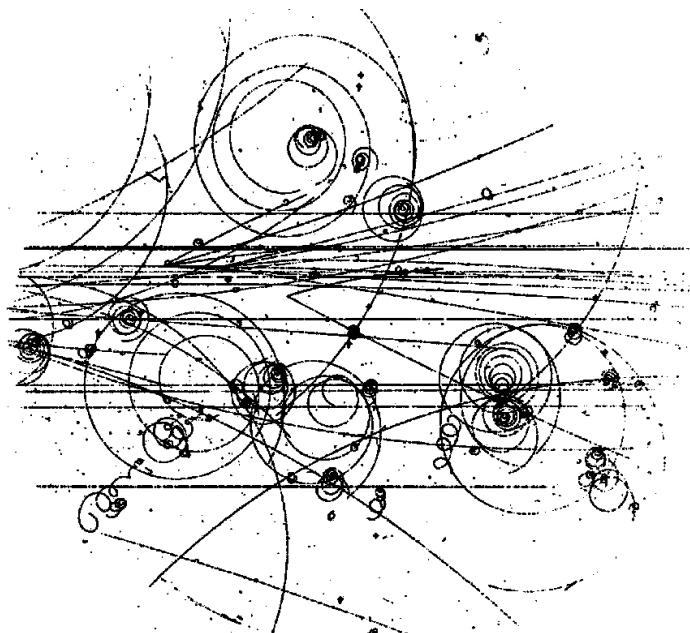


DAVIDE CAMINO

QUANTUM COMPUTING FOR
LOGICAL INFERENCE

Subtitle



CONTENTS

I THEORETICAL BASES	
1 QUANTUM MECHANICS	3
1.1 Experiments	3
1.1.1 Spin	3
1.1.2 Qubit	5
1.1.3 Boolean Logic	5
1.2 Quantum states	6
1.2.1 Vector Spaces	7
1.2.2 Bra and Ket	7
1.2.3 Hidden variables	8
1.2.4 Spin states	9
1.3 Observables	11
1.3.1 Hermitian operator	11
1.3.2 Principles of quantum mechanics	12
1.3.3 Spin Operator	13
1.3.4 Theory and experiments	14
1.3.5 Operator and Measure	16
1.4 Temporal Evolution	16
1.4.1 Unitarity	16
1.4.2 Time-Development Operator	17
1.4.3 The Hamiltonian	18
1.4.4 Commutators	19
1.4.5 Conservation of Energy	20
1.5 Conclusions	20
2 QUANTUM GATE	23
3 QUANTUM ANNEALING	25
4 ONTOLOGY	27
4.1 Knowledge Base	27
4.1.1 Ontology in philosophy	27
4.1.2 Ontology in computer science	27
4.1.3 OWL Language	28
4.1.4 Importance of ontologies	29
4.2 Example ontologies	29
4.2.1 Simple ontology	29
4.2.2 DOLCE ontology	31
4.3 Rerasoning on ontology	31
4.3.1 <i>SROIQ</i> DL	31
4.3.2 Interpretation of a knowledge base	32
4.3.3 Complexity of reasoning	33
4.4 Conclusions	34
II TOOLS	
5 ENVIRONMENT SETUP	37

5.1	Python environment	37
5.2	IBM Qiskit	37
5.2.1	Hello World	38
5.2.2	Transpilation	38
5.2.3	Execution	39
5.2.4	A complete example on real hardware	40
5.3	D-Wave Ocean	40
5.3.1	Hello World	40
5.3.2	Example on real hardware	41
5.3.3	Minor embedding	41
5.4	PyQUBO and qubovert	41
5.4.1	PyQUBO	42
5.4.2	qubovert	43
5.5	Conclusion	44
6	QA-PROLOG	45
6.1	The project	45
6.2	Pipeline	45
6.3	Update to the project	45
6.4	Related Work	45
6.5	Conclusion	45
III EXPERIMENTS		
7	A QUANTUM ONTOLOGY	49
7.1	Ontology structure	49
7.2	Prolog version	49
7.3	Inference on the ontology	49
7.4	Conclusion	49
8	QAOA	51
8.1	QAOA	51
8.2	From QUBO to Pauli operator	51
8.3	Experiments	51
8.4	Conclusion	51
9	CONCLUSION	53

Part I
THEORETICAL BASES

1

QUANTUM MECHANICS

In this chapter we will explore the basics of quantum mechanics in order to understand what we can or cannot do with a quantum computer. The reference architecture for this work is the quantum gate-based quantum computer. In the next pages we will try to justify why the algorithms that run on this hardware need to be reversible.

The chapter starts from the very beginning with the definition of a quantum system and presents the basis to understand the evolution of a quantum system, trying to justify why every evolution needs to be reversible and what exactly reversibility means. At the end of the chapter we will put all of our new knowledge together to derive the famous Schrödinger equation.

With all this work we will be able to imagine the function of a quantum gate and understand the limitations that are imposed when we develop an algorithm for a quantum computer.

1.1 EXPERIMENTS

We start our introduction to quantum mechanics with an experiment. Experiments are not only an excuse to introduce the topic, but the essential key of physics, both classical and modern.

Theory and models need to adapt to the experiments, and when the experimental results are in contradiction with the actual model it means that the model needs to be changed to respect the behavior of the world.

1.1.1 Spin

We analyze the experiment of an electron in a magnetic field. An electron is an electrically charged particle; when some electrons are shot in an electric field all of them are influenced by Coulomb's law; if all electrons have the same initial velocity the beam of electrons remains intact.

What happens to the same beam in a magnetic field? Again electrons are deflected by a force, but this time the beam splits. If the initial velocity was parallel to the x axis and the magnetic field is oriented along the z axis some electrons are deflected upward, some downward, but the intensity of the deflection is the same for all the electrons. This means that no electrons are deflected less or more than the others and the beam splits exactly into two parts.



Figure 1.1: Experiment's schema

38 Starting from this experiment we can make a measuring instrument: this
 39 apparatus \mathcal{A} can be oriented along an arbitrary axis, and in the previous con-
 40 figuration \mathcal{A} displays $+1$ if the electron is deflected upward, -1 otherwise.
 41 We call this number the spin of the electron.

Repeatability of measure
 42 If we measure the spin of an electron and \mathcal{A} displays $+1$, we can confirm
 43 the experiment's results by measuring the spin again and we obtain spin $+1$
 44 every time. This means that the measurements are repeatable (an essential
 45 property to construct models and make predictions). We can think, and it
 46 will be clear later why it is useful, that the first experiment prepares the spin
 47 $+1$ and the others confirm this result.

48 Spin is a quantum property and all the visual representations such as the
 49 rotation of the electron around its axis would lead to misrepresentation. Spin
 50 and rotation, however, have some similarities. Let's analyze what would
 51 happen if we consider a charged sphere in a magnetic field with the laws of
 52 classical physics. We consider a sphere rotating around its axis, and this axis
 53 is parallel to the z axis. The x or y component of the angular momentum is
 54 zero. Measuring the component along a generic axis, oriented like the versor
 55 \hat{n} ¹, we would obtain a result proportional to the projection of \hat{z} on \hat{n} . This
 56 projection can be found with the scalar product $\hat{z} \cdot \hat{n} = \cos \theta^2$, where θ is the
 57 angle between the axes.

58 Now we consider the quantum version of this phenomenon. Let's start by
 59 measuring the z component of the spin and assume that the result is $\sigma_z = +1$;
 60 if we rotate the apparatus \mathcal{A} around, for example, the x axis, we can measure
 61 σ_x . This component would not be zero, and \mathcal{A} keeps displaying only $+1$
 62 or -1 . The single result is not helpful, but we can repeat the experiment,
 63 namely:

- 64 1. orienting \mathcal{A} along the z axis and preparing a spin $\sigma_z = +1$
- 65 2. rotating \mathcal{A} around x
- 66 3. measuring the x component of the spin

67 statistically we would observe the same number of $\sigma_x = +1$ and $\sigma_x = -1$.

68 If we start the experiments with a spin prepared as $\sigma_z = +1$ and then
 69 orient \mathcal{A} along a generic axis \hat{n} each measure would be binary and unpre-
 70 dictable, but the mean of the measures tends to $\hat{z} \cdot \hat{n} = \cos \theta$ where θ is the
 71 angle between \hat{z} and \hat{n} . In the most general case we can start with the ap-
 72 paratus oriented like m and prepare the spin $\sigma_m = +1$, then we rotate \mathcal{A}
 73 around \hat{n} without interfering with the spin and measure again; we would
 74 obtain the statistical result $\langle \sigma_n \rangle = \hat{m} \cdot \hat{n}$ ³.

75 The result of a single measure is non deterministic, but we can make pre-
 76 dictions over the mean values of the measures: the expected values behave
 77 as the single results of the classic experiment.

Invasive experiments
 78 Considering now a sequence of three measures: starting with \mathcal{A} oriented
 79 along z we prepare the spin $\sigma_z = +1$, then we rotate \mathcal{A} to measure σ_x
 80 obtaining, let's say, $+1$ (the reasoning is the same if we obtain -1); lastly
 81 returning with \mathcal{A} parallel to z we cannot make any prediction on the single

¹ A versor is a vector of magnitude 1 (unit vector), it is normally used to specify a direction.

² We can use directly the angle because we are considering versors.

³ The Dirac bracket $\langle \rangle$ denotes the statistical mean of a quantity. We call that expectation value.

82 result, the initial configuration (with $\sigma_z = +1$) is lost forever, the only result
 83 we can predict is that $\langle \sigma_z \rangle = 0$.

84 **1.1.2 Qubit**

85 We have introduced the spin referring to electrons in a magnetic field. How-
 86 ever, we can study the spin without examining the associated electron; we
 87 have isolated a simple physical system, the simplest we can study.

88 Spin belongs to a class of simple physical systems called *qubit*; in all of
 89 these systems the result of a measure is binary. We will see that, even if the
 90 result of a measure is equal to the classical *bit*, the qubit system is described
 91 in a very different way compared to its classical alter ego.

92 **1.1.3 Boolean Logic**

93 In this paragraph we try to understand why we need two different ways
 94 to describe a classical and a quantum state space. To do so we analyze the
 95 results of some logical propositions, both basic and composed via logical
 96 connectives.

97 Starting with the classical case we consider a bag of colored and numbered
 98 balls. We can construct the state space by enumerating all states, namely
 99 taking each ball from the bag and annotating the pair number–color. The
 100 basic propositions we analyze are:

- 101 • The extracted ball is red.
- 102 • The number on the extracted ball is even.

103 If we consider a particular state we can say if a proposition is true or
 104 false; we can also define two subsets of balls, the first with all the red balls
 105 (for this subset the first proposition is true), the second one with the balls
 106 that show an even number (subset that makes the second proposition true).
 107 Considering now disjunction and conjunction:

- 108 • The extracted ball is red *or* even.
- 109 • The extracted ball is red *and* even.

110 Again it is simple to associate a truth value to these propositions if we con-
 111 sider a single state; also we can construct two subsets that satisfy the propo-
 112 sitions from the subsets we defined before: the new subsets are respectively
 113 the union and intersection of the old ones.

114 In the quantum world the situation is very different. Let's start from
 115 propositions that can be verified with a simple experiment:

- 116 • The z component of the spin is $+1$.
- 117 • The x component of the spin is $+1$.

118 If we want to check the first proposition we can orient the apparatus \mathcal{A} along
 119 z and make a measurement; the same procedure can be followed for the
 120 second proposition. The disjunction and conjunction of these propositions
 121 are:

- 122 • The z component of the spin is $+1$ or the x component is $+1$.
 123 • The z component of the spin is $+1$ and the x component is $+1$.

124 Starting with the disjunction. Considering a state prepared, without our
 125 knowledge, with $\sigma_z = +1$. If our first measure is along the z axis, \mathcal{A} will
 126 always display $+1$ and we can immediately conclude that the proposition
 127 is true. If we start measuring the x component, we have a 50% chance that
 128 \mathcal{A} displays $+1$ or -1 ; also this measurement destroys the initial state and
 129 the measure of σ_z becomes non predictable. In this scenario we have a 25%
 130 chance of deducing that the proposition is false; figure 1.2 shows all the
 The disjunction is not
 commutative^{g2} possible measurement results in this case. The logical value of a proposition
 depends on the order in which we perform the measurements.

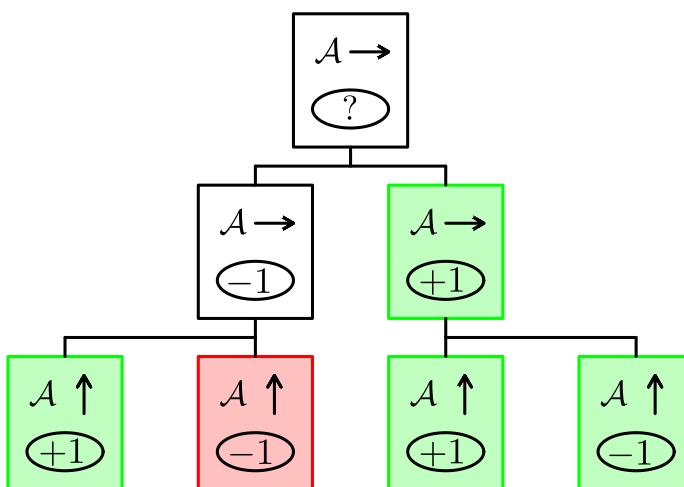


Figure 1.2: The apparatus \mathcal{A} is represented as a box, the arrow represents the direction along which the apparatus is oriented, the display (ellipse) shows the result of measurement. We have highlighted in green the cases in which we can immediately conclude that the disjunction of the propositions is true

133 The conjunction is even worse: no matter the order of the measurements,
 134 the second one destroys the result of the first. The disjunction is true if at
 135 least one of the sub-propositions is true, and if we find a spin component
 136 that is $+1$ we can always confirm this result with another measurement. In
 137 the conjunction the two sub-propositions must be true *at the same time*, but
 The conjunction loses its
 meaning^{g3} with the second measurement we lose all the knowledge of the first one. We
 can never conclude that the conjunction is true.

140 1.2 QUANTUM STATES

141 In the previous section we have understood that a state space of a quantum
 142 system cannot be represented in the same way as a classical state space. Now
 143 we present a formal mathematical model to describe the state space for spin.

144 **Axiom 1.** *The state space for a quantum system is a complex vector space.*

145 This is a physical axiom, which means that it is true because there are a lot
 146 of experiments that confirm this model and none that shows a contradiction.

147 **1.2.1 Vector Spaces**

148 A vector space is a mathematical and abstract construction that can have
 149 multiple dimensions (even infinite) and has, as components, integers, real
 150 or complex numbers, or other elements. An example that shows well how
 151 abstract a vector space can be is the complex-valued continuous function of
 152 variable x ; the set of these functions generates a vector space.

153 In quantum mechanics the state space is described by a vector space hav- *Hilbert space*
 154 ing as element $|A\rangle$ called *ket*. The properties of this space are:

- 155 • the sum of two kets is a ket;
- 156 • addition is commutative;
- 157 • addition is associative;
- 158 • existence of identity element for addition;
- 159 • existence of inverse elements for addition;
- 160 • existence of identity element for scalar multiplication;
- 161 • linearity property.

162 **1.2.2 Bra and Ket**

163 An example of ket that we will find often is the column vector of two dimen-
 164 sions:

$$|A\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$$

165 where α_1 and α_2 are complex numbers. With this simple example of ket it
 166 is easy to verify the validity of all previously described properties.

167 If, for complex numbers, exists the complex conjugate, for every ket there
 168 exists a *bra*. The set of bra generates a dual conjugate space with respect
 169 to the state space of ket. We denote a bra as $\langle A|$. If $|A\rangle$ is the ket of the
 170 previous example the corresponding bra is a row vector having as elements
 171 the complex conjugate of $|A\rangle$:

$$\langle A| = (\alpha_1^*, \alpha_2^*).$$

172 Name and symbol associated with elements of Hilbert spaces become clear
 173 when we define the product *bra-ket*, this is the corresponding scalar product
 174 of an ordinary vector and is called inner product. Considering bra and ket of
 175 two dimensions we can evaluate the inner product by adding the products
 176 of corresponding components:

$$\langle A | B \rangle = (\alpha_1^*, \alpha_2^*) \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \alpha_1^* \beta_1 + \alpha_2^* \beta_2.$$

177 Having the inner product we can define:

178 **VERSOR** normalized vector $|A\rangle$ in which $\langle A | A \rangle = 1$;

179 **ORTHOGONAL VECTOR** vectors that have a null inner product: $\langle A | B \rangle = 0$.

Inner product

We are familiar with these concepts in two and three dimensions, the first one is a vector of length one, the second is the right angle between two vectors. This representation is misleading in our case, we cannot imagine a ket like an arrow and the state space is completely abstract even if there are properties and operations in common between this space and the 3D space that we are familiar with.

We have lost the geometric interpretation, and it seems that we have defined two completely abstract and useless concepts, we will see next that these are key concepts in the description of quantum systems and have a precise and important physical meaning.

Orthonormal basis

By having a vector space is possible to build a set of orthogonal versors that generates all vectors in the given space. This set is called orthonormal basis and the cardinality of the set is equal to the dimension of the space.

Formally having a basis $\mathcal{B} = \{|i_1\rangle, |i_2\rangle, \dots, |i_N\rangle\}$ of a space with N dimensions, we can write a generic vector in that space as

$$|A\rangle = \sum_{n=1}^N \alpha_n |i_n\rangle = \sum_{n=1}^N |i_n\rangle \langle i_n | A \rangle \quad (1.1)$$

this is the linear combination of the basis versors; where kets $|i_n\rangle$ are the versors in the basis and α_n are the vector components. We can obtain those components with the inner product between the vector $|A\rangle$ and the basis versors:

$$\alpha_i = \langle i | A \rangle. \quad (1.2)$$

1.2.3 Hidden variables

In a classical system we can measure all the variables associated to a physical system and then make a deterministic prediction of the evolution of that system. From the experiments described in the first section we have learned that a quantum system is not completely predictable even if we can make all the measurements that we want⁴. We can ask ourselves if our measurements aren't enough, if there are other variables that can make the prediction completely deterministic. About that topic we don't have any experimental proof, the opinion of physicists is divided in two main visions:

OPINION ONE : there are hidden variables and, if we manage to measure them, the prediction of results become deterministic. These variables can be

- very difficult to measure
- unknowable to us because also we are constituted by quantum material.

OPINION TWO : hidden variables don't exist, we already know all the information about a given system and quantum mechanics is intrinsically non deterministic.

No hidden variables

Probably no experiment could determine which vision is correct, but this doubt doesn't worsen our comprehension of the physical world. We can

⁴ We remember that a measure along one axis destroys our knowledge about the result along another axis.

simply choose one vision and build our model coherently. We choose the simpler one, without hidden variables, all that we have to model are the quantities that we can measure and the measurements allow us to know all the information about a given system.

Even if we have lost complete determinism, knowing the state of a system gives us some information about the system and the successive measurements. In the next section we will see what we can deduce about spin.

1.2.4 Spin states

Let's start enumerating all possible spin states along the coordinate axes. If we rotate the apparatus \mathcal{A} around z , we can obtain $\sigma_z = \pm 1$; we call these states *up* and *down* and label them with kets $|u\rangle$ and $|d\rangle$. Orienting \mathcal{A} along x , we obtain *left* $|l\rangle$ and *right* $|r\rangle$. Lastly, along the y axis, we measure the states *in* $|i\rangle$ and *out* $|o\rangle$.

The hypothesis that there aren't hidden variables allows us to represent the space state in a simple way: each spin state can be represented as a ket in a two-dimensional complex vector space.

To express a vector we need a basis; we choose $\mathcal{B} = \{|u\rangle, |d\rangle\}$ ⁵ and try to obtain all states as a linear combination (*superposition*) of the basis vectors. A generic state $|\mathcal{A}\rangle$ can be expressed as:

$$|\mathcal{A}\rangle = \alpha_u |u\rangle + \alpha_d |d\rangle$$

where α_u and α_d are the components of $|\mathcal{A}\rangle$ along $|u\rangle$ and $|d\rangle$, and can be obtained by projection: $\alpha_u = \langle u | \mathcal{A} \rangle$ and $\alpha_d = \langle d | \mathcal{A} \rangle$ (as in equation 1.2).

$|\mathcal{A}\rangle$ components are complex numbers and their physical meaning is: having a spin prepared in the state $|\mathcal{A}\rangle = \alpha_u |u\rangle + \alpha_d |d\rangle$ ⁶; $\alpha_u^* \alpha_u$ is the probability of measuring $\sigma_z = +1$, while $\alpha_d^* \alpha_d$ is the probability that a measurement of σ_z will yield -1 . Formally we can denote the probability of measuring $+1$ and -1 as P_u and P_d respectively and write:

$$\begin{aligned} P_u &= \langle \mathcal{A} | u \rangle \langle u | \mathcal{A} \rangle \\ P_d &= \langle \mathcal{A} | d \rangle \langle d | \mathcal{A} \rangle. \end{aligned} \quad (1.3)$$

Spin space states have two dimensions

Probability amplitudes

Components α_u and α_d are called probability amplitudes, and their physical meaning is given by the square of the magnitude. This is the actual probability, and we want the sum of all probabilities to be one. This is equivalent to requiring that $|\mathcal{A}\rangle$ is normalized: $\langle \mathcal{A} | \mathcal{A} \rangle = 1$.

Now we will show why $|u\rangle$ and $|d\rangle$ have to be orthogonal:

$$\begin{aligned} \langle u | d \rangle &= 0 \\ \langle d | u \rangle &= 0. \end{aligned}$$

We try to give an idea with a *reductio ad absurdum*: if $|u\rangle$ and $|d\rangle$ were not orthogonal, the projection of one on the other would not be null. This means

⁵ We will show that these vectors are in fact orthogonal and why they need to be.

⁶ From now on we use "prepared" or "measured" as synonyms: every measurement is invasive and can change the spin state, so no matter what was the previous state, after a measurement the state is the one we have measured.

that if we orient \mathcal{A} along z and measure $\sigma_z = +1 = |\mathbf{u}\rangle$, we would have $\alpha_d = \langle \mathbf{d} | \mathbf{u} \rangle \neq 0$, which is a contradiction to experimental results. If $\alpha_d \neq 0$, then $\alpha_d^* \alpha_d > 0$; we started with a state prepared as $\sigma_z = +1$ and ended with a nonzero probability of measuring $\sigma_z = -1$: this is absurd.

Orthogonal states are mutually exclusive

We can extend the reasoning to a general and key concept of quantum mechanics: two orthogonal states are distinct and mutually exclusive. If the system is in the first state, the probability of finding it in the second is zero.

Now we are ready to express spin states as linear combinations of the basis vectors $\mathcal{B} = \{|\mathbf{u}\rangle, |\mathbf{d}\rangle\}$. The representation of the basis vectors themselves is naturally easy:

$$|\mathbf{u}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (1.4)$$

$$|\mathbf{d}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (1.5)$$

To construct vector *right*, let's consider a spin prepared in the state $|\mathbf{r}\rangle$. If we measure σ_z , we have a 50% chance of obtaining +1 (and 50% for -1); this means that for $|\mathbf{r}\rangle$ we have $\alpha_u^* \alpha_u = \alpha_d^* \alpha_d = 1/2$. A vector that satisfies this constraint is:

$$|\mathbf{r}\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}. \quad (1.6)$$

The reasoning is the same for state *left*; we also add the constraint that a state *left* cannot be *right* and vice versa: $\langle \mathbf{r} | \mathbf{l} \rangle = \langle \mathbf{l} | \mathbf{r} \rangle = 0$. We can express *left* as:

$$|\mathbf{l}\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}. \quad (1.7)$$

Lastly, the constraints to find explicit forms for *in* and *out* are:

- states must be orthogonal: $\langle \mathbf{i} | \mathbf{o} \rangle = \langle \mathbf{o} | \mathbf{i} \rangle = 0$;
- if we have a spin prepared as *in* or *out*:
 - equiprobability of measuring $\sigma_z = +1$ and $\sigma_z = -1$;
 - equiprobability of measuring $\sigma_x = +1$ and $\sigma_x = -1$.

Two vectors that satisfy these constraints are:

$$|\mathbf{i}\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix} \quad (1.8)$$

$$|\mathbf{o}\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{pmatrix}. \quad (1.9)$$

This last derivation shows why it is important that the state space is complex: if we only accepted real components for our vectors, the system of equations we have implicitly defined would not have any solution⁷.

⁷ To avoid confusion, we point out that $|\mathbf{i}\rangle$ is the ket of state *in*. *i*, instead, is the imaginary unit.

278 1.3 OBSERVABLES

279 We have learned that in classical mechanics we can trust our intuition, and
 280 we can do one or more measurements to know exactly the state of a system:
 281 a measurement does not perturb the state, which is the same before, during,
 282 and after the measurement.

283 In quantum mechanics the situation is more complex; our intuition is mis-
 284 leading, and we need mathematical tools to describe what we can measure:
 285 the observables. These tools are mathematical operators called *machines* (**M**)
 286 and have as both input and output state vectors.

287 **Axiom 2.** *Machines associated with observables are described by linear operators.*

288 We will show that machines are Hermitian operators, so let's start defining
 289 these operators and describing their properties⁸.

290 1.3.1 Hermitian operator

291 Formally, machines modify a state vector in this way:

$$\mathbf{M}|\mathbf{A}\rangle = |\mathbf{B}\rangle$$

292 The linearity of machines implies that:

$$\mathbf{M}|\mathbf{A}\rangle = |\mathbf{B}\rangle \Rightarrow \mathbf{M}z|\mathbf{A}\rangle = z|\mathbf{B}\rangle$$

293 and:

$$\mathbf{M}(|\mathbf{A}\rangle + |\mathbf{B}\rangle) = \mathbf{M}|\mathbf{A}\rangle + \mathbf{M}|\mathbf{B}\rangle.$$

294 If we choose a basis to represent machines and state vectors, we can write
 295 explicitly the linear operator as an $N \times N$ matrix, where N is the dimension
 296 of the vector space of the state vectors. A generic machine that transforms
 297 spins can be expressed as:

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}.$$

298 When we fix a basis, we are forced to express all state vectors and opera-
 299 tors in that basis, but now we have a set of rules to define the application of
 300 the operator to a state vector, i.e. the matrix multiplication:

$$\mathbf{M}|\mathbf{A}\rangle = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \times \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = |\mathbf{B}\rangle$$

301 When we consider a linear operator, we can search for eigenvalues and
 302 eigenvectors (if they exist). Eigenvectors are vectors that don't change their
 303 direction when multiplied by the operator; their magnitude is scaled by a
 304 constant factor called the eigenvalue. Formally:

Eigenvalues and eigenvectors

$$\mathbf{M}|\lambda\rangle = \lambda|\lambda\rangle$$

305 where $|\lambda\rangle$ is the eigenvector and λ the eigenvalue.

8 The reason why we need this kind of operator will be clear in 1.3.2.

306 Considering the transformation between ket $|A\rangle$ and $|B\rangle$: $\mathbf{M}|A\rangle = |B\rangle$,
 307 taking into account the dual space of bras and searching for a machine that
 308 transforms the bra $\langle A|$ into $\langle B|$, we cannot simply use the matrix having as
 309 elements the complex conjugate of \mathbf{M} ; the correct operator is the *Hermitian*
 310 *conjugate* of \mathbf{M} , which is the transpose of the matrix having as elements
 311 the complex conjugates of \mathbf{M} . We denote the Hermitian conjugate with the
 312 dagger \dagger :

$$\mathbf{M}^\dagger = [\mathbf{M}^*]^T = [\mathbf{M}^T]^*.$$

313 We can now write:

$$\mathbf{M}|A\rangle = |B\rangle \Rightarrow \langle A|\mathbf{M}^\dagger = \langle B|.$$

314 An operator that is equal to its Hermitian conjugate is called a *Hermitian*
 315 *operator*. Formally, \mathbf{M} is Hermitian if and only if

$$\mathbf{M} = \mathbf{M}^\dagger.$$

316 Hermitian operators have some important properties:

- 317 • all eigenvalues are real;
- 318 • eigenvectors form a *complete set*: all vectors obtained with the applica-
 319 tion of the operator can be expressed as a linear combination of eigen-
 320 vectors;
- 321 • if λ_1 and λ_2 are different eigenvalues, the associated eigenvectors are
 322 orthogonal;
- 323 • if two eigenvalues are equal (*degeneracy*), it is always possible to find
 324 two associated eigenvectors that are orthogonal.

Fundamental theorems 325 The last three properties can be summed up in the following way:

326 **Theorem 1.** *The eigenvectors of a Hermitian operator form an orthonormal basis.*

327 1.3.2 Principles of quantum mechanics

328 Let's introduce the first four principles of quantum mechanics, the ones
 329 about observables⁹.

330 **Principles 1.** *Observables in quantum mechanics are described by linear operators*
 331 \mathbf{L} .

332 \mathbf{L} must also be a Hermitian operator: we can consider this proposition an
 333 axiom itself or deduce it from the other principles.

334 **Principles 2.** *The results of a measurement can only be the eigenvalues associated*
 335 *with the observable operator.*

336 Calling λ_i a generic eigenvalue and $|\lambda_i\rangle$ the associated eigenvector, if the
 337 system is in the *eigenstate* $|\lambda_i\rangle$, the measurement always returns λ_i . Since
 338 all λ_i must be physical quantities they must be real, a peculiar property of
 339 Hermitian operators.

⁹ The fifth, and last one, concerns the temporal evolution. It will be discussed later on (1.4).

340 **Principles 3.** *Unambiguously distinguishable states are represented by orthogonal
341 vectors.*

342 Distinguishable states can be separated without ambiguity by a measure-
343 ment. For example, if we want to distinguish between $|u\rangle$ and $|d\rangle$, we mea-
344 sure σ_z : *up* and *down* are distinct. We cannot, instead, say if a certain system
345 is in state *up* or *right*, because even if the system is in the state $|u\rangle$ we can
346 still measure σ_x and find (with 50% chance) that the system is in state $|r\rangle$.

347 The inner product is a measure of how much two states are indistin- *Overlap*
348 guishable; for that reason it is also called overlap. Two states are physically dis-
349 tinct if the overlap is zero.

$$\begin{aligned}\langle u | d \rangle &= 0 \\ \langle u | r \rangle &\neq 0\end{aligned}$$

350 **Principles 4.** *If the system is in state $|A\rangle$ and we measure the observable L , the
351 probability of obtaining λ_i is:*

$$P(\lambda_i) = \langle A | \lambda_i \rangle \langle \lambda_i | A \rangle .$$

352 where λ_i is a generic eigenvalue of L and $\langle \lambda_i |$, $|\lambda_i \rangle$ are the bra and ket asso-
353 ciated with that eigenvalue (eigenvector of λ_i).

354 1.3.3 Spin Operator

355 The principles tell us what properties a machine must have to represent an
356 observable. Let's construct the spin operator σ .

357 Until now, we have measured spins with the apparatus \mathcal{A} , orienting \mathcal{A}
358 along the component of our interest. σ is a mathematical tool that allows
359 us to make predictions about the result of a measurement with \mathcal{A} (fourth
360 principle); as we can rotate \mathcal{A} , we must also rotate σ (mathematically). For
361 this spatial property, σ is called a *3-vector operator*.

362 **OPERATOR σ_z :** Let's start with the simplest operator¹⁰. The second prin-
363 ciple says that all eigenvectors of σ_z are $|u\rangle$ and $|d\rangle$, with associated eigen-
364 values +1 and -1. We can write this assertion as equations:

$$\begin{aligned}\sigma_z |u\rangle &= |u\rangle \\ \sigma_z |d\rangle &= -|d\rangle .\end{aligned}$$

365 In matrix form:

$$\begin{aligned}\begin{pmatrix} (\sigma_z)_{11} & (\sigma_z)_{12} \\ (\sigma_z)_{21} & (\sigma_z)_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} (\sigma_z)_{11} & (\sigma_z)_{12} \\ (\sigma_z)_{21} & (\sigma_z)_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= -\begin{pmatrix} 0 \\ 1 \end{pmatrix}.\end{aligned}$$

366 The solution of this system is¹¹:

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

¹⁰ This is because we have chosen $\mathcal{B} = \{|u\rangle, |d\rangle\}$ as the basis.

¹¹ It is easy to verify that this operator is also linear.

367 **OPERATOR σ_x :** With the same reasoning, we can construct the operator 368 along the x axis. We have already deduced the representations of *right* and 369 *left* in equations [1.6](#) and [1.7 on page 10](#). The equations that allow us to 370 construct σ_x are:

$$\begin{pmatrix} (\sigma_x)_{11} & (\sigma_x)_{12} \\ (\sigma_x)_{21} & (\sigma_x)_{22} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\begin{pmatrix} (\sigma_x)_{11} & (\sigma_x)_{12} \\ (\sigma_x)_{21} & (\sigma_x)_{22} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix} = - \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix}.$$

371 The solution of this system is:

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

372 **OPERATOR σ_y :** The last direction is along the y axis. Considering the 373 expressions for *in* and *out* given in equations [1.8](#) and [1.9 on page 10](#), and 374 following the second principle, we can write:

$$\begin{aligned} \sigma_y |i\rangle &= |i\rangle \\ \sigma_y |o\rangle &= -|o\rangle. \end{aligned}$$

375 We can rewrite this in matrix form, and the solution we would obtain is:

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

*Pauli matrices*366 We have obtained a matrix representation of the three spin operators σ_z , 377 σ_x , and σ_y :

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (1.10)$$

378 These famous and important matrices are named after their inventor, Wolfgang 379 Ernst Pauli.

380 1.3.4 Theory and experiments

381 Thanks to the operators σ_z , σ_x , and σ_y , if we know the state vector, we can 382 statistically predict the result of a measurement of the spin along one of 383 the three coordinate axes. What can we say about a measurement taken by 384 orienting the apparatus \mathcal{A} along a generic direction?

385 Considering \mathcal{A} oriented along the unit vector \hat{n} , if σ behaves as a 3-vector, 386 in order to obtain σ_n we only need the inner product:

$$\sigma_n = \vec{\sigma} \cdot \hat{n}$$

387 Expanding the components:

$$\sigma_n = \sigma_x n_x + \sigma_y n_y + \sigma_z n_z.$$

388 If we choose the basis $\mathcal{B} = \{|u\rangle, |d\rangle\}$, we can use the Pauli matrices to
 389 express in matrix form the expression for σ_n :

$$\sigma_n = n_x \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + n_y \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + n_z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix}.$$

390 Given a direction (expressed by the unit vector \hat{n}), we can construct the
 391 matrix we have now made explicit, and then, after finding eigenvalues and
 392 eigenvectors, we can know all possible results of a measurement and ob-
 393 tain the probability associated with each result. For example, considering a
 394 direction in the x - z plane, the operator σ_n would be:

$$\sigma_n = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$$

395 where θ is the angle between \hat{n} and z . For this matrix, the eigenvalues and
 396 eigenvectors are:

$$\lambda_1 = 1 \quad |\lambda_1\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix}$$

397 and

$$\lambda_2 = -1 \quad |\lambda_2\rangle = \begin{pmatrix} -\sin \frac{\theta}{2} \\ \cos \frac{\theta}{2} \end{pmatrix}.$$

398 It should be pointed out that the theory is in agreement with experimen-
 399 tal results¹². Eigenvalues are $+1$ and -1 , exactly the only results that the
 400 apparatus \mathcal{A} can retrieve. The probability of obtaining a certain result can
 401 be evaluated as:

$$\begin{aligned} P(+1) &= |\langle u | \lambda_1 \rangle|^2 = \cos^2 \frac{\theta}{2} \\ P(-1) &= |\langle u | \lambda_2 \rangle|^2 = \sin^2 \frac{\theta}{2} \end{aligned}$$

402 Lastly, let's calculate the average value for the measurement σ_n . From the
 403 first experiment we have seen in 1.1.1, we already know that the result of
 404 repeated measurements with \mathcal{A} is $\cos \theta$. Let's verify if our model is coherent
 405 with the world.

Expectation value

406 Expected values are obtained as:

$$\langle L \rangle = \sum_i \lambda_i P(\lambda_i)$$

407 Specifically:

$$\langle \sigma_n \rangle = (+1) \cos^2 \frac{\theta}{2} + (-1) \sin^2 \frac{\theta}{2} = \cos \theta.$$

408 This is in complete agreement with the experimental results.

409 Before going on, we present, without proof, a useful theorem about expec-
 410 tation values:

411 **Theorem 2.** *To know the expectation value of an observable, we can simply place the
 412 operator associated with the observable between the bra and ket of the state vector:*

$$\langle L \rangle = \langle A | L | A \rangle \tag{1.11}$$

413 where L is an observable, $|A\rangle$ is a state vector, and $\langle A|$ is the corresponding
 414 bra.

¹² If not, we must abandon this model and build another one.

415 **1.3.5 Operator and Measure**

416 Operators allow us to know the probability of measuring a certain spin given
 417 the direction of the measurement and the state vector. This probability is
 418 expressed by the state vector that we obtain when we apply the operator σ
 419 to the initial state.

420 It is important not to confuse the measurement act with the application
 421 of a machine that represents the observables. The spin state after the mea-
 422 surement is not the same as the one we obtain after the application of the
 423 operator. The operator is only an abstract mathematical construct that allows
 424 us to make statistical predictions about results, but doesn't have physical im-
 425 plications.

426 Let's consider an example to clarify the previous assertion. Having a spin
 427 prepared in the *up* state, its state vector is $|u\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. If we apply the operator
 428 σ_z , we would obtain again $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and if we measure the spin with \mathcal{A} oriented
 429 along z , it will always display +1, and we conclude that the state after the
 430 measurement is $|u\rangle$.

431 Consider now a spin prepared *right*, i.e. $|r\rangle = 1/\sqrt{2}|u\rangle + 1/\sqrt{2}|d\rangle$. Applying
 432 again the operator σ_z , the new state vector is $1/\sqrt{2}|u\rangle - 1/\sqrt{2}|d\rangle$. This
 433 vector tells us the probability of measuring $\sigma_z = +1$ (50%), but it is not the
 434 spin state after the measurement. Using the apparatus \mathcal{A} , we could measure:

- 435 • +1: the final state will be *up*;
- 436 • -1: the final state will be *down*.

437 No matter the result of the measurement, the final state will be different
 438 from the one we obtained by applying the operator.

439 **1.4 TEMPORAL EVOLUTION**

440 Let's explore the laws that describe the temporal evolution of a quantum
 441 system. In particular, we will see how the state vector can evolve over time.

442 **1.4.1 Unitarity**

443 In classical mechanics we are used to having a motion law that links differ-
 444 ent states of our system deterministically; this means being able to know
 445 precisely the following state given the previous one. A good law, however,
 446 doesn't allow us only to know the future, but also the past states that brought
 447 the system to the current state¹³.

Reversibility¹³ 448 In other words, we want physical transformations to be reversible. This
 449 requirement is so important that we call this property the *minus first law*,
 450 because it underlies everything else. If we think about the system states as
 451 nodes in an oriented graph, reversibility imposes that each node has exactly
 452 one input edge and one output edge. This fundamental law is also true

¹³ For example, if we observe a ball in free fall touching the floor with a certain speed and at a certain time, we can know exactly when and from what height the ball started its fall.

453 in quantum mechanics and is called *unitarity*¹⁴, and it assures us that no
454 information is lost. The unitarity law can be expressed as:

455 **Axiom 3.** *If two identical isolated systems are in different states, they stay in differ-
456 ent states, and they were in different states in the past.*

457 **1.4.2 Time-Development Operator**

458 Considering a system in the state $|\Psi(t)\rangle$, where the t indicates that the state
459 vector evolves over time, quantum motion equations allow us to obtain the
460 state at time t given the initial state:

$$|\Psi(t)\rangle = \mathbf{U}(t) |\Psi(0)\rangle. \quad (1.12)$$

461 Thanks to the operator $\mathbf{U}(t)$ we can know exactly the state vector $|\Psi(t)\rangle$ at *Determinism*
462 time t , given $|\Psi(0)\rangle$. This assertion can be rephrased as:

463 **Axiom 4.** *The temporal evolution of the state vector is deterministic.*

464 Quantum mechanics is still non-deterministic, because knowing the state
465 vector doesn't mean knowing the result of a measurement.

466 In order for $\mathbf{U}(t)$ to behave as we want, it has to:

- 467 • be a linear operator;
- 468 • respect reversibility.

469 The second constraint allows us to define the mathematical properties of
470 $\mathbf{U}(t)$. Considering two initially different states $|\Psi(0)\rangle$ and $|\Phi(0)\rangle$, since there
471 exists an experiment capable of certainly distinguishing the states, $|\Psi(0)\rangle$
472 and $|\Phi(0)\rangle$ must be orthogonal:

$$\langle\Psi(0)|\Phi(0)\rangle = 0.$$

473 The minus first law assures that during the entire temporal evolution of the
474 two systems, the state vectors $|\Psi(t)\rangle$ and $|\Phi(t)\rangle$ will continue to be distin-
475 guishable (orthogonal):

Conservation of Distinctions

$$\langle\Psi(t)|\Phi(t)\rangle = 0 \quad \forall t \geq 0.$$

476 If we rewrite this equation using formula 1.12, we obtain:

$$\langle\Psi(0)|\mathbf{U}^\dagger(t)\mathbf{U}(t)|\Phi(0)\rangle = 0.$$

477 From this we can see that $\mathbf{U}^\dagger(t)\mathbf{U}(t)$ must behave as the identity operator,
478 that is:

$$\mathbf{U}^\dagger(t)\mathbf{U}(t) = \mathbf{I}. \quad (1.13)$$

479 An operator that behaves as \mathbf{U} is *unitary*.

480 **Principles 5.** *The temporal evolution of state vectors is unitary.*

481 From the unitarity of \mathbf{U} descends the *conservation of overlaps*: the overlap
482 between two states (their inner product), subjected to the same temporal-
483 development operator, is preserved over time.

¹⁴ We will see in the next paragraph the reason for this name

484 1.4.3 The Hamiltonian

485 Often, in classical physics, a motion law is the result of a differential equation
 486 where we have exchanged a finite time interval with an infinite number of
 487 infinitesimal intervals.

Continuity 488 In quantum mechanics we can follow the same path and consider time
 489 intervals ϵ close to zero. In this scenario, after an ϵ amount of time, the state
 490 vector will change slightly and “smoothly”, and the operator $\mathbf{U}(\epsilon)$ will be
 491 very similar to the identity. We can rewrite $\mathbf{U}(\epsilon)$ in order to highlight the
 492 difference with the identity \mathbf{I} as:

$$\mathbf{U}(\epsilon) = \mathbf{I} - i\epsilon\mathbf{H}. \quad (1.14)$$

493 For now, i is a mere scale factor that later will help us recognize in \mathbf{H} the
 494 quantum version of the classical Hamiltonian.

495 We can now express the infinitesimal evolution of a quantum system by
 496 combining equations 1.12 and 1.14:

$$|\Psi(\epsilon)\rangle = |\Psi(0)\rangle - i\epsilon\mathbf{H}|\Psi(0)\rangle.$$

497 Bringing to the left the time interval:

$$\frac{|\Psi(\epsilon)\rangle - |\Psi(0)\rangle}{\epsilon} = -i\mathbf{H}|\Psi(0)\rangle.$$

498 Now considering the limit for $\epsilon \rightarrow 0$, we can see in the left member the time
 499 derivative of the state vector:

$$\frac{\partial |\Psi(t)\rangle}{\partial t} = -i\mathbf{H}|\Psi(0)\rangle.$$

500 Before using \mathbf{H} as the quantum Hamiltonian, we have to verify the di-
 501 mensional correctness. As in classical mechanics, the Hamiltonian is the
 502 mathematical construct that represents the energy. In our formula, however,
 503 ignoring the state vector, we have the inverse of time on the left and the
 504 energy on the right. To resolve this problem, let's introduce an important
 505 physical constant: the reduced Planck constant, \hbar .

506 The equation becomes:

$$\hbar \frac{\partial |\Psi\rangle}{\partial t} = -i\mathbf{H}|\Psi\rangle \quad \text{or} \quad \frac{\partial |\Psi\rangle}{\partial t} = \frac{-i\mathbf{H}|\Psi\rangle}{\hbar}. \quad (1.15)$$

Time-dependent Schrödinger equation

507 The constant \hbar has units of $\text{kg} \cdot \text{m}^2/\text{s}$ and resolves the incompatibility between
 508 the two members. This equation is fundamental and is called the *generalized*
 509 *Schrödinger equation*, or time-dependent Schrödinger equation. If we know
 510 the Hamiltonian of an undisturbed system, we can know the evolution of
 511 the state vector.

512 If \mathbf{H} represents the energy of the system, we should be able to measure it,
 513 so \mathbf{H} has to be an observable. If \mathbf{H} is an observable, it must be a Hermitian
 514 operator; let's verify it. Starting from 1.13 and substituting \mathbf{U} with 1.14, we
 515 obtain:

$$(\mathbf{I} + i\epsilon\mathbf{H}^\dagger)(\mathbf{I} - i\epsilon\mathbf{H}) = \mathbf{I}.$$

516 Expanding to first order in ϵ , we find:

$$\mathbf{H}^\dagger - \mathbf{H} = 0 \Rightarrow \mathbf{H}^\dagger = \mathbf{H}.$$

517 We have concluded that \mathbf{H} is an Hermitian operator that represents an
 518 observable: the energy of the system. Eigenvalues of \mathbf{H} are the results of all
 519 possible direct measurements of the energy of the system.

Quantum Hamiltonian

520 1.4.4 Commutators

521 In a system that evolves with time, we expect that the expectation values for
 522 a certain observable \mathbf{L} will also change. Thanks to equation 1.11 on page 15,
 523 we can write explicitly the time dependence of expectation values:

$$\langle \mathbf{L} \rangle = \langle \Psi(t) | \mathbf{L} | \Psi(t) \rangle.$$

524 The time derivative¹⁵ is:

$$\frac{d}{dt} \langle \Psi(t) | \mathbf{L} | \Psi(t) \rangle = \langle \dot{\Psi}(t) | \mathbf{L} | \Psi(t) \rangle + \langle \Psi(t) | \mathbf{L} | \dot{\Psi}(t) \rangle.$$

525 Substituting bra and ket with the time-dependent Schrödinger equation 1.15
 526 (namely $|\dot{\Psi}(t)\rangle = -\frac{i}{\hbar} \mathbf{H} |\Psi(t)\rangle$), we obtain:

$$\frac{d}{dt} \langle \Psi(t) | \mathbf{L} | \Psi(t) \rangle = \frac{i}{\hbar} \langle \Psi(t) | \mathbf{H}\mathbf{L} | \Psi(t) \rangle - \frac{i}{\hbar} \langle \Psi(t) | \mathbf{L}\mathbf{H} | \Psi(t) \rangle.$$

527 That can be rewritten as:

$$\frac{d}{dt} \langle \Psi(t) | \mathbf{L} | \Psi(t) \rangle = \frac{i}{\hbar} \langle \Psi(t) | [\mathbf{H}, \mathbf{L}] | \Psi(t) \rangle.$$

528 The quantity $\mathbf{HL} - \mathbf{LH}$ is called the *commutator*, and since, in general, the
 529 product between operators (matrices) is not commutative, the commutator
 530 is not zero (when it is zero, we say that \mathbf{H} and \mathbf{L} commute). Commutators
 531 are important in physics, and the commutator between two operators, in this
 532 case \mathbf{H} and \mathbf{L} , is denoted by:

$$\mathbf{HL} - \mathbf{LH} = [\mathbf{H}, \mathbf{L}].$$

533 With the commutator we can express concisely the derivative of the expec-
 534 tation value for the observable \mathbf{L} :

$$\frac{d}{dt} \langle \mathbf{L} \rangle = \frac{i}{\hbar} \langle [\mathbf{H}, \mathbf{L}] \rangle \quad (1.16)$$

535 or equivalently:

$$\frac{d}{dt} \langle \mathbf{L} \rangle = -\frac{i}{\hbar} \langle [\mathbf{L}, \mathbf{H}] \rangle. \quad (1.17)$$

536 This equation links variations of the expectation values of an observable
 537 (\mathbf{L}) to the expectation values of another physical observable $(-\frac{i}{\hbar} [\mathbf{L}, \mathbf{H}])$ ¹⁶.

¹⁵ Derivative of a product: \mathbf{L} doesn't depend on time and the dot denotes the time derivative (Newton notation).

¹⁶ It is possible to demonstrate that if \mathbf{L} and \mathbf{H} are Hermitian, then $[\mathbf{L}, \mathbf{H}]$ is also Hermitian.

538 **1.4.5 Conservation of Energy**

539 In quantum mechanics, when we say that a quantity is conserved, we mean
 540 that the expectation value of that quantity doesn't change. If we look at
 541 equation 1.17, the condition for the expectation value not to change is that
 542 the commutator between this quantity and the Hamiltonian is zero. It is
 543 possible to demonstrate that:

544 **Theorem 3.** *Having an observable \mathbf{Q} , if $[\mathbf{Q}, \mathbf{H}] = 0$, then every power satisfies*
 545 *$[\mathbf{Q}^n, \mathbf{H}] = 0$. This means that the expectation value $\langle \mathbf{Q} \rangle$ is conserved, and any*
 546 *power of the expectation value $\langle \mathbf{Q}^n \rangle$ does not change with time.*

547 The most obvious quantity that is conserved is the Hamiltonian \mathbf{H} and,
 548 since every operator commutes with itself, we always have:

$$[\mathbf{H}, \mathbf{H}] = 0.$$

549 We can conclude that, under very general conditions, energy is conserved
 550 in quantum mechanics.

551 **1.5 CONCLUSIONS**

552 We conclude this chapter with a recap of what we have discovered in these
 553 pages, trying to put everything together to answer the question that opened
 554 this chapter: what are the physical limits of quantum computing, and why
 555 must our algorithms be reversible?

556 We started the chapter with an experiment that shows that quantum me-
 557 chanics is not deterministic. We can, however, make some predictions if we
 558 consider the expectation value of a measurement instead of a single result.

559 We have built state vectors and understood their mathematical meaning,
 560 focusing on the fact that knowing the state vector doesn't allow us to know
 561 the result of a measurement. We have defined the inner product between
 562 state vectors, observed that it is a measure of the overlap between states,
 563 and concluded that two distinguishable states must be orthogonal.

564 We have linked a state vector to the result of a measurement –to be precise,
 565 to the average of the results of multiple measurements– with machines, Her-
 566 mitian operators that represent observables. We have built the spin operator
 567 and used it to predict the result of a simple experiment, showing how the
 568 theory we have built so far is in accordance with experimental results.

569 Our introduction continues with the analysis of the temporal evolution of
 570 a quantum system. We have described the evolution of a state vector with an
 571 unitary operator; the application of this operator to a state vector produces
 572 the new state in which the system will be. We understood that the tempo-
 573 ral evolution of the state vector is deterministic and that indeterminacy is
 574 caused only by the act of measuring.

575 Considering infinitesimal time intervals, we have deduced the time-dependent
 576 Schrödinger equation and, thanks to this equation, we have shown how to
 577 describe the temporal evolution of expectation values for a certain observ-
 578 able. During this analysis, we also introduced the Hamiltonian of the system,
 579 a Hermitian operator that describes the energy of the system.

580 The discussion ends with a comforting result: as in classical physics, the
 581 energy of a closed system is conserved. We have obtained this result by pre-
 582 senting the commutator and linking the temporal evolution of an observable
 583 with the commutator between the observable and the Hamiltonian (energy)
 584 of the system. The commutator of the Hamiltonian with itself is trivially
 585 zero, so the expectation value for the energy doesn't change.

586 All the information that we have learned allows us to understand the con-
 587 straint of writing only reversible algorithms for quantum-gate-based quan-
 588 tum computers. Quantum gates operate on qubits through physical transfor-
 589 mations¹⁷. These transformations, like all transformations in quantum me-
 590 chanics, are described by unitary operators that are intrinsically reversible.
 591 This means that all quantum gates are reversible.

592 In other words, we can build only quantum gates that, having as input
 593 different (distinguishable) states, return orthogonal states; also, due to the
 594 conservation of overlaps, the inner product between input states is conserved
 595 during the quantum gate transformation.

596 Reversibility doesn't mean that we can go forward and backward in time
 597 as we please, but that all quantum gates express injective functions: if we
 598 know the output, we can know the input, or in more physical terms, if
 599 we know the final state of qubits¹⁸ and the transformations applied to this
 600 system (i.e., those implemented by the quantum gates), we can determine
 601 the initial state.

602 Since every quantum algorithm has to be implemented as a path through
 603 quantum gates, and every quantum gate is reversible, the algorithms as a
 604 whole must also be reversible.

¹⁷ How depend strongly on the particular physical implementation.

¹⁸ This is a complex system (composed of more than one qubit); to fully understand these systems, we should take into account entanglement. Since our discussion is already quite long, and the temporal evolution of an entangled system is still unitary (reversible), we exclude entanglement from our introduction.

605 2 | QUANTUM GATE

606 | 3 QUANTUM ANNEALING

607 | 4 ONTOLOGY

608 In this chapter we explain what kind of knowledge base (KB) an ontology
609 is, how to build an ontology, and why this knowledge representation is
610 important. To clarify and demonstrate why ontologies are useful, we present
611 an example of an important ontology, briefly discussing its utility.

612 The rest of the chapter is about reasoning on ontologies; we discuss the
613 semantics of the formal language used to represent knowledge, what we
614 mean when saying interpretation of a KB, and the complexity of finding an
615 interpretation.

616 4.1 KNOWLEDGE BASE

617 In the field of information technologies, an ontology is a structured represen-
618 tation of knowledge about a certain domain of interest; however, the study
619 of knowledge began much before informatics. To better understand what an
620 ontology is, let's start with the philosophical definition and then point out
621 the differences between this vision and the IT one.

622 4.1.1 Ontology in philosophy

623 Ontology was born as a branch of philosophy. In this context it is the sci-
624 ence of what is, of the kinds and structures of objects, properties, events,
625 processes, and relations in every area of reality[1].

626 The goal of an ontology is to give a definitive and exhaustive classification
627 of entities in all spheres of being. With the term "definitive" we mean that
628 an ontology should answer questions such as: "What classes of entities are
629 needed for a complete description and explanation of all the goings-on in the
630 universe?" With the term "exhaustive", instead, we mean that all types of
631 entities and relations between these entities are included in our ontology[1].

632 4.1.2 Ontology in computer science

633 Thanks to the advent of the internet and the development of bigger and
634 bigger software used by bigger and bigger groups of users, what we might
635 call the Tower of Babel problem emerged. Each research group develops
636 its KB with terms and concepts shared and accepted only inside the group.
637 For example, different databases may use identical labels but with different
638 meanings, and the same meaning may be expressed with different names[1].

639 To address the incompatibility problem between software, databases, and
640 research groups, ontologies have become an important research topic in com-
641 puter science where the goal is to define standards for data exchange, infor-
642 mation integration, and interoperability[2].

643 In this field the term ontology gains a new meaning:

644 **Definition 1.** *Ontologies represent a formal and explicit specification of a shared*
 645 *conceptualization[3].*

646 In this definition the keywords are:

647 **CONCEPTUALIZATION** an ontology creates an abstract model identifying
 648 and defining only the relevant concepts;

649 **EXPLICIT** the types of concepts and constraints on their use are explicitly
 650 defined;

651 **FORMAL** an ontology should be machine-readable;

652 **SHARED** the knowledge represented by the ontology has to be accepted
 653 by a group of people, ideally by everyone.

654 When we use an ontology to represent knowledge we are describing a
 655 graph where entities are bound together through relationships, and classi-
 656 fied according to a formal description of the world[4]. Knowledge bases
 657 expressed with this formalism are divided into two components[5]:

658 **T-BOX** stores a set of universally quantified assertions (inclusion asser-
 659 tions) stating general properties of concepts and roles;

660 **A-BOX** contains assertions on individual objects (instance assertions).

661 We can see some similarities between an ontology and a database: the T-
 662 Box can be seen as the Entity-Relation schema and the A-Box as the set of all
 663 entries of the database. There is, however, a logical difference between the
 664 world represented by an ontology and the world represented by a database.

665 Databases make the *closed world assumption*: everything that is not present
 666 in the database is automatically false; for example, if a person does not
 667 appear in a bank registry it means that that person is not a client of the
 668 bank.

669 Ontologies, on the other hand, make the *open world assumption*[6], which
 670 means, for example, that we can assert that a certain person is a parent even
 671 if we have not specified any son or daughter.

672 4.1.3 OWL Language

673 OWL 2 Web Ontology Language is an ontology language for the Semantic
 674 Web with a formally defined meaning[7]. Thanks to OWL we can model
 675 classes and relations between classes (T-Box) and individuals with their spe-
 676 cific properties and relations between individuals (A-Box). The T-Box is the
 677 conceptualization of the world, while the A-Box is a certain instance of the
 678 world we have modelled in the T-Box.

679 OWL is a declarative language and defines the state of the world in a
 680 logical way. In particular, we are interested in OWL DL where the meaning
 681 of ontologies expressed with this language is assigned in a Description Logic
 682 style. OWL DL is, therefore, decidable and an appropriate tool (so-called

683 reasoner) can then be used to infer further information about that state of
 684 the world[7].

685 OWL per se doesn't specify any syntax, it states only what can or cannot
 686 be expressed in an ontology. The World Wide Web Consortium (W3C) stan-
 687 dardizes various syntaxes, some inspired by functional languages, others
 688 more suitable for storing on web pages. The only syntax that must be imple-
 689 mented by all tools to be compliant with the OWL standard is the RDF/XML
 690 syntax[7] (examples of this syntax are provided in 4.2.1).

691 4.1.4 Importance of ontologies

692 Ontologies are important in various fields, from interoperability to machine
 693 learning.

694 In the Semantic Web context, ontologies are a main vehicle for data integra-
 695 tion, sharing, and discovery[8]. Different research groups can use the same
 696 ontology to share a unified vocabulary that helps build common knowledge
 697 and helps to better integrate the results obtained by each group.

698 In a more commercial scenario, an ontology can be used as a transla-
 699 tion layer between different databases or software that are built by different
 700 teams and use different vocabularies.

701 In the machine learning field an ontology could be used to support the
 702 sharing and reuse of formally represented knowledge among AI systems[3].
 703 In the last years, we have become used to training AI agents on unstructured
 704 data, but formal knowledge can help to fine-tune these models or to check
 705 their answers.

706 4.2 EXAMPLE ONTOLOGIES

707 To help understanding the structure of ontologies and to show a practical
 708 example of ontology, we present two ontologies: a simple ontology about
 709 family relationships and DOLCE, a foundational ontology.

710 4.2.1 Simple ontology

711 This simple ontology about parental relationships shows the basic structure
 712 of an ontology, helping to understand the graph structure of these KBs and
 713 the relations between the T-Box and A-Box.

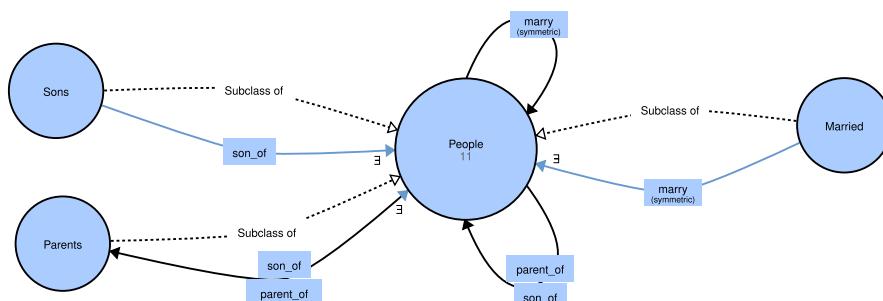


Figure 4.1: Graph for T-Box

In figure 4.1, we can see the T-Box of the ontology: this structure specifies what our domain of interest is, and what entities could possibly populate our world. This ontology is about people, so the main class/concept is `People`: this class has several subclasses that represent parents, children, and married people. We can assert that a person belongs to the married class without specifying the partner (open world assumption) but we can also infer that a person belongs to the parents class because we have created a relationship of type `parent_of` between that person and another person.

OWL allows us to express rules to infer when a member of a class belongs also to another class. The following code shows (in the RDF/XML syntax) the definition of the class `Parent`¹:

```

1 <owl:Class rdf:about="http://people#Parent">
2   <owl:equivalentClass>
3     <owl:Restriction>
4       <owl:onProperty rdf:resource="http://people#parent_of"/>
5       <owl:someValuesFrom rdf:resource="http://people#Person"/>
6     </owl:Restriction>
7   </owl:equivalentClass>
8   <rdfs:subClassOf rdf:resource="http://people#Person"/>
9 </owl:Class>
```

Listing 1: Definition of parents

At line 8 we can see that `Parent` is a subclass of `People`, and at lines 4 and 5 it is specified that a parent is a person that is `parent_of` another person.

From figure 4.1 we can also see some properties of the relations:

- relation `marry` is symmetric;
- relation `parent_of` is the inverse of `son_of`;
- we can specify a domain and a range for relations.

OWL gives us constructs for all of these specifications (and other more complex ones).

Now we can populate the ontology by adding individuals and relations between individuals. For this small example we take inspiration from the Simpson family, and in the family tree on the right we can see the small portion of the family represented. To show what we mean by open world assumption we have asserted that Jackie is a married person even if in our representation there is no husband.

Our ontology covers a small domain, the types of entities that populate our model are very limited; the next example shows the commitment of engineering an ontology to represent virtually anything in the universe.

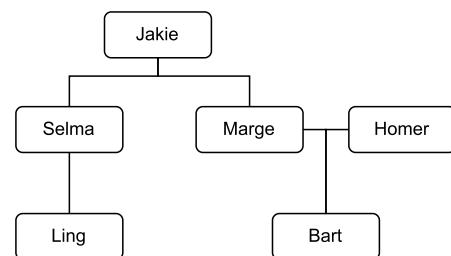


Figure 4.2: Simpson family tree

¹ The complete code of the ontology can be seen at [url](#).

747 **4.2.2 DOLCE ontology**

748 DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) is a
 749 top-level (foundational) ontology[9]; this means that this ontology describes
 750 fundamental aspects of reality and should be used as a base for constructing
 751 an ontology about a particular domain of interest. For this reason DOLCE
 752 defines only the T-Box; the user will then expand the T-Box with specific
 753 classes and relations of interest, and lastly will populate the A-Box.

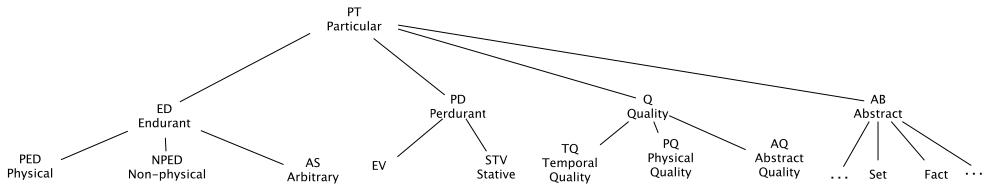


Figure 4.3: First layer of DOLCE taxonomy

754 **STRUCTURE OF DOLCE:** in DOLCE we can model the modification of ob-
 755 jects during time; for this reason DOLCE distinguishes between endurants
 756 and perdurants. Endurants may acquire and lose properties and parts through
 757 time, perdurants are fixed in time[9]. With a simplification we can see en-
 758 durants as the physical entities that are modified by the passing of time (like
 759 objects, animals, and people) and perdurants as events that, once they have
 760 passed, cannot be changed anymore (like a tennis match or a conference).

761 The relation connecting endurants and perdurants is called participation.
 762 A physical entity can be in time by participating in a perdurant, and perdu-
 763 rants happen in time by having endurants as participants[9].

764 Another important aspect of DOLCE is the way we attribute a property to
 765 an entity; this is done by using qualities, which are what can be perceived
 766 and measured. To do so we can assert that a certain entity has a specific
 767 quality and then, when it is possible, quantify that quality.

768 **IMPORTANCE OF DOLCE:** foundational ontologies can be useful in several
 769 fields, from conceptual modeling to natural language processing. DOLCE,
 770 today, is used in a variety of domains where it provides the general cate-
 771 gories and relations needed to give a coherent view of reality[9].

772 **4.3 RERASONING ON ONTOLOGY**

773 In 4.1.3 we have introduced the standard language to encode an ontology; in
 774 order to infer new information, starting from the one we already have, we
 775 need to better specify the semantics of OWL DL.

776 **4.3.1 SROIQ DL**

777 The semantics of OWL DL extends the semantics of the description logic
 778 (DL) *SROIQ* to provide support for datatypes and punning[10]. For con-

779 structs available both in OWL DL and in \mathcal{SROIQ} the semantics correspond
 780 exactly.

781 Description logics allow the modeling of the domain of interest with three
 782 kinds of entity: concepts, roles, and individual names. These entities cor-
 783 respond to unary predicates, binary predicates, and constants in first-order
 784 logic[6]. From the point of view of ontology and OWL, concepts are classes,
 785 roles are relationships, and individual names are the individuals that can
 786 belong to one or more classes.

787 \mathcal{SROIQ} is one of the most expressive description logics where we have
 788 constructors for:

- 789 • transitive roles: \mathcal{S}
- 790 • role inclusions, local reflexivity, universal role, symmetry, asymmetry,
 role disjointness, reflexivity, and irreflexivity: \mathcal{R} ;
- 791
- 792 • nominals: \mathcal{O} ;
- 793
- 794 • inverse roles: \mathcal{I} ;
- 795 • qualified number restrictions: \mathcal{Q} .

795 For example, we can construct the ontology shown in figures 4.1 and 4.2
 796 with a set of assertions like:

797 `person(selma) married(jackie) parent_of(marge, bart)`

798 Each of these statements is called an axiom and the set of all axioms consti-
 799 tutes our KB.

800 4.3.2 Interpretation of a knowledge base

801 An interpretation I consists of a domain Δ^I and an interpretation function $.^I$
 802 that maps:

$$\begin{aligned} \text{concept } A &\rightarrow A^I \subseteq \Delta^I \\ \text{role } R &\rightarrow R^I \subseteq \Delta^I \times \Delta^I \\ \text{named individual } a &\rightarrow a^I \in \Delta^I \end{aligned}$$

803 In other words I assigns a fixed meaning to all entities in the KB[6]. By
 804 having a fixed meaning, we can say if an axiom α holds in I or not; in the
 805 first case we say that I satisfies α and we write $I \models \alpha$.

806 If all axioms in an ontology are satisfied by I we say that I is a *model* of the
 807 ontology. An ontology is consistent if it accepts at least one model.

808 A reasoner should at least be capable of saying if an ontology is consistent,
 809 but we are also interested in querying knowledge to retrieve new informa-
 810 tion.

811 **QUERY INTERPRETATION:** Considering a KB K , a query q consists of ax-
 812 iom templates where \mathcal{SROIQ} axioms are composed of concept names, role
 813 names, and individual names, but also of concept variables, role variables,
 814 and individual variables. A solution for the query is an interpretation μ that

815 allows us to rewrite all variables in q with names; we denote with $\mu(q)$ the
816 result of the substitution.

817 The evaluation of q over K is a set of solutions μ with:[11]

$$818 \quad \{ \mu | K \cup \mu(q) \text{ is a } \mathcal{SROIQ} \text{ knowledge base and } K \models \mu(q) \}$$

819 In other words μ binds all free variables of q to names present in K [11].

820 A naive approach to find the solution to a query is to simply test for each
821 possible solution mapping μ , if $K \models \mu(q)$; however, in the worst case, the
822 number of mappings that have to be tested is exponential in the number of
823 variables in the query[11].

824 4.3.3 Complexity of reasoning

825 Since presenting an actual algorithm for reasoning on ontologies is out of
826 the scope of this work, we only give some hints about the reasons for the
827 complexity and then present the theoretical results.

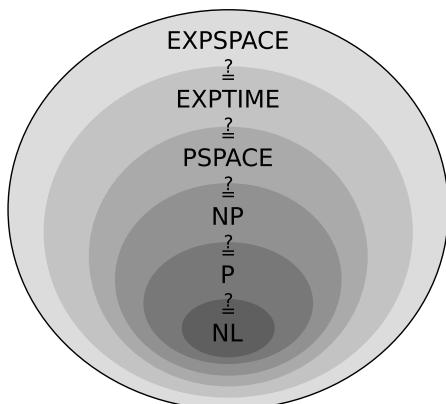
828 It is easy to convince oneself that the more axioms there are in an ontology,
829 the fewer interpretations exist that satisfy all axioms. On the other hand, if
830 an ontology has fewer models, the more axioms hold in all of them and the
831 more logical consequences follow from the ontology.

832 We can rephrase these two statements by saying that the semantics of de-
833 scription logics are *monotonic*: the more knowledge we embed in an ontology,
834 the more results it returns[6].

835 A more formal view is given in [12], where two *sources of complexity* are
836 identified:

- 837 • OR-branching: the presence of disjunctive constructors;
- 838 • AND-branching: the presence of qualified existential and universal
839 quantifiers.

840 The AND-branching is responsible for the exponential size of a single inter-
841 pretation, and the OR-branching is responsible for the exponential number
842 of different interpretations.



855 **Figure 4.4:** Complexity classes

856 To discuss the complexity of reasoning we take into account the
description logic \mathcal{ALC} ; this DL is a restriction of \mathcal{SROIQ} [6], so its
complexity is a lower bound for \mathcal{SROIQ} . It is possible to prove the
 \mathcal{ALC} [12], therefore also \mathcal{SROIQ} DL is at least PSpace-hard.

This result shows that, unless $\text{PSpace} = \text{PTime}$, the exponential time complexity of any algorithm that makes inference on an ontology cannot be improved .

857 For those interested in some numerical examples to better understand
858 what this class of complexity means in a real context, [13] presents the rea-
859 soner Hermit² and evaluates its performance on some real ontologies.

860 **4.4 CONCLUSIONS**

861 In this chapter we have explained what an ontology is and we have moti-
862 vated the interest in this field. We have shown both theoretically and with
863 examples what can be expressed in an ontology and what cannot. We have
864 formally defined what the interpretation of a KB is and showed what a query
865 and its results are.

866 Lastly, we have characterized the complexity of reasoning on ontologies.
867 This complexity is what motivated us to search for other paradigms to infer
868 new knowledge starting from an ontology. In the next chapters we will build
869 the tools necessary to achieve this goal.

² <http://www.hermit-reasoner.com/>.

870

Part II

871

TOOLS

5 | ENVIRONMENT SETUP

873 In this chapter we describe the environment, libraries and tools we use to
874 execute our tests.

875 In the following sections we install the SDKs to develop and interact with
876 quantum computers from IBM and D-Wave. We also present two other use-
877 ful tools to easily write optimization problems.

878 5.1 PYTHON ENVIRONMENT

879 The language used to interface with quantum computers is usually Python.
880 In this section we create a virtual environment in Python in order to commu-
881 nicate with the IBM quantum computer and the D-Wave quantum computer.

882 For our tests we manage Python environments with `conda`. Let's start by
883 creating the virtual environment named `quantum` and activating it with:

```
1 conda create --name quantum python=3.12 pip
2 conda activate quantum
```

884 For our tests and to follow the various examples presented both by IBM and
885 D-Wave, it is also useful to be able to run a Jupyter notebook. We can install
886 Jupyter with:

```
1 pip install jupyter
```

887 5.2 IBM QISKIT

888 To program a gate-based architecture and to access IBM quantum computers
889 we use the *Qiskit* software stack. The name Qiskit is a general term referring
890 to a collection of softwares for executing programs on quantum computers.

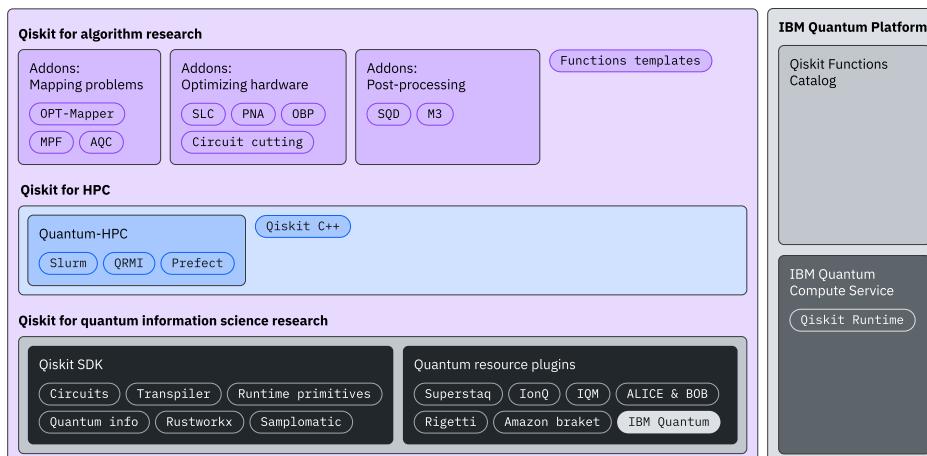


Figure 5.1: Qiskit software stack

891 The core components are *Qiskit SDK* and *Qiskit Runtime*. The first one is
 892 completely open source and allows the developer to define his circuit; the
 893 second one is a cloud-based service for executing quantum computations on
 894 IBM quantum computers.

895 **5.2.1 Hello World**

896 Following the IBM documentation¹ we can install the SDK and the Runtime
 897 with:

```
1 pip install qiskit matplotlib qiskit[visualization]
2 pip install qiskit-ibm-runtime
3 pip install qiskit-aer
```

898 Line 3 installs Aer, which is a high-performance simulator for quantum
 899 circuits written in Qiskit. Aer includes realistic noise models, and we will
 900 use it later to test our circuit.

901 Sometimes the Qiskit stack suffers from incompatibilities between the
 902 various software components that compose the environment. At the mo-
 903 ment of writing, the latest packages seem to work without any problem.
 904 For our tests we will use `qiskit: 2.2.3`, `qiskit-ibm-runtime: 0.43.1` and
 905 `qiskit-aer: 0.17.2`.

906 If the setup is successful we are now able to run a small test to build a Bell
 907 state (two entangled qubits). The following code assembles the gates, shows
 908 the final circuit and uses a sampler to simulate on the CPU the result of 1024
 909 runs of the program.

```
1 from qiskit import QuantumCircuit
2 from qiskit.primitives import StatevectorSampler
3
4 qc = QuantumCircuit(2)
5 qc.h(0)
6 qc.cx(0, 1)
7 qc.measure_all()
8
9 sampler = StatevectorSampler()
10 result = sampler.run([qc], shots=1024).result()
11 print(result[0].data.meas.get_counts())
12 qc.draw("mpl")
```

Listing 2: Building Bell state

910 **5.2.2 Transpilation**

911 Each Quantum Processing Unit (QPU) has a specific topology. We need to
 912 rewrite our quantum circuit in order to match the topology of the selected
 913 device on which we want to run our program. This phase of rewriting,
 914 followed by an optimization, is called transpilation.

¹ <https://quantum.cloud.ibm.com/docs/en/guides/install-qiskit>

915 Considering, for now, a fake hardware (so we do not need an API key)
 916 we can transpile the quantum circuit qc, from the code above, to match the
 917 topology of a specific QPU:

```

1  from qiskit_ibm_runtime.fake_provider import FakeWashingtonV2
2  from qiskit.transpiler import generate_preset_pass_manager
3
4  backend = FakeWashingtonV2()
5  pass_manager = generate_preset_pass_manager(backend=backend)
6
7  transpiled = pass_manager.run(qc)
8  transpiled.draw("mpl")

```

Listing 3: Transpilation

918 The following picture shows (5.2a) the quantum circuit that builds a Bell
 919 state, and (5.2b) the transpiled version where the Hadamard gate is replaced
 to match the actual topology of the QPU.

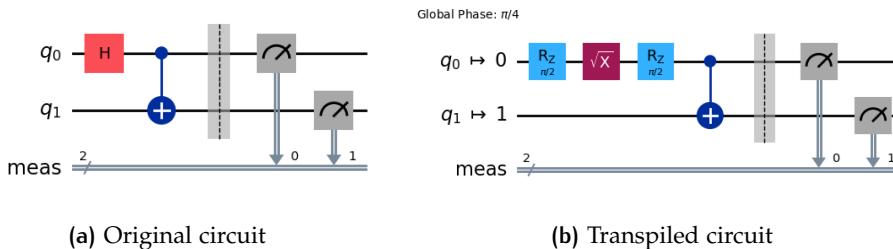


Figure 5.2: Transpilation example

920

921 5.2.3 Execution

922 To test our transpiled circuit we use Aer, which allows us to simulate also
 923 the noise of real quantum hardware. We can execute our program with:

```

1  from qiskit_aer.primitives import SamplerV2
2
3  sampler = SamplerV2.from_backend(backend)
4  job = sampler.run([transpiled], shots=1024)
5  result = job.result()
6  print(f"counts for Bell circuit : {result[0].data.meas.get_counts()}")

```

Listing 4: Simulated execution

924 If we look at the results of the execution we can observe that some answers
 925 present non-entangled qubits; this is caused by the (simulated) noise of the
 926 quantum device. A typical output of the execution could be:

```
> counts for Bell circuit : {'00': 504, '11': 503, '01': 10, '10': 7}
```

927 Where states 01 and 10 should not be present in an ideal execution with no
 928 errors.

929 5.2.4 A complete example on real hardware

930 5.3 D-WAVE OCEAN

931 To define an optimization problem that can be solved on a D-Wave quantum
 932 computer we use the Ocean software stack. Ocean also allows us to interact
 933 with D-Wave hardware, submit a problem, and simulate the execution on a
 classical CPU.

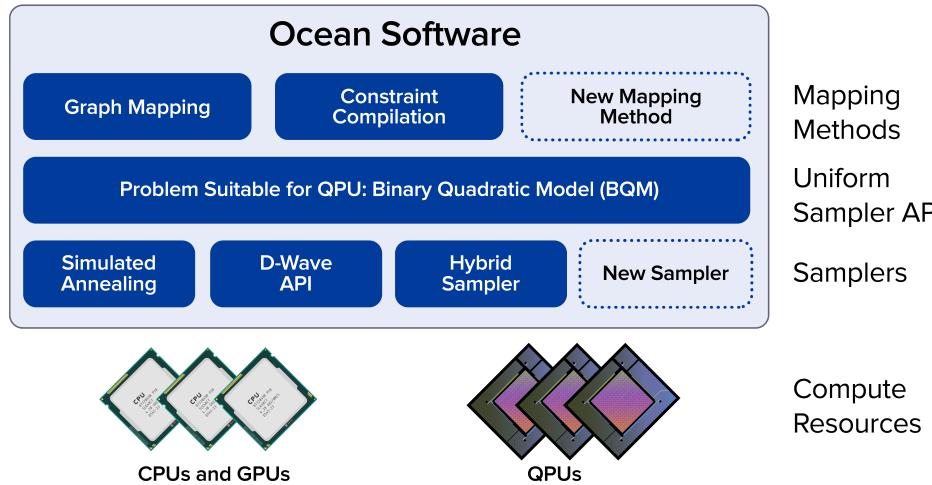


Figure 5.3: Ocean software stack

934
 935 All tools that implement the steps needed to solve your problem on a CPU,
 936 a D-Wave quantum computer, or a quantum-classical hybrid solver can be
 937 installed with:

```
1 pip install dwave-ocean-sdk
```

938 After the installation, running the command `dwave setup` will start an in-
 939 teractive prompt that guides us through a full setup. During the setup it
 940 is also possible to add an API token or connect to the D-Wave account to
 941 import a key directly to use the quantum hardware.

942 5.3.1 Hello World

943 To present a simple optimization program we consider the minimum vertex
 944 cover (MVC) problem. Given a graph $G = (V, E)$, the problem asks to find
 945 a subset $V' \subseteq V$ such that, for each edge $\{u, v\} \in E$, at least one of u or v
 946 belongs to V' , and the number of nodes in V' ($|V'|$) is the lowest possible.

947 The reduction from MVC to an Ising formulation is well known. The cost
 948 function that we want to minimize can be expressed by:

$$\text{cost} = \sum_{i=1}^{|V|} v_i + 2 \cdot \sum_{\{i,j\} \in E} (1 - v_i - v_j + v_i v_j)$$

949 where $v_i \in \{-1, 1\}$ and $v_i = 1$ means that $v_i \in V'$, otherwise $v_i = -1$.

950 Like all problems in Ising form we can express the cost as a symmetric
 951 matrix, so our function becomes

$$\text{cost} = v^T \times \mathbf{M} \times v$$

952 where v is the vector containing the binary variables v_i .

953 The figure shows an example graph (5.4a) and the corresponding matrix
 954 (5.4b) expressing the cost function.

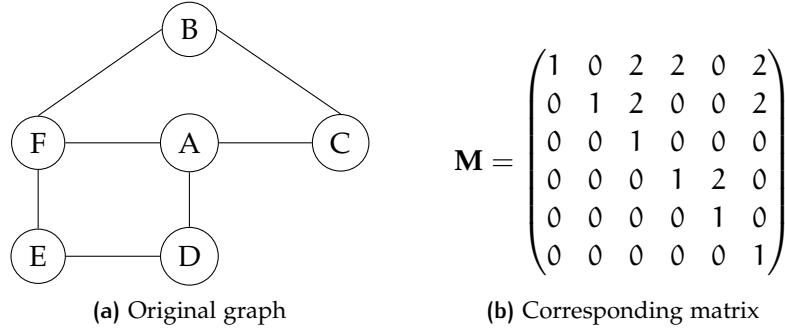


Figure 5.4: Ising formulation

955 The following code presents a possible implementation of the Ising model
 956 described above. We have defined two dictionaries to store the matrix coeffi-
 957 cients. The last line of code finds ten possible answers to the problem using
 958 the simulated annealing function implemented by D-Wave.

```

1 from dwave.samplers import SimulatedAnnealingSampler
2 linear = {'A': 1, 'B': 1, 'C': 1, 'D': 1, 'E': 1, 'F': 1}
3 quadratic = {('B', 'C'): 2, ('B', 'F'): 2, ('C', 'A'): 2, ('D', 'A'): 2,
4   ↪  ('E', 'D'): 2, ('E', 'F'): 2, ('F', 'A'): 2}
5 sampler = SimulatedAnnealingSampler()
5 result = sampler.sample_ising(linear, quadratic, num_reads=10)

```

Listing 5: Ising example

959 If we print the results with `print(result.aggregate())` we can observe
 960 something similar to this:

```

1   A  B  C  D  E  F energy num_oc.
2   0 -1 -1 +1 +1 -1 +1  -14.0      6
3   1 +1 +1 -1 -1 +1 -1  -14.0      4
4   ['SPIN', 2 rows, 10 samples, 6 variables]

```

961 The two different results represent the two correct answers to our particular
 962 instance of the MVC problem.

963 5.3.2 Example on real hardware

964 5.3.3 Minor embedding

965 5.4 PYQUBO AND QUBOVERT

966 In listing 5 we have manually built the matrix representing the function that
 967 we want to minimize. It can be useful to have some tools that allow us to

968 work at a higher level, defining cost functions like ?? that we defined in the
 969 section about quantum annealing (??).

970 Considering again the MVC problem, the objective function tends to minimize
 971 the number of nodes in our subset, while the penalty increases the
 972 cost if we leave out some edges. This interpretation allows us to transform
 973 the Ising model into the more familiar —from the point of view of a com-
 974 puter scientist— QUBO model, where all variables $x_i \in \{0, 1\}$. Let's see how
 975 PyQUBO and qubovert help us in this task.

976 5.4.1 PyQUBO

977 Reading from the documentation on the PyQUBO site², PyQUBO allows
 978 us to create QUBOs or Ising models from flexible mathematical expressions
 979 easily. Some of the features of PyQUBO are:

- 980 • Python based (C++ backend);
- 981 • Fully integrated with Ocean SDK;
- 982 • Automatic validation of constraints;
- 983 • Placeholder for parameter tuning.

984 We can install PyQUBO with `pip install pyqubo` and rewrite our MVC
 985 problem by defining the Hamiltonian that we want to minimize.

```

1  from pyqubo import Binary, Placeholder, Constraint
2  from dwave.samplers import SimulatedAnnealingSampler
3
4  A, B, C, D, E, F = Binary('A'), Binary('B'), Binary('C'), Binary('D'),
5    ↳ Binary('E'), Binary('F')
6
7  H_objective = (A + B + C + D + E + F)
8  H_penalty = Constraint(((1 - A - C + A*C) +
9    (1 - A - D + A*D) +
10   (1 - A - F + A*F) +
11   (1 - B - C + B*C) +
12   (1 - B - F + B*F) +
13   (1 - D - E + D*E) +
14   (1 - E - F + E*F)), label='cnstr0')
15
16  L = Placeholder('L')
17  H = H_objective + L*H_penalty
18  H_internal = H.compile()
19  bqm = H_internal.to_bqm(feed_dict={'L': 2})
20
21  sampler = SimulatedAnnealingSampler()
22  result = sampler.sample(bqm, num_reads=10)
```

Listing 6: Rewriting MVC with pyQUBO

² <https://pyqubo.readthedocs.io/en/latest/>

986 Listing 6 presents a possible re-implementation of listing 5, where we also
 987 see how PyQUBO interfaces with the Ocean SDK (line 17), and how to create
 988 (lines 14–16) and instantiate (line 17) a parametric Hamiltonian.

989 **5.4.2 qubovert**

990 As written in the documentation³, qubovert is the one-stop package for for-
 991 mulating, simulating, and solving problems in boolean and spin form. Using
 992 our nomenclature, boolean and spin form are respectively QUBO and Ising
 993 form.

994 Qubovert allows us to define various types of optimization problems that
 995 can be solved by brute force, with qubovert’s simulated annealing, or with
 996 D-Wave’s solver. Models defined in qubovert are:

997 **QUBO**: Quadratic Unconstrained Boolean Optimization;
 998 **QUSO**: Quadratic Unconstrained Spin Optimization (Ising model);
 999 **PUBO**: Polynomial Unconstrained Boolean Optimization;
 1000 **PUSO**: Polynomial Unconstrained Spin Optimization;
 1001 **PCBO**: Polynomial Constrained Boolean Optimization;
 1002 **PCSO**: Polynomial Constrained Spin Optimization.

1003 In addition to generic models, qubovert has a library of famous NP-complete
 1004 problems mapped to QUBO and Ising forms.

```

1 from qubovert import boolean_var
2 from dwave.samplers import SimulatedAnnealingSampler
3
4 A, B, C, D, E, F = boolean_var('A'), boolean_var('B'),
5   ↪ boolean_var('C'), boolean_var('D'), boolean_var('E'),
6   ↪ boolean_var('F')
7
8 model = A + B + C + D + E + F
9 model.add_constraint_OR(A, C, lam=2)
10 model.add_constraint_OR(A, D, lam=2)
11 model.add_constraint_OR(A, F, lam=2)
12 model.add_constraint_OR(B, C, lam=2)
13 model.add_constraint_OR(B, F, lam=2)
14 model.add_constraint_OR(D, E, lam=2)
15 model.add_constraint_OR(E, F, lam=2)
16
17 qubo = model.to_qubo()
18 dwave_qubo = qubo.Q
19 sampler = SimulatedAnnealingSampler()
20 result = sampler.sample_qubo(dwave_qubo, num_reads=10)
```

Listing 7: Rewriting MVC with qubovert

3 <https://qubovert.readthedocs.io/en/latest/index.html>

1005 Listing 7 shows a possible implementation of the MVC problem using the
1006 tools provided by qubovert. Qubovert allows us to express our problem as a
1007 PCBO; we use this formulation to express constraints in a more natural way.
1008 In our example we ensure that each edge is covered simply by enforcing that
1009 at least one of the nodes linked by the edge is present in the solution. This
1010 constraint is repeated for each edge in the graph (lines 7–13). To specify the
1011 Lagrange multiplier (equation ??) we use the keyword `lam`.
1012 Qubovert, like PyQUBO, can interface with the Ocean SDK, transforming
1013 a PCBO problem into a QUBO problem (line 15) and then rewriting it in the
1014 format accepted by the D-Wave solver (or sampler).

1015 **5.5 CONCLUSION**

1016 In this chapter we have set up an environment to run our future experiments
1017 and tests. We have also shown some small examples to present the main
1018 characteristics and test the tools we will use in our work.

1019 Following this setup allows anyone to recreate exactly the same configura-
1020 tion we use, avoiding (for what we know and test) incompatibilities between
1021 Python packages.

1022 **6** | QA-PROLOG

1023 **6.1 THE PROJECT**

1024 **6.2 PIPELINE**

1025 **6.3 UPDATE TO THE PROJECT**

1026 **6.4 RELATED WORK**

1027 **6.5 CONCLUSION**

1028

Part III

1029

EXPERIMENTS

1030 **7** | A QUANTUM ONTOLOGY

1031 **7.1** ONTOLOGY STRUCTURE

1032 **7.2** PROLOG VERSION

1033 **7.3** INFERENCE ON THE ONTOLOGY

1034 **7.4** CONCLUSION

1035 **8** | QAOA

1036 **8.1 QAOA**

1037 **8.2 FROM QUBO TO PAULI OPERATOR**

1038 **8.3 EXPERIMENTS**

1039 **8.4 CONCLUSION**

9 | CONCLUSION