantum systems into simpler probabilistic systems which nevertheless have the same behaviour.

Example 3.7. We reconsider Example 3.1. Recall that we represented this system as the coalgebra

$$\begin{aligned} f : \mathcal{DM}(\mathbb{C}^4) &\to \mathcal{D}(4) \times \mathcal{DM}(\mathbb{C}^4) \\ \rho &\mapsto \left((\mathrm{tr}(\rho |k\rangle\langle k|))_{k=0,1,2,3}, U\rho U^\dagger \right). \end{aligned}$$

Assume that $|0\rangle\langle 0| \in \mathcal{DM}(\mathbb{C}^4)$ is the initial state. The resulting output stream is

$$\sigma = \left(|0\rangle, \frac{1}{2}|1\rangle + \frac{1}{2}|2\rangle \right)^\omega.$$

The minimal coalgebra with this behaviour is the smallest subcoalgebra of $\mathcal{D}(4)^\mathbb{N}$ containing σ. We shall compute this subcoalgebra by determining the states of $\mathcal{D}(4)^\mathbb{N}$ that can be reached from σ with the transition structure, and then taking the convex set generated by the reachable states.

The transition structure of the final coalgebra is given by the function $\mathrm{tail} : \mathcal{D}(4)^\mathbb{N} \to \mathcal{D}(4)^\mathbb{N}$. Applying the tail function repeatedly to σ gives the streams σ and $\sigma' = \left(\frac{1}{2}|1\rangle + \frac{1}{2}|2\rangle, |0\rangle\right)^\omega$. The convex algebra generated by σ and σ' inside $\mathcal{D}(4)^\mathbb{N}$ is

$$X = \left\{ \left(p|0\rangle + \tfrac{1}{2}(1-p)|1\rangle + \tfrac{1}{2}(1-p)|2\rangle, (1-p)|0\rangle + \tfrac{1}{2}p|1\rangle + \tfrac{1}{2}p|2\rangle \right)^\omega \mid p \in [0,1] \right\}.$$

This is the minimal coalgebra with behaviour σ and hence the minimization of f. The coalgebra structure is obtained as restriction of the final coalgebra $\mathcal{D}(4)^\mathbb{N}$. By projecting onto the first coordinate twice we obtain a convex isomorphism between X and $[0,1]$. Therefore a more elementary display of the minimization is

$$\begin{aligned} [0,1] &\to \mathcal{D}(4) \times [0,1] \\ p &\mapsto \left(p|0\rangle + \tfrac{1}{2}(1-p)|1\rangle + \tfrac{1}{2}(1-p)|2\rangle, 1-p \right) \end{aligned}$$

Observe that the state space $[0,1]$ of the minimization is more manageable than the original state space $\mathcal{DM}(\mathbb{C}^4)$. The minimization is not a quantum system anymore, since $[0,1]$ is not isomorphic to a convex set of density matrices. It can be seen as a probabilistic system with two states, since $[0,1] \cong \mathcal{D}(2)$. Thus the behaviour of this quantum mechanical system can be reproduced with a simpler probabilistic system.

We usually ignore initial states when working with coalgebras, in which case reduction is a more appropriate procedure for simplifying systems. Determining the reduction of the coalgebra in the above example is a more difficult task than determining the minimization, due to the subtle structure of $\mathcal{DM}(\mathbb{C}^4)$. Therefore we will discuss an easier, two-dimensional example. It will be useful to have an explicit description of the density matrices on the Hilbert space $\mathbb{C}^2$. The positivity condition of density matrices is sometimes difficult to check, so before giving the explicit description we will give a characterization of positive operators.

Lemma 3.8. *A self-adjoint operator A on a finite-dimensional Hilbert space H is positive if and only if all eigenvalues of A are non-negative.*

Proof. Suppose that A is positive, and let λ be an eigenvalue of A with eigenvector x. Then $0 \leq \langle x|A|x\rangle = \langle x|\lambda|x\rangle = \lambda\langle x|x\rangle$, hence $\lambda \geq 0$.

Conversely, suppose that all eigenvalues of A are non-negative. Let (e_i) be an orthonormal basis of eigenvectors of A, with corresponding eigenvalues λ_i. Then an arbitrary vector $x \in H$ can be written as linear combination $x = \sum_i c_i e_i$. Therefore

$$\begin{aligned}\langle x|A|x\rangle = \langle \sum_i c_i e_i|A|\sum_j c_j e_j\rangle = \langle \sum_i c_i e_i|\sum_j c_j A e_j\rangle &= \langle \sum_i c_i e_i|\sum_j c_j \lambda_j e_j\rangle \\ &= \sum_i |c_i|^2 \lambda_i \geq 0,\end{aligned}$$

where we used that (e_i) is an orthonormal basis. □

Lemma 3.9. *The matrix $\rho = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is a density matrix on $\mathbb{C}^2$ if and only if the following conditions hold:*

1. $a + d = 1$
2. $c = \overline{b}$
3. $a \in [0, 1]$
4. $|b|^2 \leq a(1-a)$

Proof. Suppose that ρ is a density matrix. Condition 1 follows from $\mathrm{tr}(\rho) = 1$ and condition 2 holds because ρ is self-adjoint. Since ρ is positive, all of its eigenvalues are non-negative. Calculating the eigenvalues ρ, using conditions 1 and 2, gives the eigenvalues $\lambda_\pm = \frac{1}{2}(1 \pm \sqrt{1 - 4(a(1-a) - |b|^2)})$. Since $\lambda_- \geq 0$, we get $\sqrt{1 - 4(a(1-a) - |b|^2)} \leq 1$, so $|b|^2 \leq a(1-a)$, which is condition 4. Finally, for condition 3, use that condition 4 implies $a(1-a) \geq 0$, so either $a \geq 0$ and $1 - a \geq 0$, or $a \leq 0$ and $1 - a \leq 0$. In the first case it follows that $a \in [0, 1]$, and the second case is impossible.

Conversely, suppose that ρ is of the form $\left(\begin{smallmatrix} a & b \\ \overline{b} & 1-a \end{smallmatrix}\right)$ with $a \in [0, 1]$ and $|b|^2 \leq a(1-a)$. It is immediate that ρ is self-adjoint and has trace 1. To confirm positivity, we use again that the eigenvalues of ρ are $\lambda_\pm$. The number a lies in $[0, 1]$, so $a(1-a) \in [0, \frac{1}{4}]$, hence the expression $1 - 4(a(1-a) - |b|^2)$

below the square root is always non-negative. Thus $\lambda_+ \geq 0$. Furthermore, from $|b|^2 \leq a(1-a)$ it follows that $1 - 4(a(1-a) - |b|^2) \leq 1$, and therefore λ_- is also non-negative. Both eigenvalues of ρ are non-negative, so ρ is positive. □

Example 3.10. We will compute the reduction for a system with underlying Hilbert space $\mathbb{C}^2$. The dynamics is given by the Hadamard matrix

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

and for the measurement we take the projection $|0\rangle\langle 0|$. This gives a coalgebra

$$\begin{aligned} c : \mathcal{DM}(\mathbb{C}^2) &\to [0,1] \times \mathcal{DM}(\mathbb{C}^2) \\ \rho &\mapsto \left(\rho_{11}, U\rho U^\dagger\right). \end{aligned}$$

The behaviour morphism is

$$\begin{aligned} \mathrm{beh}_c : \mathcal{DM}(\mathbb{C}^2) &\to [0,1]^{\mathbb{N}} \\ \rho &\mapsto \left(\rho_{11}, (U\rho U^\dagger)_{11}\right)^{\omega}. \end{aligned}$$

The state space of the reduction is the image of beh_c, so we have to find all possible values of $\mathrm{beh}_c(\rho)$ for $\rho \in \mathcal{DM}(\mathbb{C}^2)$. An arbitrary density matrix ρ in $\mathcal{DM}(\mathbb{C}^2)$ can be written as $\left(\begin{smallmatrix} a & b \\ \overline{b} & 1-a \end{smallmatrix}\right)$. This matrix satisfies $\rho_{11} = a$ and $(U\rho U^\dagger)_{11} = \frac{1}{2}(b + \overline{b} + 1) = \Re(b) + \frac{1}{2}$, where $\Re(b)$ denotes the real part of the complex number b. From Lemma 3.9 it follows that $a \in [0,1]$ and $\Re(b) \in [-\sqrt{a(1-a)}, \sqrt{a(1-a)}]$. All of these values are indeed attained: take for b a real number in the interval $[-\sqrt{a(1-a)}, \sqrt{a(1-a)}]$, then this gives a density matrix by the explicit description in Lemma 3.9. Thus

$$\mathrm{im}(\mathrm{beh}_c) = \left\{ (a, b + \tfrac{1}{2}) \mid a \in [0,1], b \in [-\sqrt{a(1-a)}, \sqrt{a(1-a)}] \right\}.$$

Lemma 3.9 specializes to a description of density matrices on the real vector space $\mathbb{R}^2$: these are precisely the real matrices $\left(\begin{smallmatrix} a & b \\ b & 1-a \end{smallmatrix}\right)$ such that $a \in [0,1]$ and $b \in [-\sqrt{a(1-a)}, \sqrt{a(1-a)}]$. It follows that the image $\mathrm{im}(\mathrm{beh}_c)$ is isomorphic to $\mathcal{DM}(\mathbb{R}^2)$. We conclude that the state space of the reduction is the real part of the state space of the original coalgebra.

3.4 The Heisenberg picture

The coalgebras in Section 3.2 are based on the Schrödinger picture of quantum mechanics: the states are density matrices that evolve unitarily, while the observables are constant. It is also possible to form an abstract representation in Heisenberg style. For this the observables should be variable, which can be represented via morphisms in **EMod**. There are two ways to connect both pictures of quantum mechanics, corresponding to the two ways of lifting the dual adjunction between **Conv** and **EMod** from Corollary 1.13:

$$\begin{tikzcd} \mathbf{Alg}(F) \arrow[r, shift left] \arrow[r, phantom, "\perp"] \arrow[d] & \mathbf{CoAlg}(G)^{\mathrm{op}} \arrow[l, shift left] \arrow[d] \\ \mathbf{Conv} \arrow[r, shift left] \arrow[r, phantom, "\perp"] \arrow[loop below, "F"] & \mathbf{EMod}^{\mathrm{op}} \arrow[l, shift left] \arrow[loop below, "G^{\mathrm{op}}"] \end{tikzcd} \qquad \begin{tikzcd} \mathbf{CoAlg}(F) \arrow[r, shift left] \arrow[r, phantom, "\perp"] \arrow[d] & \mathbf{Alg}(G)^{\mathrm{op}} \arrow[l, shift left] \arrow[d] \\ \mathbf{Conv} \arrow[r, shift left] \arrow[r, phantom, "\perp"] \arrow[loop below, "F"] & \mathbf{EMod}^{\mathrm{op}} \arrow[l, shift left] \arrow[loop below, "G^{\mathrm{op}}"] \end{tikzcd}$$

In both cases we have to choose suitable functors for F and G. We will first describe the situation with coalgebras in **Conv** and algebras in **EMod**, as in the diagram on the right. For simplicity we will restrict to quantum automata, instead of general quantum systems. Let F be the functor on **Conv** given by $F(X) = [0,1] \times X^A$ for a fixed set A. To satisfy the condition in Corollary 1.13, we have to find a functor G on **EMod** such that $F(\mathbf{EMod}(X,[0,1]))$ is naturally isomorphic to $\mathbf{EMod}(G(X),[0,1])$. This will be the content of Lemma 3.12.

Lemma 3.11. *The free effect module on 1 generator is* $[0,1]^2 = [0,1] \times [0,1]$.

Proof. Let E be an arbitrary effect module. We have to establish a bijective correspondence between points in E and effect module homomorphisms $[0,1]^2 \to E$. Define the map $\Phi : \mathbf{EMod}([0,1]^2, E) \to E$ by $\Phi(f) = f(1,0)$. To prove injectivity of Φ, suppose that $\Phi(f) = \Phi(g)$. Then, for all $r, s \in [0,1]$,

$$\begin{aligned} f(r,s) = f(r(1,0) \boxplus s(1,0)^\perp) = rf(1,0) \boxplus sf(1,0)^\perp &= rg(1,0) \boxplus sg(1,0)^\perp \\ &= g(r,s). \end{aligned}$$

Therefore $f = g$. To establish surjectivity, pick an arbitrary $x \in E$. Define $f : [0,1]^2 \to E$ by $f(r,s) = rx \boxplus sx^\perp$. Since f preserves addition, scalar multiplication, and the constants 0 and 1, it is an effect module homomorphism. Furthermore $\Phi(f) = f(1,0) = x$, finishing the proof that Φ is an bijection. $\square$

Lemma 3.12. *Let X be a convex set and A a set. Then the convex sets* $\mathbf{EMod}([0,1]^2 + A \cdot X, [0,1])$ *and* $[0,1] \times \mathbf{EMod}(X,[0,1])^A$ *are isomorphic, naturally in X. Here, + denotes the coproduct in* **EMod** *and $A \cdot X$ represents the coproduct of A copies of X.*

Proof. By general results on sets of morphisms,

$$\mathbf{EMod}([0,1]^2 + A \cdot X, [0,1]) \cong \mathbf{EMod}([0,1]^2, [0,1]) \times \mathbf{EMod}(X,[0,1])^A.$$

Hence it suffices to show that $\mathbf{EMod}([0,1]^2, [0,1])$ is isomorphic to $[0,1]$ as a convex set. Lemma 3.11 asserts that these are isomorphic as sets, and it is obvious that the map Φ produced in the proof of that Lemma is a convex map. □

Let G be the endofunctor $G(X) = [0,1]^2 + A \cdot X$ on **EMod**. Then

$$F(\mathbf{EMod}(X,[0,1])) \cong \mathbf{EMod}(G(X),[0,1]),$$

so we can apply Corollary 1.13 to obtain an adjunction between $\mathbf{CoAlg}(F)$ and $\mathbf{Alg}(G)$.

Example 3.13. In Example 3.3 we represented an automaton as an F-coalgebra in **Conv**. Here we will determine the dual algebra in **EMod**. The coalgebra in **Conv** has type $\mathcal{DM}(\mathbb{C}^2) \to [0,1] \times \mathcal{DM}(\mathbb{C}^2)^A$, so since $\mathbf{Conv}(\mathcal{DM}(\mathbb{C}^2),[0,1]) \cong \mathcal{E}f(\mathbb{C}^2)$, the dual algebra has type $[0,1]^2 + A \cdot \mathcal{E}f(\mathbb{C}^2) \to \mathcal{E}f(\mathbb{C}^2)$. It can be split in an effect module homomorphism $[0,1]^2 \to \mathcal{E}f(\mathbb{C}^2)$ and an A-indexed family of maps on $\mathcal{E}f(\mathbb{C}^2)$. The first map is obtained from the output map of the coalgebra as follows: since $[0,1]^2$ is the free effect module on one generator, the morphism $[0,1]^2 \to \mathcal{E}f(\mathbb{C}^2)$ corresponds to an element ϵ_0 of $\mathcal{E}f(\mathbb{C}^2)$. The first component of the F-coalgebra is $\rho \mapsto \mathrm{tr}(\rho|0\rangle\langle 0|)$. Using the isomorphism between $\mathcal{E}f(\mathbb{C}^2)$ and $\mathbf{Conv}(\mathcal{DM}(\mathbb{C}^2),[0,1])$, we see that the effect ϵ_0 is $|0\rangle\langle 0|$. For the second component of the algebra, the map $\mathcal{E}f(\mathbb{C}^2) \to \mathcal{E}f(\mathbb{C}^2)$ indexed by $a \in A$ is the dual of the component $\mathcal{DM}(\mathbb{C}^2) \to \mathcal{DM}(\mathbb{C}^2)$ of the coalgebra indexed by a. Concretely, since the coalgebra component is given by $\rho \mapsto \mathcal{DM}(\delta_a)(\rho) = \delta_a \rho \delta_a^\dagger$, the algebra component is $\epsilon \mapsto \mathcal{E}f(\delta_a)(\epsilon) = \delta_a^\dagger \epsilon \delta_a$.

The final F-coalgebra has underlying convex set $[0,1]^{A^*}$, and the initial G-algebra has underlying effect module $A^* \cdot [0,1]^2$. The unique algebra morphism $A^* \cdot [0,1]^2 \to \mathcal{E}f(\mathbb{C}^2)$ to the algebra described above can be identified with a family of effects on $\mathcal{E}f(\mathbb{C}^2)$, indexed by A^*. The effect with index $u \in A^*$ is $\delta_u^\dagger |0\rangle\langle 0| \delta_u$. Since right adjoints preserve final objects, the dual of this algebra morphism is the behaviour morphism $\mathcal{DM}(\mathbb{C}^2) \to [0,1]^{A^*}$.

Now we will look at the second way to relate the Schrödinger and the Heisenberg picture, which involves coalgebras in **EMod** and emphasizes the dynamics of the effects. First we will define the shape of the coalgebras in **EMod**, in such a way that it corresponds to the dynamics given by the Heisenberg picture of quantum mechanics.

As before, we assume a Hilbert space H and a set of unitaries S. In the Schrödinger representation, we furthermore needed a set of possible measurements. Dually, for the Heisenberg representation we will require a subset D of $\mathcal{DM}(H)$, which can be thought of as the possible preparations of the system. It is often inconvenient to include all density matrices in the resulting coalgebra, so we restrict to a subset. Using these ingredients, form the functor $G(X) = [0,1]^D \times X^S$ on **EMod** and the G-coalgebra

$$\begin{aligned} g : \mathcal{E}f(H) &\to [0,1]^D \times \mathcal{E}f(H)^S \\ \epsilon &\mapsto \left((\mathrm{tr}(\rho\epsilon))_{\rho \in D}, (\mathcal{E}f(U)(\epsilon))_{U \in S} \right). \end{aligned} \tag{3.3}$$

Recall that $\mathcal{E}f$ is a contravariant functor from $\mathbf{Hilb_{Isomet}}$ to **EMod**, so that the effects evolve backwards. From Lemma 1.29 we obtain a unique effect module homomorphism into the final coalgebra $\mathcal{E}f(H) \to ([0,1]^D)^{S^*}$.

Example 3.14. In Section 3.2 we saw how to model automata in the Schrödinger picture. Here we will express the automaton from Example 3.3 in the Heisenberg framework above. The initial state, in this case $|0\rangle\langle 0|$, will be the only possible preparation. The coalgebra (3.3) becomes

$$\begin{aligned} \mathcal{E}f(\mathbb{C}^2) &\to [0,1] \times \mathcal{E}f(\mathbb{C}^2)^{\{a,b\}} \\ \epsilon &\mapsto \Big(\mathrm{tr}(|0\rangle\langle 0|\epsilon), (\delta_a^\dagger \epsilon \delta_a, \delta_b^\dagger \epsilon \delta_b)\Big). \end{aligned}$$

This results in a behaviour map $\mathcal{E}f(\mathbb{C}^2) \to [0,1]^{\{a,b\}^*}$ sending an effect ϵ to the reverse of the language recognized by the automaton with ϵ as output. The reversal occurs because the effects "move backwards" under the dynamics of the coalgebra. We will elaborate on this later.

Using a similar technique as above, we can define a functor F on **Conv** such that there is a dual adjunction between F-algebras and G-coalgebras. But it is also interesting to establish a connection between the coalgebras defined by (3.2) and those defined by (3.3). Recall that the Schrödinger coalgebra (3.2) has type $\mathcal{DM}(H) \to [0,1]^E \times \mathcal{DM}(H)^S$ and that the Heisenberg coalgebra (3.3) has type $\mathcal{E}f(H) \to [0,1]^D \times \mathcal{E}f(H)^S$. Intuitively it is clear that both coalgebras represent the same physical phenomena. In the Schrödinger coalgebra, the state moves forward and is finally measured using a constant effect. In the Heisenberg coalgebra, the effect moves backward and is finally used as a measurement on a constant state. More formally we have a result relating the behaviours of both coalgebras.

Proposition 3.15. *Consider the coalgebras defined in* (3.2) *and* (3.3)*. Take a preparation $\rho \in D$ and an effect $\epsilon \in E$. Then, for each word $u \in S^*$, the component of* $\mathrm{beh}_f(\rho)$ *indexed with u and ϵ equals the component of* $\mathrm{beh}_g(\epsilon)$ *indexed with u^R and ρ.*

Proof. The component of $\mathrm{beh}_f(\rho)$ with indices $u = u_1 \dots u_n$ and ϵ is

$$\mathrm{tr}(u_n \dots u_1 \rho u_1^\dagger \dots u_n^\dagger \epsilon).$$

The component of $\mathrm{beh}_g(\epsilon)$ with indices $u^R = u_n \dots u_1$ and ρ is

$$\mathrm{tr}(\rho u_1^\dagger \dots u_n^\dagger \epsilon u_n \dots u_1).$$

These two numbers are equal by the cyclic property of the trace. □

Thus for a fixed initial state and a fixed output effect, we can construct two automata, one in **Conv**, one in **EMod**. Then the above result states that the probabilistic languages recognized by these automata are each other's reverse.

3.5 Generalized duality

In Section 3.4 we saw how the two pictures of quantum dynamics are related via the dual adjunction between **Conv** and **EMod**. If one of the viewpoints is modeled using algebras and the other using coalgebras, then the relation amounts to the adjunction lifting property from Corollary 1.13. We will now explore which categorical structure is needed to obtain the duality between coalgebras in **Conv** and coalgebras in **EMod**, as described in Proposition 3.15. This duality can be generalized: it will turn out that for any dual adjunction satisfying certain properties, the coalgebras on both sides of the adjunction can be connected. Specializing this connection to the duality between **Conv** and **EMod** gives the equivalence of the Schrödinger and the Heisenberg picture. Unfortunately the construction in this section works only for functors of a special shape.

The setting of this section is as follows. Assume that we have three categories **B**, **C**, and **D**. The category **B** is thought of as a base category, which is often **Sets**. Furthermore, **C** is the category of state spaces, and **D** the category of spaces of observables. We stipulate that the three categories are related via the adjunctions

$$
\begin{tikzcd}
\mathbf{D}^{\mathrm{op}} \arrow[d, bend left, "G"] \\
\mathbf{C} \arrow[u, bend left, "F"] \arrow[u, phantom, "\dashv"]
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{C} \arrow[d, bend left, "U"] \\
\mathbf{B} \arrow[u, bend left, "\Phi"] \arrow[u, phantom, "\dashv"]
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{D} \arrow[d, bend left, "U'"] \\
\mathbf{B} \arrow[u, bend left, "\Phi'"] \arrow[u, phantom, "\dashv"]
\end{tikzcd}
$$

A system with states in **C** and observables in **D** is determined by four parameters. Let $C \in \mathbf{C}$ be the state space. Take a morphism $d : D \to U(C)$ in **B** representing the possible preparations and a morphism $e : E \to U'F(C)$ in **B** representing the possible measurements. Finally S is a set of morphisms $C \to C$, standing for the dynamics. We assume that the data specified here satisfies some requirements.

Assumption 3.16.

1. The categories $\mathbb{C}$ and $\mathbb{D}$ have all products.

2. Let $\eta : \mathrm{Id} \Rightarrow GF$ be the unit of the adjunction $F \dashv G$. The morphism $\eta_C : C \to GF(C)$ is an isomorphism in **C**, for the particular object C.

Under this assumption, we can form two coalgebras. Firstly, construct the following coalgebra in **C**:

$$f = \langle f_1, \langle u \rangle_{u \in S} \rangle : C \to G\Phi'(E) \times C^S. \tag{3.4}$$

The morphism f_1 is obtained from the measurement morphism e using the adjunction correspondences:

$$\begin{array}{cc}
E \xrightarrow{e} U'F(C) & \text{in } \mathbf{B} \\
\hline\hline
\Phi'(E) \longrightarrow F(C) & \text{in } \mathbf{D} \\
\hline\hline
F(C) \longrightarrow \Phi'(E) & \text{in } \mathbf{D}^{\mathrm{op}} \\
\hline\hline
C \xrightarrow[f_1]{} G\Phi'(E) & \text{in } \mathbf{C}
\end{array}$$

Secondly, there is a coalgebra

$$g = \langle g_1, \langle F(u)\rangle_{u\in S}\rangle : F(C) \to F\Phi(D) \times F(C)^S \tag{3.5}$$

in $\mathbf{D}$. The morphism g_1 is constructed by applying F to the adjoint transpose of d.

By Lemma 1.29 and Assumption 3.16 we obtain behaviour morphisms $\mathrm{beh}_f : C \to (G\Phi'(E))^{S^*}$ and $\mathrm{beh}_g : F(C) \to (F\Phi(D))^{S^*}$. The components of beh_f and beh_g with index $u \in S^*$ are denoted $(\mathrm{beh}_f)_u$ and $(\mathrm{beh}_g)_u$, respectively. The behaviour morphisms are connected in the following way.

Proposition 3.17. *Under Assumption 3.16, the following diagram commutes for each $u \in S^*$.*

$$\begin{tikzcd}
\Phi(D) \arrow[r, "\overline{(\mathrm{beh}_g)_{u^R}}"] \arrow[d, "\overline{d}"'] & GF(C) \arrow[r, "\eta_C^{-1}", "\cong"'] & C \arrow[d, "f_1"] \\
C \arrow[rr, "(\mathrm{beh}_f)_u"'] & & G\Phi'(E)
\end{tikzcd}$$

Some remarks about the notation: a bar over a morphism denotes the adjoint transpose, where the adjunction is left implicit. Recall that u^R denotes the reversal of the word u.

Proof. Write the word u as $u_1 \dots u_n$ with $u_i \in S$. The proof of Lemma 1.29 provides an explicit description of the behaviour morphisms:

$$(\mathrm{beh}_f)_u = f_1 \circ u_n \circ \dots \circ u_1,$$

$$(\mathrm{beh}_g)_{u^R} = g_1 \circ F(u_1) \circ \dots \circ F(u_n).$$

Using $g_1 = F(\overline{d})$ and the fact that F is a contravariant functor, we get $(\mathrm{beh}_g)_{u^R} = F(u_n \circ \dots \circ u_1 \circ \overline{d})$. Therefore we wish to prove that

$$f_1 \circ u_n \circ u_1 \circ \overline{d} = f_1 \circ \eta_C^{-1} \circ \overline{F(u_n \circ \dots \circ u_1 \circ \overline{d})}.$$

For this it suffices to show

$$\eta_C \circ u_n \circ \dots \circ u_1 \circ \overline{d} = \overline{F(u_n \circ \dots \circ u_1 \circ \overline{d})}.$$

By naturality of η, $\eta_C \circ u_n \circ \dots \circ u_1 \circ \overline{d} = GF(u_n \circ \dots \circ u_1 \circ \overline{d})$. This equals $\overline{F(u_n \circ \dots \circ u_1 \circ \overline{d})}$ by a general property of adjunctions. □

Example 3.18. First we will check that the equivalence of the Schrödinger and the Heisenberg representation is an instance of Proposition 3.17. Take the categories $\mathbf{B} = \mathbf{Sets}$, $\mathbf{C} = \mathbf{Conv}$, and $\mathbf{D} = \mathbf{EMod}$. The categories **Conv** and **EMod** have all products. In Theorem 2.28 we established an adjunction between **Conv** and **EMod**. Furthermore the forgetful functors $\mathbf{Conv} \to \mathbf{Sets}$ and $\mathbf{EMod} \to \mathbf{Sets}$ have left adjoints. For the functor $\mathbf{Conv} \to \mathbf{Sets}$ this is true because **Conv** is the category of algebras for the distribution monad on **Sets**. The left adjoint of the forgetful functor $\mathbf{EMod} \to \mathbf{Sets}$ can be understood as the composition of two free constructions:

$$\begin{tikzcd} \mathbf{Sets} \arrow[r, bend left, "MO"] \arrow[r, phantom, "\perp"] & \mathbf{EA} \arrow[l, bend left] \arrow[r, bend left, "{[0,1]\otimes -}"] \arrow[r, phantom, "\perp"] & \mathbf{EMod} \arrow[l, bend left] \end{tikzcd}$$

The free constructions were described explicitly in Examples 2.23 and 2.25.

If H is a finite-dimensional Hilbert space, then the unit η of the adjunction between **Conv** and **EMod** is an isomorphism on $\mathcal{DM}(H)$, because

$$\mathbf{EMod}(\mathbf{Conv}(\mathcal{DM}(H),[0,1]),[0,1]) \cong \mathbf{EMod}(\mathcal{E}f(H),[0,1]) \cong \mathcal{DM}(H).$$

Hence we can take $\mathcal{DM}(H)$ as the state space. The corresponding space of observables is $\mathbf{Conv}(\mathcal{DM}(H),[0,1]) \cong \mathcal{E}f(H)$. We assume a set of possible preparations $D \subseteq \mathcal{DM}(H)$, a set of measurements $E \subseteq \mathcal{E}f(H)$, and a set of unitary operators S on H. The morphisms d and e are the inclusions $D \hookrightarrow U(\mathcal{DM}(H))$ and $E \hookrightarrow U'(\mathcal{E}f(H))$. The set of unitary operators S gives a set of endomorphisms on $\mathcal{DM}(H)$ by applying the functor $\mathcal{DM}$ to each operator in S.

Filling in this data in (3.4) gives a coalgebra of type

$$\mathcal{DM}(H) \to \mathbf{EMod}(\Phi'(E),[0,1]) \times \mathcal{DM}(H)^S.$$

Since $\Phi'(E)$ is a free effect module, the convex set $\mathbf{EMod}(\Phi'(E),[0,1]$ is isomorphic to $[0,1]^E$. The inclusion $e : E \hookrightarrow U'(\mathcal{E}f(H))$ gives a morphism $f_1 : \mathcal{DM}(H) \to \mathbf{EMod}(\Phi'(E)) \cong [0,1]^E$ given by $f_1(\rho) = (\mathrm{tr}(\rho\epsilon))_{\epsilon\in E}$. Thus the coalgebra (3.4) coincides with the Schrödinger representation of quantum systems defined in (3.2). Similarly it can be checked that the coalgebras (3.5) and (3.3) agree with the above choice of parameters.

We now turn our attention to the meaning of the commuting diagram in Proposition 3.17. It states that the morphisms $f_1 \circ \eta^{-1} \circ \overline{(\mathrm{beh}_g)_{u^R}}, (\mathrm{beh}_f)_u \circ \overline{d} :$ $\Phi(D) \to \mathbf{EMod}(\Phi'(E),[0,1]) \cong [0,1]^E$ are equal, so their adjoint transposes $D \to [0,1]^E$ in **Sets** are also equal. The transpose of $f_1 \circ \eta^{-1} \circ \overline{(\mathrm{beh}_g)_{u^R}}$ maps $\rho \in D$ to $\big(\mathrm{beh}_g(\epsilon)(u^R)(\rho)\big)_{\epsilon\in E}$. Here we identified the power $([0,1]^D)^{S^*}$ in the codomain of beh_g with the set of functions $S^* \to (D \to [0,1])$. The transpose of $(\mathrm{beh}_f)_u \circ \overline{d}$ maps $\rho \in D$ to $\big(\mathrm{beh}_f(\rho)(u)(\epsilon)\big)$. Saying that both morphisms are equal is exactly the statement of Proposition 3.15.

Example 3.19. We will apply the above framework to deterministic finite automata. This leads to two different views on automata: the first view emphasizes the states, and the second one emphasizes the output set.

We work again over the base category **Sets**, but now we use the duality between sets and Boolean algebras for the adjunction between **C** and **D**, see Example 1.14. For the other two adjunctions we take the identity adjunction on **Sets** and the free Boolean algebra adjunction between **Sets** and **BA**. The first criterion of Assumption 3.16 is satisfied, since **Sets** and **BA** have products.

Now we will specify the states, preparations, measurements, and dynamics. Let $\mathbb{A} = (X, \delta, x_0, U)$ be a deterministic automaton over the alphabet A, and assume that X is a finite set. Let X act as state space. The ultrafilters of $\mathcal{P}(X)$ are in bijective correspondence with the elements of X, since X is finite. Hence the unit η_X is an isomorphism, as desired by the second criterion of Assumption 3.16. For the possible preparations we take the inclusion $\{x_0\} \hookrightarrow X$, and the inclusion $\{U\} \hookrightarrow \mathcal{P}(X)$ assumes the role of the measurements. The set of morphisms $\{\delta_a : X \to X \mid a \in A\}$ gives the dynamics.

This choice of parameters brings about two coalgebras. The coalgebra f in **Sets** has type $X \to \mathcal{UF}(\Phi'(\{U\})) \times X^A$, where $\Phi'(\{U\})$ is the free Boolean algebra generated by the singleton set $\{U\}$. Explicitly, it is the Boolean algebra $\mathcal{P}(2)$. Since $\mathcal{P}(2)$ has two ultrafilters, we can simplify the type to $X \to 2 \times X^A$. The coalgebra f maps $x \in X$ to $(\chi_U(x), (\delta_a(x))_{a \in A})$, which is the standard coalgebraic description of automata.

The type of the other coalgebra g in **BA** is $\mathcal{P}(X) \to \mathcal{P}(\{x_0\}) \times \mathcal{P}(X)^A \cong 2 \times \mathcal{P}(X)^A$. Its first component maps $S \in \mathcal{P}(X)$ to 1 if $x_0 \in S$, and to 0 otherwise. The second component is $S \to (\delta_a^{-1}(S))_{a \in A}$. The behaviour morphism $\mathrm{beh}_g : \mathcal{P}(X) \to 2^{A^*}$ can be thought of as starting in a subset of accepting states S, pulling this subset back along the transitions, and finishing by checking if x_0 is in the subset. This account is analogous to the Heisenberg picture in quantum mechanics, i.e. the state x_0 is constant and the subset of accepting states varies. We expect the behaviour of g to be equivalent to the behaviour of f. This equivalence is embodied in Proposition 3.17. The diagram in this Proposition states the equality of two morphisms $\{x_0\} \to \mathcal{UF}(\Phi'(\{U\})) \cong 2$. The morphism $f_1 \circ \eta^{-1} \circ \overline{(\mathrm{beh}_g)_{u^R}}$ maps x_0 to 1 if and only if $x_0 \in \delta_u^{-1}(U)$. The morphism $(\mathrm{beh}_f)_u \circ \overline{d}$ maps x_0 to 1 if and only if $\delta_u(x_0) \in U$. Thus both coalgebraic descriptions of automata are indeed equivalent.

Conclusion

Quantum systems can be represented by using density matrices as states, with unitary operators acting on them as dynamics. We have exploited this fact to model quantum systems as coalgebras in the category **Conv** of convex sets, with a set of density matrices as state space. We have looked at two consequences of this representation. Firstly, minimization of coalgebras gives a method to transform a quantum system into a probabilistic system that has the same behaviour, but a simpler structure. Secondly, the duality between algebras and coalgebras gives a connection between the Schrödinger and the Heisenberg picture of quantum mechanics. The equivalence of these pictures can also be seen as an instance of the connection between two types of coalgebras, described in Proposition 3.17.

There are several possibilities for future research. The duality between convex sets and effect modules is only worked out completely for finite-dimensional systems, because Theorems 2.29 and 2.30 are only formulated for finite-dimensional Hilbert spaces. Generalization to infinite-dimensional systems would for instance put the quantum walk from Example 3.4 in our abstract framework. Furthermore, we have only studied minimization of quantum systems empirically through examples. It would be nice to have more general results about the structure of minimal realizations. Moreover it is unclear if quantum systems can also be modeled in such a way that minimization gives a quantum coalgebra again.

Bibliography

[1] S. Abramsky. Coalgebras, Chu spaces, and representations of physical systems. In *Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science*, pages 411–420, 2010.

[2] S. Abramsky and B. Coecke. A categorical semantics of quantum protocols. In *Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science*, pages 415–425, 2004.

[3] J. Adámek, F. Bonchi, M. Hülsbusch, B. König, S. Milius, and A. Silva. A coalgebraic perspective on minimization and determinization. In Lars Birkedal, editor, *Foundations of Software Science and Computational Structures*, volume 7213 of *Lecture Notes in Computer Science*, pages 58–73. Springer Berlin Heidelberg.

[4] J. Adámek and V. Koubek. On the greatest fixed point of a set functor. *Theoretical Computer Science*, 150(1):57–75, 1995.

[5] P. Busch. Quantum states and generalized observables: a simple proof of Gleason's theorem. *Phys. Rev. Lett.*, 91(12), 2003.

[6] J. Butterfield and C. Isham. A topos perspective on the Kochen–Specker theorem: I. Quantum states as generalized valuations. *Int. Jour. Theor. Physics*, 37(11):2669–2733, 1998.

[7] C. Cohen-Tannoudji. *Quantum Mechanics*. John Wiley, 1977.

[8] D. Deutsch. Quantum theory, the Church–Turing principle and the universal quantum computer. *Proc. R. Soc. Lond. A*, 400(1818):97–117, 1985.

[9] D. Deutsch and R. Jozsa. Rapid solution of problems by quantum computation. *Proc. R. Soc. Lond. A*, 439(1907):553–558, 1992.

[10] A. Dvurečenskij and S. Pulmannová. *New Trends in Quantum Structures*. Kluwer Academic Publishers, 2000.

[11] S. Eilenberg. *Automata, Languages, and Machines, Volume A*. Academic Press, 1974.

[12] D. Foulis and M. Bennett. Effect algebras and unsharp quantum logics. *Foundations of Physics*, 24(10), 1994.

[13] J. Goguen. Minimal realizations of machines in closed categories. *Bull. Amer. Math. Soc.*, 78(5):777–783, 1972.

[14] M. de Gosson. *Symplectic Geometry and Quantum Mechanics*. Birkhäuser Verlag, 2006.

[15] S. Gudder and S. Pulmannová. Representation theorem for convex effect algebras. *Commentationes Mathematicae Universitatis Carolinae*, 39(4):645–659, 1998.

[16] I. Hasuo, B. Jacobs, and A. Sokolova. Generic trace semantics via coinduction. *Logical Methods in Computer Science*, 3(4:11):1–36, 2007.

[17] B. Jacobs. Convexity, duality and effects. In Cristian Calude and Vladimiro Sassone, editors, *Theoretical Computer Science*, IFIP Advances in Information and Communication Technology, pages 1–19. Springer Boston, 2010.

[18] B. Jacobs. Coalgebraic walks, in quantum and Turing computing. In *Lecture Notes in Computer Science*, volume 6604 of *Foundations of Software Science and Computational Structures*, pages 12–26, 2011.

[19] B. Jacobs. *Introduction to Coalgebra. Towards Mathematics of States and Observations*. Unpublished, 2011.

[20] B. Jacobs and J. Mandemaker. Relating operator spaces via adjunctions. *arXiv:1201.1272v1*, 2012.

[21] B. Jacobs, A. Silva, and A. Sokolova. Trace semantics via determinization. In *Lecture Notes in Computer Science, to appear*, Proceedings of CMCS 2012, 2012.

[22] J. Jauch. *Foundations of Quantum Mechanics*. Addison-Wesley Pub. Co., 1968.

[23] E. Kreyszig. *Introductory Functional Analysis with Applications*. John Wiley, 1978.

[24] S. Mac Lane. *Categories for the Working Mathematician*. Springer, 1998.

[25] C. Moore and J. Crutchfield. Quantum automata and quantum grammars. *Theoretical Computer Science*, 237(1):275–306, 2000.

[26] A. Nerode. Linear automaton transformations. *Proc. Amer. Math. Soc.*, 9(4):541–544, 1958.

[27] J. von Neumann. *Mathematical Foundations of Quantum Mechanics.* Princeton University Press, 1955.

[28] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information.* Cambridge university press, 2000.

[29] J. Rutten. Automata and coinduction (an exercise in coalgebra). In D. Sangiorigi and R. de Simone, editors, *Proceedings of CONCUR '98*, volume 1466 of *LNCS*, pages 194–218. Springer, 1998.

[30] J. Rutten. Universal coalgebra: a theory of systems. *Theoretical Computer Science*, 249(1):3–80, 2000.

[31] A. Silva, F. Bonchi, M. Bonsangue, and J. Rutten. Generalizing the powerset construction, coalgebraically. In *Proceedings of the IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science*, volume 8 of *Leibniz International Proceedings in Informatics*, pages 272–283, 2010.

[32] S. Elías Venegas-Andraca. *Quantum Walks for Computer Scientists.* Morgan & Claypool, 2008.

[33] J. Weidmann. *Lineare Operatoren in Hilberträumen.* B.G. Teubner, 2000.