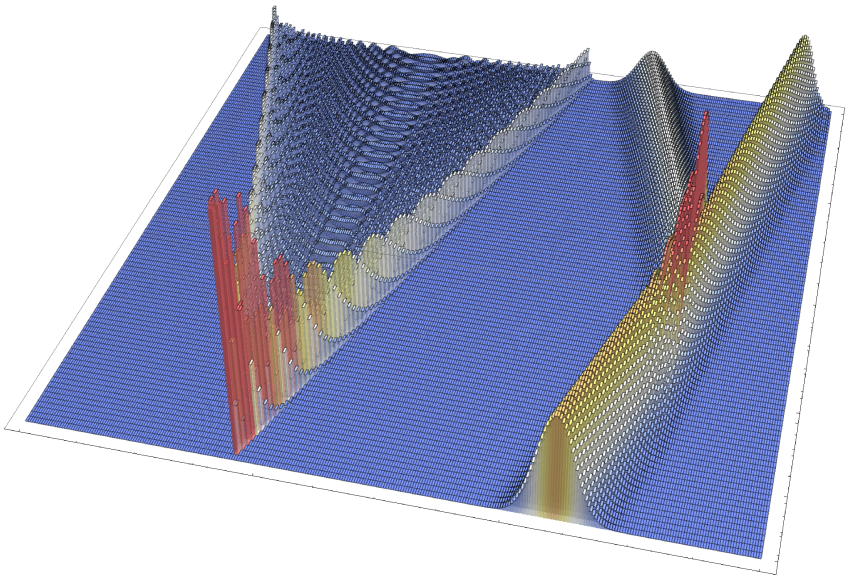
UNIVERSITÀ DEGLI STUDI DI PAVIA

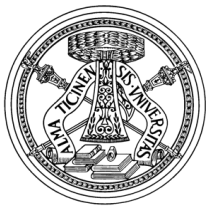
DOTTORATO DI RICERCA IN FISICA XXV CICLO

A Quantum Cellular Automata Framework for Quantum Fields Dynamics

Alessandro Tosini



TESI PER IL CONSEGUIMENTO DEL TITOLO



UNIVERSITÀ DEGLI STUDI DI PAVIA

DOTTORATO DI RICERCA IN FISICA XXV CICLO

A Quantum Cellular Automata Framework for Quantum Fields Dynamics

Alessandro Tosini

Advisor: Prof. Giacomo Mauro D'Ariano
Referees: Prof. Giovanni Amelino-Camelia and Prof. João Magueijo

SUBMITTED TO THE GRADUATE SCHOOL IN PHYSICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOTTORE DI RICERCA IN FISICA
DOCTOR OF PHILOSOPHY IN PHYSICS
AT THE
UNIVERSITY OF PAVIA

Cover
The one dimensional Dirac Quantum Cellular Automaton evolution. On the left the free evolution for a localized state. On the right the scattering of a smooth state with a potential barrier.

Title
A Quantum Cellular Automata Framework for Quantum Fields Dynamics

Author
Alessandro Tosini

PhD thesis — University of Pavia

Typeset in LaTeX
Printed in Pavia, Italy
October 2013

ISBN 978-88-95767-67-3

A Quantum Cellular Automata Framework for Quantum Fields Dynamics

Alessandro Tosini
October 2013

Contents

List of Figures

CHAPTER 1

Introduction

According to *operationalism* the meaning of a concept is given by a method of measurement for it. This intuition was first formalized in 1927 by the American physicist P.W. Bridgman [1] who stated "we mean by any concept nothing more than a set of operations; the concept is synonymous with the corresponding set of operations". Beside its obvious philosophical impact, operationalism was many times the germ for reconsidering the foundations of existing physical theories pointing out their weakness and paving the way to new more in principle coherent description of physical phenomena.

At the beginning of the XX century the conflict between the electromagnetism and the laws of mechanics was solved by the deep reconsiderations of the notions of space and time originated by the Einstein synchronization protocol. This taught us that the only reasonable way to define a physical entity is through an experimental measurement procedure. Einstein relativity also attested the relevance of physical principles and their ability in predicting unexpected new phenomena. A prime example of the predicting power of solid physical principles has been the discovery of antimatter. Guided only by the basic rules of *Quantum Theory* (QT) along with the *relativity principle*, Dirac wrote the equation which first predicted antiparticles, later discovered in the Anderson's experiment. The route from the elementary principles to the sophisticated laws of physics may require a tangled formal theoretical derivation, which usually involves abstract mathematics or even the development of new mathematical tools. The challenge is thus to pursue the principles with the guidance of logic while mathematically deriving all their physical consequences.

In the last decades it was QT that benefitted from such an operational ap-

proach with a renewed interest of the physicists community in his foundational aspects. The need for a deeper understanding of QT in terms of fundamental principles was already clear to von Neumann who expressed his dissatisfaction with the mathematical formulation based on the abstract framework of Hilbert spaces and self-adjoint operators with the words "I don't believe in Hilbert space anymore". The mathematical structure of Hilbert spaces (or C*-algebras), which is adopted as a prescription that "works well" if used as a black box to produce experimental predictions, should emerge as a consequence of postulates formulated in the language of physics, namely referring to notions like physical system, experiment and physical process. Hence QT should emerge from a set of rules that allow the experimenter to predict future events on the basis of suitable tests, having local control and low experimental complexity. Driven by this belief Ludwig initiated an operational axiomatization program [2] with the postulates corresponding to rules about how preparation and measurement devices combine to give the probabilities of experimental outcomes. However, Ludwig did not succeed in deriving the whole apparatus of QT from purely operational axioms. In the following years the raising field of Quantum Information provided a significant boost in the understanding of QT and new light was shed on *entanglement* and *nonlocality*. These achievements enforced the idea that QT could have been regarded as a theory of information, namely asserting basic properties of information-processing, such as the possibility or impossibility to carry out certain tasks by manipulating physical systems. In this scenario Lucien Hardy [3] reopened the debate about the operational axiomatizations with the renewed interest of the community [4–9]. The finite dimensional mathematical structure of QT was thus derived by Giacomo Mauro D'Ariano, Paolo Perinotti and Giulio Chiribella from purely operational and informational principles [10] the most crucial ones being *causality* and the conservation of information. However, the principles that thoroughly interpret the probabilistic structure of QT cannot explain the emergence of dynamical quantities and the equations of their evolution.

The research activity carried out in this thesis stems from the last consideration and investigates the consequences of assuming an elementary informational framework not only for predicting the outcomes' probability of an experiment but also as a description of the physical dynamics occurring between the preparation and the measurement devices. Although physical quantities still lie outside the sole quantum framework, QT will be considered as the primitive notion to be extended for many reasons. QT is arguably regarded as the most successful quantitative theory of nature and no violation has ever been detected in any laboratory despite a huge number of experimental tests involving light, atoms, molecules, and solids, as well as nuclei, electrons, and other subatomic particles. Moreover QT lies at the core of *Quantum Field Theory* (QFT), the most detailed description of the dynamics of physical systems available nowadays. Finally, we already know QT to be a theory of information processing.

Taking QT as the elementary probabilistic description of any physical experiment, and pursuing the general idea that also the "mechanical" side of physics should be described in terms of information processing, one need to identify new informational principles for seeking dynamics in a minimal quantum setting. Within this perspective the natural stage for dynamics—say spacetime—and the usual "objects" evolving in such a stage—say particles or more precisely quantum fields—are not primitive notions and should be recovered from the above minimal informational description. The long term aim of this program is thus a deep reconsideration of the foundations of QFT which firstly cured the tension between QT and Relativity providing a unified model for describing the dynamics of physical systems.

Despite its impressive predictive power QFT still lacks a satisfactory interpretation and during the last two decades the conceptual aspects behind its mathematical structure became a very discussed topic both in physics and in philosophy of science (see Refs. [11–15]). Among the interpretative issues of QFT we can mention the ultraviolet divergences, the causality violation due to the Hamiltonian description leading to the wave-function superluminal tails [16] and the *localization problem* [11, 17]. The latter plays a major role in the still present tension between Relativity and QT since it prevents the formulation of a rigorous theory of measurement for quantum fields. The incompatibility between QT and relativistic spacetime sharpens in the attempt of developing a quantum theory of gravity. In order to overcome this incompatibility complete, but still not falsifiable, theories have been proposed as *String Theory* [18, 19] and *Loop Quantum Gravity* by Rovelli, Smolin and Ashtekar. Other alternative models of spacetime are the *causal sets* of Bombelli et al. [20], the *non-commutative spacetime* of Connes [21], the *quantized spacetime* of Synder [22], the *Deformed Special Relativity* of Camelia [23,24], Smolin and Magueijo [25]. Some consequences of these approaches are even considered for experimental tests, see for example the recent experiment proposals by Hogan [26, 27] and Brukner et al. [28] for probing a possible noncommutative spacetime. A guideline in the recent reconsiderations of spacetime are the Hawking-Bekenstein results about the finiteness of the black hole entropy [29, 30], which implies that the number of bits of information that "can be stored" is finite. This led to consider the chance that spacetime at the *Planck scale* could be discrete and that the amount of information in a finite volume should be finite.

The most rigorous attempt to reformulate QFT is the so called *Algebraic Quantum Field Theory*, formulated by Haag and Kastler in Ref. [31] and by Streater and Wightman in Ref. [32]. The key notion in this approach is the *principle of locality* as stated in the title *"Local Quantum Field Theory"* of the Haag popular book [33] on the subject. The only physical observables are the ones connected with finite regions of spacetime and the local algebra of observables on two space-like separated regions must commute. This principle resembles the Einstein causality and is the main relativistic ingredient of

algebraic QFT. This approach has some similarities with the aforementioned operational point of view regarding QFT as a probabilistic theory about local measurements in finite regions of spacetime. However, the notion of locality at the basis of the algebraic approach is strongly connected to the underlying notion of continuous spacetime which is assumed as primitive, with fields "operator valued distributions" over the Minkowski spacetime carrying a unitary representation of the Poincaré covering group.

Giacomo Mauro D'Ariano, one of the authors who derived QT from informational principles [10], has recently undertaken the program of reconsidering the foundations of QFT, along with the notion of spacetime, starting from an elementary theory of interacting quantum systems [34–38]. From an operational perspective, an informational theory of dynamics should describe any physical evolution via a quantum algorithm and, more importantly, the complexity of the algorithm cannot invalidate the full computability of the theory. Moreover the framework must ensure the universality of the physical law expressed by the algorithm. These two requirements translate into two features of the interacting quantum systems. On one hand systems must have minimal Hilbert-space dimension with a finite set of interacting systems carrying finite information. In view of recovering the notion of spacetime from the systems in interaction, and identifying a finite region of the emergent spacetime with a finite number of interacting systems, one could say that the volume-density of information must be finite. On the other hand we have the homogeneity of the computation, namely the graph made of the interacting systems must be topologically homogeneous. These considerations lead to the idea of replacing quantum fields with *Quantum Cellular Automata* and to consider spacetime as emergent from the automata discrete *causal network*.

Cellular Automata (CAs) were first introduced by Stanislaw Ulam and John von Neumann [39] in the early 1950s and a big boost to their popularization came from the John Conway's Game of Life [40] in 1970 and later by the Wolfram work in Ref. [41]. A CA consists in a discrete set of identical cells where each cell can be in a fine set of states. The cells' configuration, namely the state of each cell, is then evolving in discrete time steps according to a transition rule satisfying certain notions of reversibility, locality and causality. In other words the single cell state at time $t+1$ just depends on the state of the neighbor cells at time t ensuring a finite speed propagation. Since CAs clearly depict a typical model of physical evolution Richard Feynman considered their extension to the quantum world [42]. This resulted in the so called Quantum Cellular Automata (QCAs) with cells of quantum systems interacting with a finite numer of other cells via a unitary operator describing the single step evolution. The first QCA appeared in Ref. [43] while the first QCA as we know it nowadays was introduced in [44]. In the following QCAs have been mostly a computer-science object of investigation, especially in the field of Quantum Information, with interesting general results [44–46] lying at the basis of the present thesis.

The idea of reproducing the evolution of a macroscopic system starting from a simple rule of local interaction among its elementary constituents is not completely new. In 1990 Gerard 't Hooft [47] imagined to model the physical evolution at the Planck scale as information processing. However, here the automaton is a classical one, and the aim of the work was to describe a deterministic discrete theory underlying QT. A classical automaton description for the Dirac field is also hidden in the nonstandard analysis formulation of the path-integral of Nakamura [48] where the "infinitesimal" plays the role of the "Planck" length. Then, the word "automaton" first appeared in relation to relativistic field theory in the pioneering work of Bialynicki-Birula [49] where the automaton is a unitary matrix representing an updating rule of classical fields evaluated on a lattice of cells. A discrete description of classical relativistic fields evolution can also be found in the context of *lattice-gas* simulations of Meyer [50], where a notion of "field automaton" first appeared, and Yepez [51]. On the other hand in these cases the automaton is considered as a tool for an approximate description of the underlying continuous dynamics, in the same spirit of the *Lattice Gauge Theory* approach [52]. The dynamics of fields is then evaluated on a discrete lattice in order to make phenomenological predictions in the non perturbative regime of QFT, *e.g.* in quantum *chromodynamics* [53].

The quantum operational scenario and the pivotal role of locality in the QCA approach to quantum fields dynamics have some immediate consequences. The QCA has a precise notion of observables, accommodates localized states and measurements and, being quantum ab initio, it is endowed with a well defined probabilistic interpretation without the need of quantization rules. Moreover the QCA could provide a bridge between the axiomatic principles at the basis of algebraic QFT and the possibility of a discrete elementary description of dynamics with QFT recovered as an emergent approximate theory. It is worth emphasizing the novelty of the discretization technic in the QCA approach. The automaton is not a finite-difference version of the usual fields differential equations, namely it is not given by a finite-difference Hamiltonian or Lagrangian as in Lattice Gauge Theory. The automaton is based on an discrete exactly causal unitary evolution and the Hamiltonian has no longer physical relevance. A first consequence of taking discrete unitary evolutions in place of discretized Hamiltonians is that the automaton framework does not suffer the Fermion-doubling [54], namely the existence of states with non zero momentum corresponding to a minimum of the energy (this was already pointed out by Bialynicki-Birula in Ref. [49]).

Taking the quantum automaton evolution as a potential elementary description of fields dynamics, its discrete causal network is reasonably assumed at a length scale much smaller than the usual scale of particle physics—hypothetically the Planck length. Clearly the QCA theoretical framework cannot enjoy a continuous Lorentzian spacetime and the usual Lorentz covariance, as well as any other continuous symmetry, breaks down at the Planck scale. Within this perspective a major requirement for the automaton proposal is the

consistency with QFT in the *large-scale limit* ("low-energy limit"), where the usual dynamics and symmetries of field theory should be recovered as approximated. A notion of spacetime is also expected to emerge in the large-scale limit based on the analysis of the automaton invariance under change of reference frame and on a kind of Einsteinian synchronization protocol in the new automaton framework. The explorative approach of this work initially considers the automaton dynamics in a fixed reference frame devoting a Chapter to the construction of the emergent spacetime.

If on one side restoring QFT and the covariant spacetime in a quantum informational scenario is interesting by itself, the long term aim of the automaton approach to physical dynamics is to provide an alternative consistent framework for tackling open problems in the physics beyond the *Standard Model*.

In **Chapter 2** we shortly review the framework of operational probabilistic theories of Refs. [7, 10] and comment about the possibility of studying toy models different from the quantum one (this discussion is partially based on the material published in Ref. [55]). We will then connect the operational notions of states, effects and transformations, to the usual statistical model of QT, based on quantum states (density operators), generalized observables (POVMs) and quantum operations. Finally two characteristic traits of QT, the *local discriminability* of states and the *entanglement*, are presented in operational terms.

In **Chapter 3** we present the theoretical models which describe a discrete time local evolution of lattice quantum systems. These models underly the elementary approach to QFT proposed in this presentation. Besides the notion of *automaton* we will introduce the one of *walk*. Following the line tracked in Chapter 2, we will put emphasis on the operational aspects of the automata and walks dynamics, considering the possibility of implementing their discrete evolution via elementary local operations. The definitions and the results in this Chapter are mainly taken from the recent literature in Refs. [44–46, 56–58], and references therein. However, the exposition is the result of an effort in presenting the general features of discrete time lattice systems in a coherent way.

In **Chapter 4** we define the *Quantum Field Cellular Automaton* as the informational theoretical model describing the local discrete evolution of quantum fields. We will then derive the simplest (minimal dimension field) non trivial automaton, denoted Dirac QCA, which is *covariant* with respect to the finite symmetries of the automaton causal network in one space-dimension. Assuming the length distance between sites on the lattice to be "small", hypothetically the Planck length, we show that the usual Dirac equation is recovered in the large-scale limit of the automaton dynamics. The QCA unitary discretization of the Dirac evolution is compared with the Lattice Gauge Theory one showing that the *Fermion doubling*, which is a typical feature of lattice field theories, is avoided in the automaton framework. The presented material

is published in Ref. [59]. In the last Section it is also presented the informational derivation of the Dirac automaton in three space-dimensions recently appeared in [60].

In **Chapter 5**, based on the papers in Refs. [59,61], the one-particle sector of the Dirac QCA derived in Chapter 4 is analyzed. We provide a dispersive differential equation which describes the automaton evolution for smooth states, namely states smoothly peaked around some one-particle momentum eigenstate. The approximated analytical solutions of the dispersive equation are compared with exact simulations. Then we define dynamical quantities as the particle position, momentum and velocity, and evaluate their evolution both in the free case and in the presence of a square potential barrier. We find that the Dirac QCA exhibits typical features of the Dirac quantum field evolution—as the *Zitterbewegung* and the *Klein paradox*.

In **Chapter 6** we consider quantitatively the problem of distinguish experimentally the Dirac automaton from the usual Dirac dynamics. The comparison will be carried out in a quantum informational scenario as an "optimal discrimination between two black boxes". A lower bound for the discrimination probability is given in terms of the parameters of the experiment, namely the number of particles involved, their masses and their momenta. Finally we will discuss possible ways of testing the automaton theory *e.g.* observing the corrections to the wave-packets fly-time. These results are published in Ref. [59].

Chapter 7 presents the first results regarding the emergence of spacetime from the automaton framework. The Dirac QCA predicts a modified dispersion relation and then the breakdown of continuous symmetries, like Lorentz covariance, which must be recovered as approximate symmetries in the large scale-limit. Looking for the transformation that preserve the distorted dispersion relation we find a model of *Deformed Special Relativity* [23, 25, 62] with an invariant energy scale. We show how this model exhibits the feature of *relative locality* recently appeared in [63,64]. A not trivial characteristic of the deformed relativistic model based on the automaton dynamics is the introduction of relative locality in a quantum scenario, which is a quite unusual feature according to [65]. The Chapter concludes showing how a simple synchronization protocol on a classical causal network provides a digital version of the usual Lorentz transformations. The content of the Chapter can be found in Refs. [66,67].

In **Chapter 8** we extend the notion of superselection rule to the general framework of operational probabilistic theories presented in Chapter 2. In this scenario the probabilistic theory of Fermionic systems is a parity superselection of QT where superpositions of states having an even and an odd particle number are forbidden. We explore the consequences of the parity superselection on the informational properties of the model. It turns out that the *Fermionic Quantum Theory* lacks two fundamental traits of QT the *local discriminability* of states and the *monogamy* of the entanglement. Possible computational and

physical implications are discussed at the end of the Chapter. These results are published in Ref. [68].

In **Chapter 9** we conclude the thesis with some final remarks and with future perspectives.

1.1 Notations and Conventions

The beginning of each Chapter is provided with a short introduction that summarizes the aim and the main ideas that are going to be presented.

Some notations and acronyms are extensively used as a standard throughout the whole thesis. For the reader's convenience we report most of them below:

- $\mathcal{H}$ — Hilbert space
- $\mathcal{A}, \mathcal{F}$ — Algebras
- $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ — Pauli matrices
- $\mathrm{A}, \mathrm{B}, \mathrm{C}, \ldots$ — systems of a probabilistic theory
- $\mathsf{St}(\mathrm{A})$ — set of states of the system A
- $\mathsf{Eff}(\mathrm{A})$ — set of Effects of the system A
- $\mathsf{Transf}(\mathrm{A}, \mathrm{B})$ — set of Transformations from the system A to the system B
- $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ — linear span of the set of states $\mathsf{St}(\mathrm{A})$
- $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$ — linear span of the set of effects $\mathsf{Eff}(\mathrm{A})$
- D_{A} — dimension of the system A
- $C_0^{\infty}[a, b]$ — set of smooth functions with compact support in $[a, b]$
- $L_1[a, b]$ — set integrable functions with support $[a, b]$
- $L_2[a, b]$ — set of square integrable functions with support $[a, b]$
- $\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \sum_{x \in \mathbb{Z}} e^{-ikx} f(x),\ k \in [-\pi, \pi]$ — discrete Fourier transform of a band-limited function f
- $U(k)$ — unitary operator in the momentum representation
- H — unphysical Hamiltonian corresponding to the discrete automaton U ($U^t = \exp(-iHt)$)
- $\omega(k)$ — dispersion relation
- $v(k) := \frac{\partial \omega(k)}{\partial k}$ — group velocity
- $D(k) := \frac{\partial^2 \omega(k)}{\partial k^2}$ — diffusion coefficient
- $|\mathsf{S}|$ — cardinality of the set S

- CA — Cellular Automaton
- CN — Causal Network
- DAG — Directed Acyclic Graph
- DSR — Deformed Special Relativity
- FQT — Fermionic Quantum Theory
- LFM — Local Fermionic Mode
- LHS — Left Hand Side
- LOCC — Local operations and classical communication
- MES — Maximally Entangled Set
- POVM — Positive Operator Valued Measure
- QCA — Quantum cellular automaton
- QFT — Quantum Field Theory
- QT — Quantum Theory
- QW — Quantum walk
- RHS — Right Hand Side
- RQT — Real Quantum Theory
- UHECR — Ultra High Energy Cosmic Ray

CHAPTER 2

Introduction to the Quantum Theory operational framework

Quantum Theory (QT) has been recently derived [10] in an operational framework from six axioms regarding the informational features of the theory. In order to achieve this remarkable result the physicists community, and especially the one of Quantum Information and Computer Science, had to move from QT to the more general scenario of *probabilistic theories*. Here it has been possible to determine the consequences of a single axiom on the whole structure of the resulting theory, leading to a set of informational principles which single out the quantum mathematical framework. Beside the QT axiomatization purpose, the operational framework gained the interest of many authors and the analysis of *probabilistic toy models* different from QT is still pursued in the literature. Typically a toy model is obtained considering a subset of the quantum axioms or a weaker version of some of them. Some example of probabilistic theories can be found in [55].

In this Chapter, after introducing the operational framework for probabilistic theories, we enumerate the six axioms leading to the quantum structure and discuss about the possibility of relaxing them. One of the main features of QT is the *entanglement* and in the last Section of the Chapter we will shortly discuss its operational quantification via the so called *entanglement of formation*. The content of this Chapter will be extensively used in the Chapter 8.

2.1 Operational framework and probabilistic theories

In this Section we summarize the framework of operational probabilistic theories usually considered in the literature [3, 7–10]. In particular here we refer to Refs. [7, 10] where the reader can find a detailed analysis of the framework. The operational notions in the following come from the description of a typical physical experiment, involving a physical system, its evolution, and measurements producing outcomes. The probabilistic structure instead comes from the fact that the possible outcomes of a given experiment occur with a certain probability.

2.1.1 Operational theories: connecting tests in a network

The primitive notion of the framework is the notion of *test*. We can regard the test in many different ways, depending on our needs and context. A test can be a man-made apparatus, such as a Stern-Gerlach setup or a beam splitter, or a nature-made "phenomenon", such as a physical interaction between different particles in some spacetime region.

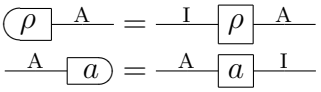
Definition 2.1 (Test) *A test is made of the following ingredients: i) an input system* A*, ii) an output system* B *and iii) a complete collection of classical outcomes* $\{\mathscr{C}_i\}_{i\in\mathsf{X}}$*, labeled by some set* X*. Diagrammatically the test* $\{\mathscr{C}_i\}_{i\in\mathsf{X}}$ *is represented as follows*

(2.1)

and each outcome $i \in \mathsf{X}$*, corresponding to a possible event, is represented as*

. (2.2)

The number of wires at the input and at the output of a device can vary, and one can also have no wire at the input and/or at the output. Among the different kinds of systems, we consider a special one called *trivial system*, denoted by I. A device with input (resp. output) system I is a device with no input (resp. no output). In the circuit it will be represented by no wire, but instead we will draw the corresponding side of the operation box convexly rounded, namely as follows



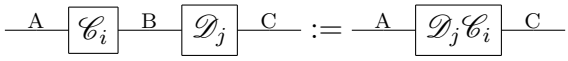
, (2.3)

and in formulas we will write $|\rho_i)_{\mathrm{A}}$ and $(a_j|_{\mathrm{A}}$. The tests corresponding to devices with no input (resp. no output) will be called *preparation-tests* (resp. *observation-tests*).

Given a couple of systems A, B, we denote by $\mathsf{Transf}(\mathrm{A},\mathrm{B})$ the set of all possible events from A to B. We write $\mathsf{Transf}(\mathrm{A})$ in place of $\mathsf{Transf}(\mathrm{A},\mathrm{A})$ if the input and output systems are the same. In particular we reserve a special name for the sets $\mathsf{Transf}(\mathrm{I},\mathrm{A})$ and $\mathsf{Transf}(\mathrm{A},\mathrm{I})$, respectively denoted as $\mathsf{St}(\mathrm{A})$ and $\mathsf{Eff}(\mathrm{A})$. The reason for this special notation is that in the following Section the elements of $\mathsf{Transf}(\mathrm{A},\mathrm{B})$ will be interpreted as the *transformations* from system A to system B while the elements of $\mathsf{St}(\mathrm{A})$ and $\mathsf{Eff}(\mathrm{A})$ will correspond to the *states* and the *effects* of system A.

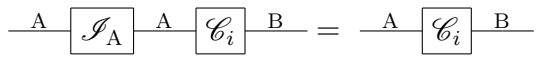
The natural place for a test-event is inside a network of other tests-events. Indeed in an operational framework it is natural to allow the sequential and the parallel compositions of different devices.

Definition 2.2 (Sequential composition) *Given two devices we can compose them in sequence, as long as the input system of the second device is equal to the output system of the first. The events in the composite test are represented as*

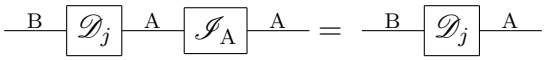


and are written in formula $\mathscr{D}_j\mathscr{C}_i$ as in the RHS *of the equation.*

In particular for every system A one can perform the *identity* test $\{\mathscr{I}_\mathrm{A}\}$ which is a single outcome test with the property

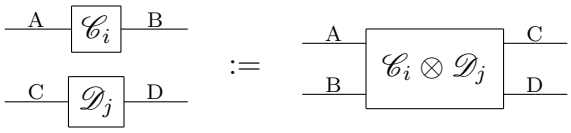


$\forall \mathscr{C}_i \in \mathsf{Transf}(\mathrm{A},\mathrm{B}),$



$\forall \mathscr{D}_j \in \mathsf{Transf}(\mathrm{B},\mathrm{A}).$

Definition 2.3 (Parallel composition) *Given a pair of systems* A *and* B *one can consider a composite system* AB*. Clearly any system* A *can be considered as the same system* A *composed with nothing, namely* $\mathrm{AI} = \mathrm{IA} = \mathrm{A}$ *with* I *the trivial system. Accordingly, two devices can be composed in parallel in a new device with input (output) system given by the composition of the two devices input (output) systems. The events in the composite test are represented as*



and are written in formula $\mathscr{C}_i \otimes \mathscr{D}_j$ as in the RHS *of the equation.*

Taking inspiration by the usual formalism of QT, in the special case of states and effects we will write $|\rho_i)\,|\sigma_j)$ in place of $\rho_i \otimes \sigma_j$ and $(a_i|\,(b_j|$ in place of $a_i \otimes b_j$. Notice that parallel and sequential composition are commuting operations.

As long as the connectivity rules are satisfied, namely i) we can connect only an input wire of a box with an output wire of another box, ii) we can

connect only wires with the same label and iii) loops are forbidden, devices can be organized in networks as in Fig. 2.1. The fact that there are no closed loops gives to the circuit the structure of a *directed acyclic graph* (DAG).

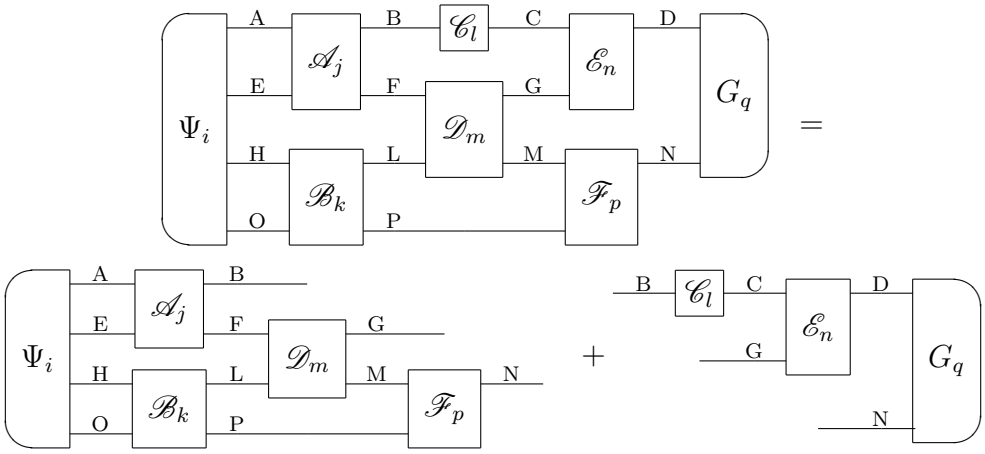


Figure 2.1: **Top figure:** Example of network obtained by composing devices in sequence and in parallel. **Bottom figure:** Splitting of the same network into a preparation and an observation test.

The notions and the simple rules described in this first part of the Chapter, define the operational circuits language and result in a precise definition of operational theory:

Definition 2.4 (Operational theory) *An operational theory is specified by a collection of systems, closed under composition, and by a collection of tests, closed under parallel and sequential composition.*

The main outcome of the operational language, which also was the motivation for its formalization, is that it can describe the schematic of the most general experiment as shown in Fig. 2.2.

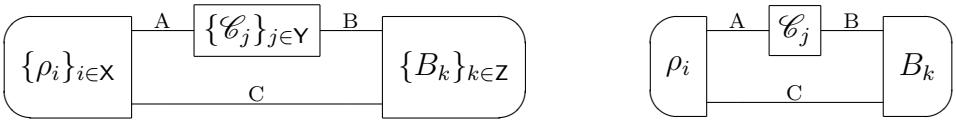
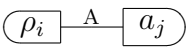


Figure 2.2: Schematic of a general experiment in the operational circuits language. **Left figure:** Schematic of the whole experiment with a preparation-test $\{\rho_i\}_{i\in\mathsf{X}}$, an intermediate test $\{\mathscr{C}_j\}_{j\in\mathsf{Y}}$ on the part A of the experiment system, and an observation-test $\{B_k\}_{k\in\mathsf{Z}}$. **Right figure:** A particular outcome of the experiment corresponding to the preparation-event ρ_i, followed by the event $\mathscr{C}_j$ from system A to system B, which is in turn followed by the observation-event B_k. The same event can be represented in formula $(B_k|_{\mathrm{BC}}\,(\mathscr{C}_j\otimes\mathscr{I}_{\mathrm{C}})\,|\rho_i)_{\mathrm{AC}}$.

2.1.2 Operational probabilistic theories: states, effects, and transformations

We have shown how to connect devices in a network and how this formalism allows to describe the general setting of an experiment. If you want to make predictions about the occurrence probability of events in the experiment, based on your current knowledge, then you need a "theory" that assigns probabilities to different events. Probabilities in the network can be introduced in an easy intuitive way, or in a more axiomatic way as in Ref. [7], assuming that the composition of a preparation-test $\{\rho_i\}_{i\in\mathsf{X}}$ with an observation-test $\{a_j\}_{j\in\mathsf{Y}}$ gives rise to a joint probability distribution $p(i,j)$, namely in the equivalent diagrammatic and formula notation

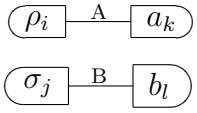


$$=: p(i,j) \qquad (a_j|\,\rho_i) =: p(i,j) \tag{2.4}$$

with

$$p(i,j) \geq 0, \qquad \sum_{i\in\mathsf{X}}\sum_{j\in\mathsf{Y}} p(i,j) = 1.$$

If the experiment consists in the parallel composition of other two experiments, we assume that the joint probability distribution is given by the product:



$$= p(i,k)q(j,l) \qquad p(i,k) := (a_k|\,\rho_i)\,,\ q(j,l) := (b_l|\,\sigma_j)\,. \tag{2.5}$$

Reminding that a preparation (observation) test is a test having as input (output) the trivial system I, namely no input (output) system, we can define an operational probabilistic theory as follows:

Definition 2.5 (Operational probabilistic theory) *An operational theory as in Definition 2.1.1 is probabilistic if to any test from the trivial system* I *to itself is associated a joint probability distribution, and to the composition of two tests from the trivial system to itself is associated the probability distribution given by the product of the two tests probability distributions.*

It is the probabilistic structure of the theory that leads to identify the set of preparation $\mathsf{St}(\mathrm{A})$ and observation $\mathsf{Eff}(\mathrm{A})$ tests for the system A with the states and the effects of the theory for the same system A, following the QT nomenclature. Indeed a preparation-event ρ_i for system A is naturally identified with a function sending observation-events of A to probabilities, namely

$$\rho_i : \mathsf{Eff}(\mathrm{A}) \to [0,1], \quad (a_j| \mapsto (a_j|\,\rho_i)\,, \tag{2.6}$$

and, analogously, observation-events are naturally identified with functions from preparation-events to probabilities

$$a_j : \mathsf{St}(\mathrm{A}) \to [0,1], \quad |\rho_i) \mapsto (a_j|\,\rho_i)\,. \tag{2.7}$$

As probability rule, two observation-events (preparation-events) corresponding to the same function are indistinguishable. We are thus led to the following notions of states and effects:

Definition 2.6 (States and effects) *Equivalence classes of indistinguishable preparation-events for system* A *are called states of* A*, and their set is denoted as* $\mathsf{St}(\mathrm{A})$*. Equivalence classes of indistinguishable observation-events for system* A *are called effects of* A*, and their set is denoted as* $\mathsf{Eff}(\mathrm{A})$*.*

Therefore, in the following we will make the identifications: 1) preparation-events $\equiv$ states; 2) observation-events $\equiv$ effects. The probabilistic structure allows for the following definition of *separability* between sets of states and sets of effects

Definition 2.7 (Separating sets) *We say that a set of effects (states) is separating for a set of states (effects), if any two states of the set have at least a different probability for two effects of the other set.*

Notice that according to our definition of states and effects as equivalence classes we have

$$|\rho_0)_{\mathrm{A}} = |\rho_1)_{\mathrm{A}} \Longleftrightarrow (a|\,\rho_0)_{\mathrm{A}} = (a|\,\rho_1)_{\mathrm{A}} \quad \forall a \in \mathsf{Eff}(\mathrm{A}), \tag{2.8}$$

$$(a_0|_{\mathrm{A}} = (a_1|_{\mathrm{A}} \Longleftrightarrow (a_0|\,\rho)_{\mathrm{A}} = (a_1|\,\rho)_{\mathrm{A}} \quad \forall \rho \in \mathsf{St}(\mathrm{A}), \tag{2.9}$$

namely fixing a system A, the set of states $\mathsf{St}(\mathrm{A})$ is separating for the set of effects $\mathsf{Eff}(\mathrm{A})$ and viceversa the set of effects $\mathsf{Eff}(\mathrm{A})$ is separating for the set of states $\mathsf{St}(\mathrm{A})$.

The set of states $\mathsf{St}(\mathrm{A})$ and effects $\mathsf{Eff}(\mathrm{A})$ are two fundamental elements of a probabilistic theory. On the other hand their "mathematical" structure is in general not trivial; usually they are convex sets but even convexity is not assumed in the most general scenario (see Remark 2.2). Therefore in the operational setting are usually introduced the linear spaces and the convex cones associated to the sets of states and effects. Since states (effects) are functions from effects (states) to probabilities, one can take linear combinations of them. This defines the dual real vector spaces $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ and $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$

$$\mathsf{St}_{\mathbb{R}}(\mathrm{A}) = \mathsf{Eff}_{\mathbb{R}}(\mathrm{A})^*, \qquad \mathsf{Eff}_{\mathbb{R}}(\mathrm{A}) = \mathsf{St}_{\mathbb{R}}(\mathrm{A})^*, \tag{2.10}$$

with V^* denoting the dual set of the set V. Since we will always restrict our attention to finite dimensions by duality one has

$$\dim(\mathsf{St}_{\mathbb{R}}(\mathrm{A})) = \dim(\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})) =: D_{\mathrm{A}}. \tag{2.11}$$

It is immediate to see that a spanning set for $\mathsf{St}(\mathrm{A})$ is separating for $\mathsf{Eff}(\mathrm{A})$ and a spanning set for $\mathsf{Eff}(\mathrm{A})$ is separating for $\mathsf{St}(\mathrm{A})$. Linear combinations of states or effects with positive coefficients define the two convex cones [1] $\mathsf{St}_+(\mathrm{A})$

[1] A set V is convex if for every $v, w \in V$ and for every $p \in [0,1]$ one has $pv + (1-p)w \in V$. Similarly a set V is a cone if for every $v \in V$ and for every $\lambda \geqslant 0$ one has $\lambda v \in V$.

and $\mathsf{Eff}_+(\mathrm{A})$ (see Fig. (2.3) for some examples) which in general are not dual but

$$\mathsf{St}_+(\mathrm{A}) \subset \mathsf{Eff}_+(\mathrm{A})^*, \qquad \mathsf{Eff}_+(\mathrm{A}) \subset \mathsf{St}_+(\mathrm{A})^*. \tag{2.12}$$

The last fundamental object for a probabilistic theory is the set of transformations from a system A to a system B denoted $\mathsf{Transf}(\mathrm{A},\mathrm{B})$. Observing that any event $\mathscr{C}_i$ from the system A to the system B maps a state of A to a state of B

one has that $\mathscr{C}_i$ induces a linear map from $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ to $\mathsf{St}_{\mathbb{R}}(\mathrm{B})$

$$|\rho) \in \mathsf{St}(\mathrm{A}) \mapsto \mathscr{C}_i\,|\rho) \in \mathsf{St}(\mathrm{B}). \tag{2.13}$$

Definition 2.8 (Transformations) *Equivalence classes of indistinguishable events from* A *to* B *are called transformations from* A *to* B.

It is worth mentioning that the transformations of a probabilistic theory are *completely state preserving*, namely given a transformation $\mathscr{C} \in \mathsf{Transf}(\mathrm{A},\mathrm{B})$, for any ancillary system C the state

$$\mathscr{C} \otimes \mathscr{I}_{\mathrm{C}}\,|\rho)_{\mathrm{AC}} \in \mathsf{St}(\mathrm{BD}) \tag{2.14}$$

is an admissible state of the theory. Therefore it is important to notice that two transformations $\mathscr{C}$ and $\mathscr{D}$ from A to B can be different events even if $\mathscr{C}\rho_{\mathrm{A}} = \mathscr{D}\rho_{\mathrm{A}}$ for any state $\rho \in \mathsf{St}(\mathrm{A})$. We say that $\mathscr{C}$ and $\mathscr{D}$ are in the same equivalence class if and only if for any ancillary system C, and for a any joint state ρ_{AC}, it is $(\mathscr{C} \otimes \mathscr{I}_{\mathrm{C}})\rho_{\mathrm{AC}} = (\mathscr{C} \otimes \mathscr{I}_{\mathrm{C}})\rho_{\mathrm{AC}}$. This means that $\mathscr{C}$ and $\mathscr{D}$ act exactly in the same way also when applied locally on any composite system AC.

Remark 2.1 (No restriction hypothesis) Sometimes, as is the case in QT, the probabilistic framework is endowed with the so called *non restriction hypothesis*, namely all completely state preserving transformations are trasformations of the theory. In this case, since also effects are a special kind of transformations $\mathsf{Eff}(\mathrm{A}) = \mathsf{Transf}(\mathrm{A},\mathrm{I})$ with no output systems, a system A is fully specified by its set of states $\mathrm{St}(\mathrm{A})$.

A remarkable feature of the framework is the possibility of equipping the vector spaces $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ and $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$ with an operational norm, whose operational interpretation is related to a scheme for optimal discrimination between two elements in the set (see [69]). From the discrimination protocol it follows that the following

$$\|\omega\| = \sup_{a_0 \in \mathsf{Eff}(\mathrm{A})} (a_0|\omega) - \inf_{a_1 \in \mathsf{Eff}(\mathrm{A})} (a_1|\omega)\,, \tag{2.15}$$

is a norm for $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$, while the norm for $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$ is given by

$$\|b\| = \sup_{\rho \in \mathsf{St}(\mathrm{A})} |\,(b|\,\rho)\,|. \tag{2.16}$$

We will always take the set of states $\mathsf{St}(\mathrm{A})$ to be closed in the operational norm. Operationally, taking $\mathsf{St}(\mathrm{A})$ as a closed set is very natural: if there is a sequence of states $\{\rho_n\}_{n=1}^{\infty}$ that converges to $\rho \in \mathsf{St}_{\mathbb{R}}(\mathrm{A})$ means that there is a procedure to prepare ρ with arbitrary precision and hence ρ should be considered a state of the theory.

Remark 2.2 (Convexity of the set of states and effects) In the literature it is usually assumed that the set of states $\mathsf{St}(\mathrm{A})$ and effects $\mathsf{Eff}(\mathrm{A})$ are convex. In an operational probabilistic theory this is almost obvious since the experimenter can randomize the choice of devices with arbitrary probabilities. On the other hand many general results can be obtained without assuming convexity as in [7, 10] or in the toy models of Spekkens in Ref. [69].

Remark 2.3 (Categorical structure) The notions involved in the operational framework can also be formalized within the *Category Theory* [70]. It is not surprising that in the Category Theory scenario, whose prerogative is to analyse mathematical structures from a very abstract and general point of view, it is possible to find a suitable mathematical framework capturing the fundamental structure of operational theories. In categorical language, an operational theory is a category, where systems and events are respectively *objects* and *arrows*. Every arrow has an input and an output object, and arrows can be sequentially composed. A test is then a collection of arrows labeled by outcomes. The fact that in an operational theory we have a parallel composition of systems, and that such a composition is symmetric (*i.e.* $\mathrm{AB} \approx \mathrm{BA}$), is expressed by saying that we have a *strict symmetric monoidal category*. If the operational theory is probabilistic then the scalars, arrows from the trivial system to itself, of the corresponding category must be probabilities.

2.2 The Quantum Theory case

Here we specify the general notions introduced in the last Section to the QT case. We introduce its sets of states, effects and transformations, and present the six axioms that single out QT from the general framework. Then we shortly comment on the consequences of relaxing the axioms and the possibility of deriving probabilistic theories different from QT.

2.2.1 The states, effects and transformation of Quantum Theory

In the textbook formulation of QT, also known as Copenhagen formulation, systems are associated with complex separable Hilbert spaces $\mathcal{H}$ of dimension $\dim \mathcal{H} < \infty$. To any system A it corresponds a Hilbert space $\mathcal{H}_{\mathrm{A}}$ of a certain dimension $\dim \mathcal{H}_{\mathrm{A}} = d_{\mathrm{A}}$. More precisely in QT it is $\mathcal{H}_{\mathrm{A}} = \mathbb{C}^d$ and

$$D_{\mathrm{A}} = \dim(\mathsf{St}_{\mathbb{R}}(\mathrm{A})) = d_{\mathrm{A}}^2, \tag{2.17}$$

where d_{A} is also the number of perfectly distinguishable states of system A.

The deterministic (usually called normalized) states $\mathsf{St}_1(\mathrm{A})$ of the system A are represented by density matrices on the corresponding Hilbert space $\mathcal{H}_{\mathrm{A}}$, namely trace-class operators $\rho \in \mathsf{T}_1(\mathcal{H}_{\mathrm{A}})$ satisfying

$$\rho \geqslant 0, \qquad \mathrm{Tr}[\rho] = 1, \tag{2.18}$$

where by $\rho \geqslant 0$ we mean that the operator ρ is positive. We know from Section 2.1.2 that the state ρ is the equivalence class of all the preparations of a system that give the same statistics for any possible experiment. A non-deterministic preparation-test $\{\rho_i\}_{i\in\mathsf{X}}$, is a collection of deterministic states and probabilities $\{\tilde{\rho}_i, p_i\}_{i\in\mathsf{X}}$ such that $\rho_i = p_i\tilde{\rho}_i$ and $\sum_{i\in\mathsf{X}} \mathrm{Tr}[\rho_i] = 1$. Accordingly the set of all states $\mathsf{St}(\mathrm{A})$ of system A is the set of all unnormalized density matrices ρ, namely

$$\rho \geqslant 0, \qquad \mathrm{Tr}[\rho] \leqslant 1. \tag{2.19}$$

Practically the states of the system A are positive $d_{\mathrm{A}} \times d_{\mathrm{A}}$ complex matrices with trace bounded by 1. It is immediate to check that $\mathsf{St}(\mathrm{A})$ is convex and the convex combination of different states can be interpreted as the randomized choice between different inequivalent procedures. The extreme points of the convex set $\mathsf{St}(\mathrm{A})$ are the *pure states* of the system while the other states are the *mixed* ones. A quantum state ρ is pure if and only if it is proportional to a rank-one projection while a quantum state is completely mixed if and only if its density matrix has full rank. An example of completely mixed state for the system A is the matrix $\frac{I_{\mathrm{A}}}{d_{\mathrm{A}}}$, with I_{A} the identity operator on A, which is clearly full rank. The pure states are the rank one projectors $\rho = |\phi\rangle\langle\phi|$, and any state ρ can be diagonalized as

$$\rho = \sum_{i=1}^{n} p_i |\phi\rangle\langle\phi|, \qquad p_i \geqslant 0, \quad \sum_i p_i = 1, \tag{2.20}$$

with $\{|\phi_i\rangle \ ; \ i = 1, \ldots, n \leqslant \infty\}$ an orthonormal basis. The diagonalization provides a possible decomposition of the state ρ as convex combination of extremal points of $\mathsf{St}(\mathrm{A})$. However, a mixed state admits many different convex decompositions.

In QT an effect for the system A is represented by a positive operator bounded by the identity on the same system A

$$P \geqslant 0, \qquad P \leqslant I_{\mathrm{A}}. \tag{2.21}$$

The probability resulting from the pairing between a state ρ and an effect P is given by the Born rule

$$(P|\,\rho) = \mathrm{Tr}[P\rho]. \tag{2.22}$$

Again an effect of the system A is simply a $d_{\mathrm{A}} \times d_{\mathrm{A}}$ positive matrix bounded by the identity I_{A}. Accordingly, in quantum theory an observation-test on

the system A is given by a *Positive Operator Valued Measure* (POVM), that is a collection of effects $\{P_j\}_{j\in Y}$ such that $\sum_{j\in Y} = I_{\rm A}$. An effect is *atomic* (extremal) in the set $\mathsf{Eff}(\rm A)$ if and only if it is proportional to a rank-one projector, while a full observation-test is said to be atomic if all the elements of the POVM are rank-one projectors.

We have defined preparation- and observation-tests, whose elements are the states and the effects of the theory. Now we need to specify which are the set of transformations $\mathsf{Transf}(\rm A, B)$ from an input to an output system. A test from A to B is given by a *quantum instrument*, namely by a collection $\{\mathscr{C}_k\}_{k\in Z}$ of completely positive trace non-increasing maps sending linear operators on $\mathcal{H}_{\rm A}$ to linear operators on $\mathcal{H}_{\rm B}$, with the property that

$$\mathscr{C}_Z := \sum_{z\in Z} \mathscr{C}_k, \qquad \mathscr{C}_k : \mathsf{Lin}(\mathcal{H}_{\rm A}) \to \mathsf{Lin}(\mathcal{H}_{\rm B}), \tag{2.23}$$

is trace-preserving. A transformation is then given by a trace non-increasing map, called *quantum operation*, whereas a deterministic transformation is given by a trace-preserving map, called *quantum channel*. Given a quantum operation $\mathscr{C} : \mathsf{Lin}(\mathcal{H}_{\rm A}) \to \mathsf{Lin}(\mathcal{H}_{\rm B})$ it is always possible to represent it in the Kraus form

$$\mathscr{C}(\rho) = \sum_i C_i \rho C_i^\dagger, \qquad C_i : \mathcal{H}_{\rm A} \to \mathcal{H}_{\rm B}, \qquad \rho \in \mathsf{St}(\rm A), \tag{2.24}$$

where the C_i's are the Kraus operators. A map $\mathscr{C}$ is atomic (extremal) if and only if its Kraus form has only one Kraus operator. In particular a reversible transformation is a unitary map whose Kraus form

$$\mathscr{U}(\rho) = U\rho U^\dagger, \qquad U : \mathcal{H}_{\rm A} \to \mathcal{H}_{\rm B}, \tag{2.25}$$

has only one Kraus unitary operator $U^\dagger U = I_{\rm A}$, $UU^\dagger = I_{\rm B}$.

Notice that the operational norm defined in Eq. (2.15) for a general probabilistic theory, in the quantum case is the usual trace-norm $||\cdot||_1$

$$||O||_1 = \mathrm{Tr}[\sqrt{O^\dagger O}]. \tag{2.26}$$

2.2.2 Informational Axioms for Quantum Theory

Here is a summary of the informational principles that, starting from the operational framework of Section 2.1, allow to single out the particular sets of States, Effects and Transformations of QT. The first five axioms are very general and could be included in the definition of the background framework since they define the simplest model of information processing. On the other hand, as we will discuss in the following, it is possible to build up probabilistic theories which do not obey all of these axioms. The sixth principle, denoted *purification*, and introduced in [7], will instead express a genuinely quantum feature and is more difficult to be satisfied in a toy model different from the quantum one.

The six axioms for QT are:

Axiom 2.1 (Causality) *The probability of preparations is independent of the choice of observations. In other words the probability of a measurement outcome at a certain time does not depend on the choice of measurements that will be performed later.*

The main consequence of causality, which is also equivalent to the causality itself, is the uniqueness of the effect e_{A} such that for every observation-test $\{a_j\}_{j\in Y}$ it is

$$e_{\mathrm{A}} = \sum_{j\in Y} a_j. \tag{2.27}$$

The effect e_{A} is denoted *deterministic effect* for system A. In a causal theory the norm of a state $\rho \in \mathsf{St}(\mathrm{A})$ is given by $||\rho|| = (e|\,\rho)$ and one can always define the normalized state

$$\tilde{\rho} = \frac{\rho}{(e|\,\rho)}. \tag{2.28}$$

For this reason, every state in a causal theory is proportional to a normalized state. The set of normalized states $\mathsf{St}_1(\mathrm{A})$, as well as the set of all states $\mathsf{St}(\mathrm{A})$, is closed in the operational norm and, as a consequence of causality, it is always convex: for every pair of normalized states $\rho_1, \rho_2 \in \mathsf{St}_1(\mathrm{A})$, and for every probability $p \in [0,1]$, the convex combination $p\rho_1 + (1-p)\rho_2$ is a normalized state (see also Remark 2.2). Notice that given a bipartite state Ψ of the system AB, the deterministic effects e_{B} and e_{A} allow to evaluate the marginal states on the component systems A and B

Axiom 2.2 (Perfect distinguishability) *Every state that is not completely mixed can be perfectly distinguished from some other state.*

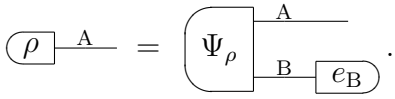
Axiom 2.3 (Ideal compression) *For every state there exists an ideal compression scheme. This means that every source of information can be encoded in a suitable physical system in a lossless and maximally efficient fashion. Lossless means that the information can be decoded without errors. A lossless compression scheme is maximally efficient if the encoding system has the smallest possible size, that is, if the system has no more states than exactly those needed to compress.*

Axiom 2.4 (Local Discriminability or Local Tomography) *Two different bipartite states give different probabilities for at least one product experiment. In other words two bipartite states can always be distinguished using local measurements (see Section 2.3).*

Axiom 2.5 (Pure conditioning) *If a bipartite system* AB *is in a pure state, then each outcome of an atomic measurement on one side—say on the system* A*—induces a pure state on the other system—say on system* B*. Remember that a measurement is said atomic if it cannot be obtained as a coarse-graining of another measurement.*

Axiom 2.6 (Purification) *Every state has a purification. For fixed purifying system, every two purifications of the same state are connected by a reversible transformation on the purifying system.*

Given a state $\rho \in \mathsf{St}_1(\mathrm{A})$, which in general can be a mixed state, a purification of ρ is a pure state Ψ_ρ of some composite system AB, with the property that ρ is the marginal state of Ψ_ρ



The purification postulate states that the ignorance about a part, in this case the ignorance about the system A which is in a mixed state ρ, is always compatible with a maximal knowledge of the whole, in this case the composite system AB. The existence of pure bipartite states with mixed marginal states is a characteristic trait of QT. On the other hand the purification axiom also requires the uniqueness of purification up to reversible transformations which is crucial for generating most of the structure of QT [7, 10].

2.2.3 Relaxing the axioms and probabilistic toy theories

In the literature it is possible to find examples of probabilistic theories different from the quantum one and then satisfying only a subset of the axioms presented in Section 2.2.2. The study of *toy theories* had a strategic role in the informational axiomatization of QT. Indeed it required to move beyond the quantum framework in the world of probabilistic theories which was largely unexplored. Moreover, in that general scenario we had poor intuition biased by our familiarity with QT. Such lack of intuition was partially overcome by the development of alternative probabilistic models for testing the axiomatic framework, and this is the main motivation for toy models.

The most popular example of non quantum probabilistic theory is the *Classical Information Theory* which is much more than a toy model representing the standard of computer technology. The set of states $\mathsf{St}(\mathrm{A})$ of the system A with n perfectly distinguishable states is the simplex having the n perfectly distinguishable states as vertexes. The *bit*, with only two extremal states, is the elementary system of the theory (see Fig. 2.3 where is depicted the case $n = 3$ also called *trit*). As pointed out in [10] the Classical Information Theory satisfies the first five axioms of QT while violates the purification one.

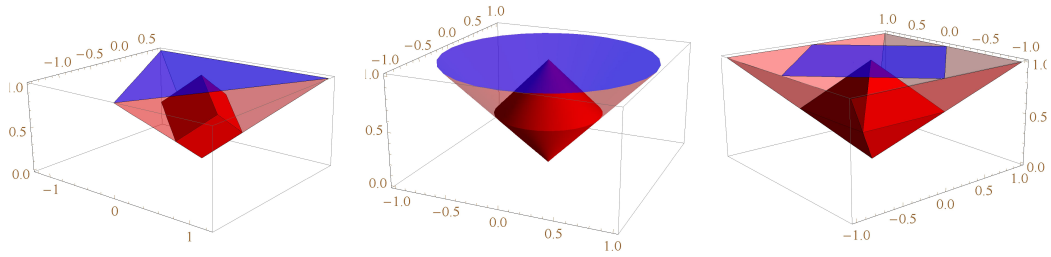


Figure 2.3: Examples of convex cone of states and effects of a probabilistic theories (see Sections 2.2.2 and 2.1.2. Here we consider a system A with dimension $D_{\mathrm{A}} = \dim(\mathsf{St}_{\mathbb{R}}(\mathrm{A})) = \dim(\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})) = 3$. The blue convex at the top represents the set of normalized states $\mathsf{St}_1(\mathrm{A})$. The red transparent cone represents the dual cone of effects $\mathsf{Eff}_+(\mathrm{A})$. The red solid inside the transparent cone represents the convex set of effects $\mathsf{Eff}(\mathrm{A})$ which is the $\mathsf{Eff}_+(\mathrm{A})$ truncation given by the condition $0 \leqslant (a|\rho) \leqslant 1$ for any state $\rho \in \mathsf{St}(\mathrm{A})$ (or $a \leqslant e_{\mathrm{A}}$ where e_{A} is the deterministic effect of the system $_{\mathrm{A}}$). **Left figure:** The trit system of the Classical Information Theory. Notice that set of states is a simplex with three vertexes, corresponding to the three perfectly distinguishable states of the system. **Middle figure:** The rebit, namely the elementary system of RQT. It corresponds to the equatorial section of the qubit, the elementary system of QT. **Right figure:** The elementary system of the Popescu-Rohrlich Boxes Theory.

Another popular probabilistic model is the *Real Quantum Theory* (RQT) [71] given by the restriction of the quantum case to real matrices. The elementary system of RQT, with two perfectly distinguishable states, is denoted *rebit* and its convex set of states (see Fig. 2.3) is the disk obtained by the equatorial section of the *qubit* (the QT elementary system corresponding to the *Bloch sphere*). RQT theory allows the unique purification of its states but violates the local tomography axiom. We will discuss this feature in some details in Chapter 8 where we will derive another probabilistic theory, the *Fermionic Quantum Theory* (FQT), which does not satisfy local tomography.

Among the toy theories we mention the Popescu-Rohrlich *Boxes Theory* [72], which was originally devised as an example of theory that exhibits stronger nonlocality than QT without violating the relativistic *no-signaling*[2]. Here the set of states of the elementary system is a square (see Fig. 2.3). In [55] it has been shown that also this probabilistic toy model violates the purification axiom.

Finally in Refs. [73,74] it is possible to find a concrete example probabilistic theory violating the causality axiom. In this theory the states are quantum operations (see Section 2.2.1), and the transformations are "supermaps" transforming quantum operations into quantum operations. In this case, transforming a state means inserting the corresponding quantum operation in a larger

[2]The Boxes Theory achieves the greatest violation of the Clauser-Horne-Shimony-Holt inequality, a Bell like inequality for quantifying the nonlocality of a probabilistic theory, compatible with the no-signaling principle.

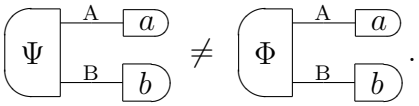
circuit, and the sequence of two such transformations is not a causal sequence.

Finally the toy theories scenario has been useful to separate the notions of causality and determinism. An example of causal non-deterministic theory is QT while in [75] it is introduced a toy theory that is deterministic and non-causal proving that the two notions are totally independent.

2.3 Holism of an operational probabilistic theory

In this Section we focus on the local tomography axiom 2.4 and on the more general notion of *n-local tomography* introduced by Hardy and Wootters in [71]. This generalization will be necessary for the analysis of the FQT in Chapter 8.

According to local discriminability, if two bipartite states are different then there is a chance of distinguishing between them by using only local devices. In general the resulting discrimination will not be optimal but here we are only interested in the possibility of discriminating. Formally we say that a theory enjoyes local discriminability if whenever two states $\Psi, \Phi \in \mathsf{St}(\mathrm{AB})$ are different, there are at least two local effects $a \in \mathsf{Eff}(\mathrm{A})$ and $b \in \mathsf{Eff}(\mathrm{B})$ such that



Local discriminability is equivalent to the local tomography axiom which states that every bipartite state can be reconstructed from the statistics of local measurements on the component systems. This is also equivalent to the fact that every state $\Psi \in \mathsf{St}(\mathrm{AB})$ can be written as

$$\Psi = \sum_{i=1}^{D_{\mathrm{A}}} \sum_{j=1}^{D_{\mathrm{B}}} \Psi_{ij} \rho_i \otimes \sigma_j, \tag{2.29}$$

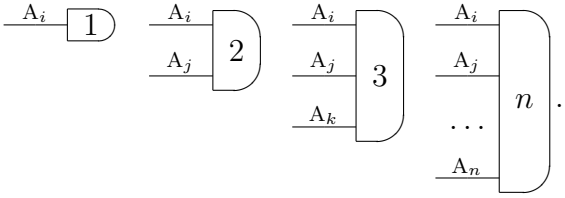
where $\{\rho_i\}_{i=1}^{D_{\mathrm{A}}}$ and $\{\sigma_i\}_{j=1}^{D_{\mathrm{B}}}$ are two basis for the vector spaces $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ and $\mathsf{St}_{\mathbb{R}}(\mathrm{B})$. The same holds for effect and every bipartite effect $B \in \mathsf{Eff}(\mathrm{AB})$ which can be written as

$$B = \sum_{i=1}^{D_{\mathrm{A}}} \sum_{j=1}^{D_{\mathrm{B}}} B_{ij} a_i \otimes b_j, \tag{2.30}$$

with $\{a_i\}_{i=1}^{D_{\mathrm{A}}}$ and $\{b_i\}_{j=1}^{D_{\mathrm{B}}}$ two basis for the vector spaces $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$ and $\mathsf{Eff}_{\mathbb{R}}(\mathrm{B})$. In the following we will use local discriminability and local tomography as synonymous.

Notice that the definition of local tomography can be restated saying that the set of local effects is separating for the set of multipartite states according to Definition 2.7. In a generic probabilistic theory one could need non-local effects for separating multipartite states.

Definition 2.9 (n-local effect) *Given the m-partite system $\mathrm{A}_1 \dots \mathrm{A}_m$, an n-local effect $(a| \in \mathrm{Eff}(\mathrm{A}_1 \dots \mathrm{A}_m)$ with $n \leqslant m$ is an effect that can be written as conic combination of tensor products of effects that are at most n-partite*[3]



Accordingly Hardy and Wootters [71] gave the following definition

Definition 2.10 (n-local tomography) *A theory is n-local-tomographic if and only if the set of n-local effects is separating for multipartite states.*

This is the natural extension of local tomography and simply states that in an n-local-tomographic theory any couple of multipartite states can be discriminated using up to n-local effects. Equivalently a theory is n-locally-tomographic if any multipartite state can be determined from the statistics of 1-local, 2-local,... n-local measurements.

Since an n-local-tomographic theory is also $(n+1)$-local-tomographic, we give the more significant definition

Definition 2.11 (Strict n-local tomography) *A theory is strictly n-local-tomographic if and only if it is n-local-tomographic but not $(n-1)$-local-tomographic.*

In [71] the authors observe that the *holism* of a strictly n-local-tomographic theory increases with n. Following this trend we define

Definition 2.12 (n-holistic theory) *We call a strictly n-local-tomographic theory n-holistic or equivalently we say that it has degree of holism n.*

2.3.1 Local and bilocal tomography

For a generic probabilistic theory we have the constraint $D_{\mathrm{AB}} \geqslant D_{\mathrm{A}} D_{\mathrm{B}}$ on the dimension of a composite system AB since

$$\mathsf{St}_{\mathbb{R}}(\mathrm{A}) \otimes \mathsf{St}_{\mathbb{R}}(\mathrm{B}) \subseteq \mathsf{St}_{\mathbb{R}}(\mathrm{AB}). \tag{2.31}$$

If the theory is local-tomographic, according to Eqs. (2.29) and (2.30), we also have

$$\mathsf{St}_{\mathbb{R}}(\mathrm{AB}) \subseteq \mathsf{St}_{\mathbb{R}}(\mathrm{A}) \otimes \mathsf{St}_{\mathbb{R}}(\mathrm{B}), \tag{2.32}$$

[3]Notice that for example an effect $(p| \otimes (q| \otimes (n| \otimes$, with $p \leqslant q \leqslant n$ is an $(p+q+n)$-partite effect which is still n-local since it is the tensor product of at most n-partite effects.

and for any couple of systems A and B, the dimension of the composite system AB must be

$$D_{\mathrm{AB}} = D_{\mathrm{A}} D_{\mathrm{B}}. \tag{2.33}$$

Examples of local-tomographic theories are QT, Classical Information Theory and the Boxes Theory.

For a strictly bilocal-tomographic theory, the definition implies two constraints on the dimension of bipartite and tripartite systems [71]. A theory is strictly bilocal-tomographic if and only if for any $\mathrm{A}, \mathrm{B}, \mathrm{C}$ we have

$$D_{\mathrm{AB}} > D_{\mathrm{A}} D_{\mathrm{B}}, \tag{2.34}$$

$$D_{\mathrm{ABC}} \leqslant D_{\mathrm{A}} D_{\mathrm{B}} D_{\mathrm{C}} + \widetilde{D}_{\mathrm{AB}} D_{\mathrm{C}} + \widetilde{D}_{\mathrm{BC}} D_{\mathrm{A}} + \widetilde{D}_{\mathrm{CA}} D_{\mathrm{B}}, \tag{2.35}$$

where

$$\widetilde{D}_{\mathrm{AB}} := D_{\mathrm{AB}} - D_{\mathrm{A}} D_{\mathrm{B}}. \tag{2.36}$$

The lower bound (2.34) on of the dimension of an arbitrary bipartite system just means that the theory is strictly bilocal—say it is not local-tomographic. The upper bound (2.35) on the dimension of an arbitrary tripartite system ensures that the theory is not more then bilocal—say we do not need 3-local effects for state tomography and discrimination.

A very popular example of bilocal-tomographic theory is RQT [71]. In Chapter 8 we will show that also the probabilistic with Fermionic modes as elementary systems is bilocal-tomographic. Moreover we will see that both the Real and the Fermionc Theory saturate the upper bound (8.14) and can be derived imposing superselection rules on the quantum states.

2.4 Operational measure of entanglement

The *entanglement* is a quantum mechanical property singled out by Schrödinger many decades ago and it is commonly considered as the characteristic trait of quantum mechanics. For this reason it has been studied extensively in the literature in connection with the others features of QT.

A pure state of a pair of quantum systems is called entangled if it is unfactorizable, as for example for the singlet state of two spin-1/2

$$\tfrac{1}{\sqrt{2}}(|10\rangle - |01\rangle). \tag{2.37}$$

A mixed state is entangled if it cannot be written as a mixture of factorizable pure states, namely if it is not *separable*. A density operator ρ_{AB} of a bipartite system is said to be separable if it can be regarded as an ensemble of product states,

$$\rho_{\mathrm{AB}} = \sum_{j=1}^{n} p_j \rho_{\mathrm{A}}^j \otimes \rho_{\mathrm{B}}^j, \qquad \sum_{j=1}^{n} p_j = 1, \tag{2.38}$$

with a separable pure state corresponding to a product state. Differently from entangled states, separable states can be created by procedures that are local to each subsystem. A first criterion for testing if a state is separable or non-separable was proposed by Peres [76]. A state is separable if taking the matrix transpose of any density operator relative to some orthonormal basis (which is equivalent to take the complex conjugate in that basis) gives another positive semi-definite operator with unit trace which is still a density operator. Similarly, if a state of a bipartite system is separable, taking the partial transpose on system B in any basis gives a good density operator. However, if taking the partial transpose we are left with an operator that is not positive semi-definite then the bipartite state was entangled. This is known in the literature as the *partial transpose condition* for establishing if the density operator of a bipartite systems corresponds to an entangled state. The Peres condition is actually a necessary and sufficient condition only for 2×2 systems (two qubits) or 2×3 systems (a qubit and a qutrit) [77] and in general cannot be considered a general criterion for testing entanglement.

Among the other measures of entanglement considered in the literature we can cite the *entanglement of formation* [78,79] the *distillable entanglement* [80] and the *relative entropy of entanglement* [81]. For a review on the entanglement measures see [82]. In the following we will focus on the entanglement of formation which, having a clear operational interpretation, will be chosen in Chapter 8 for studying the feature of quantum theories with superselection rules.

2.4.1 Entanglement of formation and concurrence

Any measure of entanglement has to consider two main aspects: what is entanglement, and what is entanglement used for. The answer to these questions is not unique and leads to different entanglement measures.

In a general operational theory, as in QT, entanglement must be quantified in operational terms. For bipartite states all measures of entanglement refer to a standard unit—*the ebit*—which is the amount of entanglement of a bipartite singlet state (2.37).

The so called entanglement of formation, which was introduced by Bennett et al. in [83] and by Wootters and Hill in [84], focus on the resources need in order to generate a given amount of entanglement when state manipulation is restricted to Local Operations assisted by Classical Communication (LOCC) (see also Fig. 2.4). Thus the entanglement of formation is the number of ebits that are needed to achieve the state by LOCC. The constraint to LOCC is crucial in elevating entanglement to the status of a resource. Indeed the entanglement of states does not increase under LOCC transformations and entanglement may be defined as the sort of correlations that may not be created by LOCC alone. Roughly speaking LOCC can only use the available entanglement inducing a hierarchy on states based their "usefulness" under LOCC operations. Allowing classical communication menas that LOCC operations are not completely local,

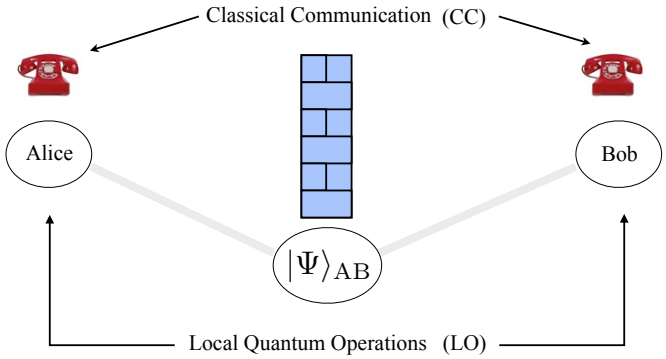


Figure 2.4: Schematic of LOCC operations, namely Alice and Bob who share a bipartite state can only perform local quantum operations (LO) on their subsystems and communicate classical messages (CC).

and can actually have a complicated structure whose complete characterization is still an open problem [4]. We will briefly discuss the operational meaning of the entanglement of formation in the following Section 2.4.2.

The entanglement of formation for an arbitrary state of two qubits was introduced by Wootters and Hill in [84] and later refined by Wootters in [78]. Here we recall the main definitions. Given a state ρ of two qubits A and B, consider all possible pure-state decompositions of ρ, that is

$$\mathcal{D}_\rho \coloneqq \left\{ \{p_i, |\Psi_i\rangle\} \mid \rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i| \right\}. \tag{2.41}$$

The entanglement E for a pure state is defined [86] as the entropy of either one of the two subsystems A and B, namely

$$E(|\Psi\rangle) = -\operatorname{Tr}(\rho_{\mathrm{A}} \log_2 \rho_{\mathrm{A}}) = -\operatorname{Tr}(\rho_{\mathrm{B}} \log_2 \rho_{\mathrm{B}}), \tag{2.42}$$

where $\rho_{\mathrm{A}} = \operatorname{Tr}_{\mathrm{B}}[|\Psi\rangle\langle\Psi|]$ and $\rho_{\mathrm{B}} = \operatorname{Tr}_{\mathrm{A}}[|\Psi\rangle\langle\Psi|]$ are the marginal states of systems A and B. The entanglement of formation is defined for a generally

[4] A complete characterization of LOCC maps is still not present in the literature. This led the community of quantum information to study some larger classes of operations that can be more easily characterized and that retain many features of LOCC maps. The most important class is the set of separable operations. These are the operations that can be written in terms of Kraus operators with a product decomposition:

$$\frac{A_i \otimes B_i \rho_{\mathrm{AB}} A_i^\dagger \otimes B_i^\dagger}{\operatorname{Tr}[A_i \otimes B_i \rho_{\mathrm{AB}} A_i^\dagger \otimes B_i^\dagger]} \tag{2.39}$$

where

$$\sum_i A_i^\dagger A_i \otimes B_i^\dagger B_i = I \otimes I. \tag{2.40}$$

It is easy to see that any LOCC operation can always be written in the form of separable operation. However, it is remarkable that the con- verse is not true as shown in [85], where an example of separable operation is presented that cannot be implemented using LOCC .

mixed state ρ as the convex-roof extension of the pure-state entanglement entropy

$$E(\rho) = \min_{D_\rho} \sum_i p_i E(|\Psi_i\rangle), \tag{2.43}$$

corresponding to the average entanglement of the pure states of the decomposition, minimized over all decompositions. Notice that, since the entropy of a pure state is bounded between 0 and 1, we also have

$$0 \leqslant E(\rho) \leqslant 1, \tag{2.44}$$

with $E(\rho) = 0$ for separable states and $E(\rho) = 1$ for states having maximal entanglement of formation, *e.g.* the singlet state.

In the same Letters [78, 84] the authors provided a simple formula for evaluating the quantity in Eq. (2.43) in terms of the only density matrix ρ. Their formula is based on the notion of *concurrence* which is in turn defined using the *spin flip* transformation. For a pure state of a single qubit the spin flip is given by

$$|\tilde{\psi}\rangle = \sigma_y |\psi\rangle^* \tag{2.45}$$

where $|\psi\rangle^*$ is the complex conjugate of $|\psi\rangle$ in a fixed basis—say the standard one $\{|0\rangle, |1\rangle\}$— and σ_y is expressed in the same basis. For a general state ρ of two qubits the spin flip transformation acts as

$$\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y), \tag{2.46}$$

where the complex conjugate is taken in a fixed basis—say the standard one $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. As pointed out in [87] the spin flip reverses the direction of the spin and it is then intuitive how to exploit it for quantifying the entanglement of a bipartite state. Indeed if a pure state is factorized then the spin flip transforms it into an ortogonal state, while if the state is entangled its spin flipped version can have a non orthogonal component. Following this intuition in [84] the authors showed that the entanglement of a pure state of two qubits can be written as

$$E(|\Psi\rangle) = \mathcal{E}(C(|\Psi\rangle)), \tag{2.47}$$

where $C(\Psi)$ is the concurrence of $|\Psi\rangle$ defined as the overlap between $|\psi\rangle$ and its flipped state

$$C(\Psi) = |\langle\Psi|\tilde{\Psi}\rangle|, \tag{2.48}$$

while $\mathcal{E}$ is the function given by

$$\mathcal{E}(x) = h\left(\frac{1+\sqrt{1-x^2}}{2}\right) \tag{2.49}$$

with h the binary Shannon entropy $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$. Since the function $\mathcal{E}(x)$ is monotonically increasing and ranges from 0 to 1 as x goes from 0 to 1, the concurrence as well as the entanglement of formation can be considered a good measure of entanglement.

As for the entanglement of formation, the concurrence $C(\rho)$ of a mixed state of two qubits is defined as the convex-roof extension of the pure-state concurrence

$$C(\rho) = \min_{D_\rho} \sum_i p_i C(|\Psi_i\rangle). \tag{2.50}$$

Moreover, $C(\rho)$ is connected to the entanglement of formation $E(\rho)$ as in Eq. (2.49) for pure states,

$$E(\rho) = \mathcal{E}(C(\rho)), \tag{2.51}$$

and is given by the following simple formula [78, 84]

$$C(\rho) = \max\{0, \lambda_1 - (\lambda_2 + \lambda_3 + \lambda_4)\}, \tag{2.52}$$

where λ_i, $i = 1, \ldots, 4$ are the eigenvalues of the Hermitian matrix $\sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}}$ in decreasing order.

Here we have introduced the definitions of the entanglement of formation and concurrence for states of two qubits. However, the same definitions hold for bipartite systems of any dimension [79, 88].

2.4.2 The entanglement of formation operational meaning

The operational meaning of the entanglement of formation was proved by Wootters in [79] and extends to mixed states the operational meaning of the entanglement entropy for pure states.

Suppose that two parties, say Alice and Bob, share a quantum state ρ, that in principle can be both separable or entangled, and that they want to create n copies of the given state. Alice and Bob will hold respectively the part A and B of the state and quantum communication between them is not allowed, namely they can only use LOCC operations. We ask, how much entanglement Alice and Bob must share in order to achieve their task? In order to make the question clear we have to fix the elementary resource of entanglement between the systems A and B, which is the "ebit" and corresponds to the singlet state (2.37). The question becomes now how many singlet pairs do Alice and Bob need in order to create n copies of ρ by means of LOCC? If ρ is a pure state $|\Psi\rangle$ the answer is given in terms of the entanglement entropy (2.42) : Alice and Bob need $nE(|\Psi_{\mathrm{AB}}\rangle)$ pairs of singlet states in order to achieve their task [86, 89]. Similarly, if ρ is not pure, Alice and Bob need at least $nE(\rho)$ singlet pairs [79], with $E(\rho)$ the entanglement of formation of Eq. (2.43). In

this sense the entanglement of formation generalizes the operational meaning of the entanglement entropy. More precisely the operational meaning of the entanglement of formation is recovered asymptotically. For any $\epsilon > 0$ there exists a big enough n such that from m singlets the two parties can create n copies of the target state ρ with $m/n \leqslant (1+\epsilon)E(\rho)$.

In order to be a proper measure of entanglement, the entanglement of formation should correspond to the minimal resource needed for generating the amount of entanglement in a given state. It is still not known whether the entanglement of formation is additive [5]

$$E(\rho^{\otimes^N}) = NE(\rho). \tag{2.53}$$

Otherwise, namely if $E(\rho^{\otimes^N}) < NE(\rho)$, it would be possible to use less resources simply using many copies of the state ρ together. In this case a good measure of entanglement would be the regularized entanglement of formation

$$E^{reg}(\rho) = \lim_{N\to\infty} \frac{E(\rho)}{N}. \tag{2.54}$$

2.4.3 Maximally entangled sets

We know that entanglement theory is a "resource theory" in which entanglement is a resource for manipulations restricted to LOCC. Thus, LOCC convertibility induces a natural ordering in the set of entangled states and a fundamental question to ask is which states are maximally entangled, *i.e.* which states are "maximally useful" under the LOCC transformations.

For the bipartite quantum case the singlet state (2.37) is maximally useful, since it allows to reach any other pure state by means of LOCC. Indeed if we consider an arbitrary pure state having Schmidt decomposition

$$|\Phi\rangle = \alpha\,|00\rangle + \beta\,|11\rangle\,, \tag{2.55}$$

we can find a LOCC map that takes the singlet (2.37) to $|\Phi\rangle$ deterministically. It is easy to see that the Kraus operators defined by

$$\begin{aligned} A_0 &\coloneqq (\alpha|0\rangle\langle 0| - \beta|1\rangle\langle 1|) \otimes I \\ A_1 &\coloneqq (\alpha|0\rangle\langle 1| - \beta|1\rangle\langle 0|) \otimes (|0\rangle\langle 1| + |1\rangle\langle 0|) \end{aligned} \tag{2.56}$$

satisfy

$$\sum_{i=0,1} A_i^\dagger A_i = I \otimes I \qquad \sum_{i=0,1} A_i|\Psi\rangle\langle\Psi|A_i^\dagger = |\Phi\rangle\langle\Phi|. \tag{2.57}$$

[5] Clearly $E(\rho^{\otimes^N})$ is less or equal to $NE(\rho)$ since it is always possible to decompose $\rho^{\otimes N}$ into N copies of the optimal decomposition of ρ corresponding to the entanglement of formation $E(\rho)$.

Physically these operations can be realized by LOCC. First we add an ancilla in state $|0\rangle$ to Alice, namely we move from the singlet state to the state

$$\frac{1}{\sqrt{2}}(|00\rangle_{\rm A}|0\rangle_{\rm B} - |01\rangle_{\rm A}|1\rangle_{\rm B}). \tag{2.58}$$

Then we performs the local unitary operations $|00\rangle_{\rm A} \to \alpha|00\rangle + \beta|11\rangle$ and $|01\rangle_{\rm A} \to -\alpha|01\rangle - \beta|10\rangle$ on Alice's systems transforming the state in Eq. (2.58) as follows

$$\frac{1}{\sqrt{2}}\Big[|0\rangle_{\rm A}(\alpha|00\rangle_{\rm AB} + \beta|11\rangle_{\rm AB}) + |1\rangle_{\rm A}(\alpha|01\rangle_{\rm AB} + \beta|10\rangle_{\rm AB})\Big]. \tag{2.59}$$

Now Alice performs a local measurement on the ancilla qubit. If the result is $|0\rangle$ she communicates to Bob that he does not need any further operation, if instead Alice finds $|1\rangle$ she communicates to Bob to perform the map $\sigma_x \cdot \sigma_x$ on his qubit. In both cases the final state is the one in Eq. (2.55). Therefore the singlet state (2.37) is maximally useful under LOCC operations and we say that it is maximally entangled.

On the other hand the notion of maximally entangled state cannot be trivially extended to n-partite states with $n > 2$, where a single maximally useful state does not exist.

In [90] the notion of maximally entangle state has been generalized to the notion of the *Maximally Entangled Set* (MES) of n-partite states.

Definition 2.13 (maximally entangled set) *A* MES *is a set S of states with the following two properties: i) no state in S can be obtained from any other state via* LOCC *(excluding local unitaries) and ii) for any n-partite state $\Psi \notin S$, there exists a state in S such that Ψ can be obtained from it via* LOCC.

According to the definition it is the set of states, not an individual state as in the bipartite case, which is maximally entangled. For example, for the three qubits system, the $|GHZ\rangle$ state and the Werner state $|W\rangle$

$$|GHZ\rangle = |000\rangle + |111\rangle\,, \qquad |W\rangle = |100\rangle + |010\rangle + |001\rangle \tag{2.60}$$

are not connected by LOCC and are both elements of the MES.

The notion of MES can be considered also in probabilistic theories different from the quantum one. As we will see in Chapter 8, for FQT and RQT the notion of maximally entangled state must be superseded by the MES one also in the bipartite case.

We stress that the notion of maximally entangled state (or set) does not coincide with the notion of state with maximal entanglement of formation. A generic a state ρ which has maximal entanglement of formation, *i.e.* $E(\rho) = 1$, may be outside the MES. However, in the bipartite quantum case the two notions are equivalent and a state is maximally entangled if and only if it has maximal entanglement of formation.

2.4.4 Monogamy of entanglement

It has been observed by numerous authors that a quantum system being entangled with another one limits its possible entanglement with a third system: this has been dubbed the "monogamous nature of entanglement".

The sharability of correlations between many parities is one of the main differences between quantum and classical correlations. Indeed in the Classical Information Theory the correlations can be shared among many parties, while QT imposes a severe limitation. For example if a pair of qubits A and B have a perfect quantum correlation, *e.g.* they are in the maximally entangled singlet state (2.37), then neither of them can be entangled to a third system C. A big effort have been devoted to quantify the monogamy of entanglement, see for example [91–98] or [99, 100] for a recent review on the subject.

In the literature one can find many inequalities describing the monogamy of entanglement. Usually such inequalities involve one or more entanglement measures. A class of inequalities has the following form

$$M(\rho_{\mathrm{AB}}) + M(\rho_{\mathrm{AC}}) \leqslant M(\rho_{\mathrm{A(BC)}}), \tag{2.61}$$

where $M(\rho_{\mathrm{AB}})$ measures the entanglement between systems A and B.

It is worth mentioning that not all the entanglement measures satisfy the inequality (2.61) and then not all the entanglement measures are good indicators for monogamy. An entanglement measure satisfying the inequality (2.61) is called *monogamous*. In [95] it has been shown that the concurrence is monogamous and satisfies

$$C^2(\rho_{\mathrm{AB}}) + C^2(\rho_{\mathrm{AC}}) \leqslant 1. \tag{2.62}$$

Notice that if one of the two states, *e.g.* ρ_{AB}, has maximal concurrence $C(\rho_{\mathrm{AB}}) = 1$ (and then also maximal entanglement of formation $E(\rho_{\mathrm{AB}}) = 1$) then the other state ρ_{AC} must have concurrence equal to 0.

In Chapter 8 we will show that FQT is not monogamous. After extending to the Fermionic case the entanglement of formation and the concurrence we will see that the inequality (2.62) can be violated.

CHAPTER 3

Fundamental notions of Quantum Cellular Automata and Quantum Walks

The *Cellular Automata* (CAs) paradigm is simple and appealing and its inherent simplicity belies its potential complexity. The original idea consists in a discrete set of identical cells where each cell can be in a finite set of states. The cells' configuration, namely the state of each cell, is then evolving in discrete steps of time according to a transition rule satisfying certain notions of reversibility, locality and causality. In other words the single cell state at time $t+1$ just depends on the state of the neighbor cells at time t ensuring a finite speed propagation.

The CAs were first introduced by Stanislaw Ulam and John von Neumann in the early 1950s. They realized that CAs were a versatile tool for analyzing many natural phenomena, which is not surprising since most physical processes are themselves local—molecules interact locally with their neighbors, bacteria with their neighbors, ants with theirs and people likewise. Stanislaw Ulam was studying models of crystal growth, while John von Neumann, they were colleagues at Los Alamos, was trying to devise self-replicating systems, *i.e.* a mechanism by which a system produces a copy of itself. Following his intuition, and the suggestions of Ulam based on his work on crystal growth, von Neumann succeeded in engineering a self-replicating machine and published his results in the pioneering work [39].

A big boost to the popularization of the subject came from the John Con-

way's highly addictive Game of Life presented in Martin Gardner's October 1970 column in *Scientific American* [40]. Conway's automaton consists in an infinite two-dimensional grid of square cells, each of which is in one of two possible states, alive or dead. Every cell interacts with its eight neighbors according to an easy evolutionary rule: i) any live cell with fewer than two live neighbors dies, as if caused by under-population ii) any live cell with two or three live neighbors lives on to the next generation iii) live cell with more than three live neighbors dies, as if by overcrowding, iv) any dead cell with exactly three live neighbors becomes a live cell, as if by reproduction. Astonishingly such an elementary model of evolution exhibits a great variety of behaviors by changing the initial pattern, namely the set of alive cells on the grid at the time $t = 0$. In particular for a class of seeds the evolution exhibits global order—say a fixed set of alive cells, a population oscillating periodically between a number of configurations and finally a system of alive cells moving coherently on the grid—from simple local rules that in principle say nothing whatsoever about the global behavior.

Despite their promising features CAs lacked a deep analysis and could not really be called a scientific discipline till the early 1980s, when physicist Stephen Wolfram in a seminal paper [101] entitled *Statistical mechanics of cellular automata* initiated the first serious formalization program.

The CAs features seemed to resemble the properties of a physical evolution but the classical nature of CAs prevented their diffusion as a tool for theoretical physics. The natural way for overtaking this limitation was the extension of the automaton notion to the quantum word as suggested by Richard Feynman in [42]. This resulted in the so called *Quantum Cellular Automata* (QCAs) where the classical systems are superseded by quantum systems in local unitary interaction.

The CAs are not the unique available discrete model of evolution. The *Random Walks*, introduced as a simple description of the one-particle evolution on a discrete lattice, where generalized to quantum systems leading to the notion of *Quantum Walks* (QWs). It is not surprising that QWs and QCAs (or the corresponding classical counterparts) share many features, and many times in the literature the word QCA has been used in place of QW. However, the two notions coincide rigorously only in some specific cases and we will introduce them separately.

The first QCA and QW as we know them nowadays were introduced in Refs. [44] and [102]. In the following literature both QCAs and QWs have been mostly a computer-science object of investigation, especially in the field of Quantum Information where they found the definitive mathematical formalization and a thorough analysis with interesting general results [44–46, 57, 103, 104].

3.1 Generalities on Cellular Automata

In the following we highlight some of the fundamental features of the automata model of dynamics. It is easy to introduce these general properties in the classical context which, from some perspectives, is paradoxically more subtle than the corresponding quantum one. Indeed, as we will see through the Chapter, the unitarity of the quantum evolution simplifies the global structure of the automaton.

Here we provide the general definition of d-dimensional CA that is a causal local evolution of classical systems on a discrete lattice:

Definition 3.1 *A* CA *consists of a d-dimensional cubic lattice $\mathbb{Z}^d$ and a finite set A of symbols,* e.g. $\mathsf{A} = \{0, 1\}$ *in the simplest case, associated to each point x of the lattice. The cardinality $|\mathsf{A}|$ of the set of symbols is denoted the (local) cell dimension. A classical configuration c is a function $c : \mathbb{Z}^d \to \mathsf{A}$ that assigns a symbol of A to each lattice point. The space of all configurations will be denoted by $\mathsf{A}_{\mathbb{Z}^d}$ and we will write $c \in \mathsf{A}_{\mathbb{Z}^d}$ with $c_x = c(x)$ the value of the cell x in configuration c. A cellular automaton is defined by:*

(i) A finite neighborhood scheme, namely a subset $N_x \in \mathbb{Z}^d$ with $N_x = N_y$ for any $x, y \in \mathbb{Z}^d$.

(ii) A local transition rule $T_x : \mathsf{A}_{N_x} \to \mathsf{A}_x$.

The local transition rule, which provides the state of the cell x at the time step $t + 1$ given the state of the neighboring cells N_x at time t, induces a global transition rule

$$T : \mathsf{A}_{\mathbb{Z}^d} \to \mathsf{A}_{\mathbb{Z}^d} \qquad T(c)(x) = T_x(\{c(y); y \in N_x\}) \tag{3.1}$$

which transforms a configuration c by applying the local rule T_x in parallel to all cells.

Two typical neighborhood schemes in $\mathbb{Z}^2$ are depicted in Fig. 3.1. The

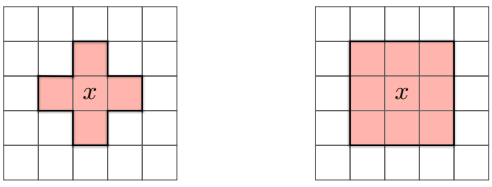


Figure 3.1: Example of neighborhood schemes on a two-dimensional automaton. The state of the cell x at time $t+1$ only depends on the states of the cells in N_x (in red in the figure) at time t according to the local rule T_x of Definition 3.1. On the left the von Neumann scheme, on the right the Moore scheme.

neighborhood scheme could in principle be different from cell to cell ($N_x \neq N_y$) and many results of the automata theory works as well in this scenario. On the

other hand the standard definition takes all cells identical, both in the classical and in the quantum case.

Looking at the definition of CA it should be clear that the most relevant topics in the automata theory are the consequences of their local evolution on the full structure of the automaton. What can we infer from the local rule T_x? We know that it uniquely defines the global action of the automaton via Eq. (3.1) but it gives apparently no constraints on the following two points:

- Is the automaton evolution *reversible* and, if an inverse exists, which is its local structure?
- Does the automaton admit an operational description? Namely does exist a set of local elementary operations such that any time step of the evolution can be represented as a sequence of these basic building blocks? In the automata literature this question usually appears as: is the automaton *locally implementable*?

These issues are at the basis of the automata theory, both in the classical and in the quantum scenario.

3.1.1 Reversibility

Since reversibility is a fundamental property of microscopic physical systems and CAs are devised for simulating such systems, they should obey the same laws, hence be reversible. However, for classical CAs the notion of reversibility is highly non-trivial. Indeed the CA is given as a local rule T_x, while reversibility is a property of the global mapping T. Accordingly in the automaton context we can distinguish between two different notions of reversibility

Definition 3.2 (Reversibility and structural reversibility) *A* CA *is said to be reversible if the global map T is bijective, while it is structurally reversible if it is reversible and the inverse map is still a* CA*, namely it is provided by a local map T_x as in Eq.* (3.1)*.*

Clearly the structural reversibility is the more interesting notion of invertibility for an automaton, whose main feature is the locality of the evolution.

It took a big efforts and many years of investigation for shedding light on the theory of reversible CAs which is well described by the following general questions:

- Does T admit an inverse map?
- Do non trivial reversible CAs exist?
- Ia a CA structurally reversible?
- Which is the neighborhood scheme of the inverse CA and how is it related to the original one?

- Is it possible to infere from the local rule whether the CA is invertible? If this is true, can one efficiently compute the local rule of the inverse CA from the original local rule?

The main results regarding these questions can be found in the Reviews of Toffoli, Margolus and Kari [56, 105, 106] of which here is a punctual summary.

Hedlund [107] and Richardson [108] proved independently that any injective CA is also surjective, and then reversible:

Theorem 3.1 (Hedlund 1969-Richardson 1972) *In any dimension, injectivety, bijectivity and surgectivity are equivalent and then any injective* CA *is reversible.*

On the other hand it is well known that not all reversible CAs are structurally reversible and in general there is no upper bound on the neighborhood scheme of the inverse automaton. A popular counterexample is provided by the *XOR automaton* (see for example [109]) which, despite being bijective, has a non-local inverse map. We will see that for a quantum automaton reversibility and structural reversibility are equivalent notions. As pointed out in [109] this is very peculiar because whilst we are used to think that any reversible computation admits a trivial quantization, this turns out not to be case in the realm of CAs. Indeed the XOR CA cannot be quantized otherwise we should obtain the absurd result of having a non structurally reversible QCA.

The non triviality of reversible CAs was first proved by Toffoli in [110] where it is shown that any d-dimensional CA can be simulated by a $d+1$-dimensional reversible CA. Even more surprisingly, in [111] the authors showed that reversible Turing machines can be simulated by one-dimensional reversible CAs:

Theorem 3.2 (Morita Harao 1989) *One-dimensional reversible* CA*s exist that are computationally universal.*

In 1972 Amoroso and Patt [112] showed that it is possible to decide if an arbitrary one-dimensional CA is reversible looking at the CA local rule:

Theorem 3.3 (Amoroso and Patt 1972) *There exist algorithms based on the automaton local rule to determine if a given one-dimensional* CA *is injective or surjective.*

Unfortunately this result is not generalizable to higher dimension as shown by the Kari *no-go theorem* [113]:

Theorem 3.4 (Kari 1994) *There are no algorithms to determine if a given two-dimensional* CA *is injective or surjective. Therefore there are no algorithms for deciding if a two-dimensional* CA *is invertible or not.*

3.1.2 Operational description: local implementability

The locality of the CA evolution does not imply that the automaton can be realized using local operations on the states of its cells according to the following definition:

Definition 3.3 (Local implementability) *An automaton is locally implementable if its single step evolution $t \to t+1$ (and then the evolution at any time t) can be decomposed in a finite sequence of local operations on non-overlapping groups of cells, also called blocks.*

In the present thesis we propose the quantum automaton as a possible description of microscopic physical dynamics starting from the operational framework of QT. Localizability is certainly an inescapable operational feature since it requires the automaton to be effectively realizable via local operations on the cells' systems. These transformations are respectively classical operations and unitary operations for a classical and a quantum automaton.

A prototypical example of localizable automata are the ones constructed via the Margolous block structure introduced in [114, 115] for the first time.

Definition 3.4 (Margolous block structure) *An automaton has Margolous block structure if the single time step is implemented by a finite number of layers of local transformations. In each layer cells are grouped into non-overlapping blocks and each block undergoes a local transformation. In the same layer the blocks are identical (see Fig. 3.2).*

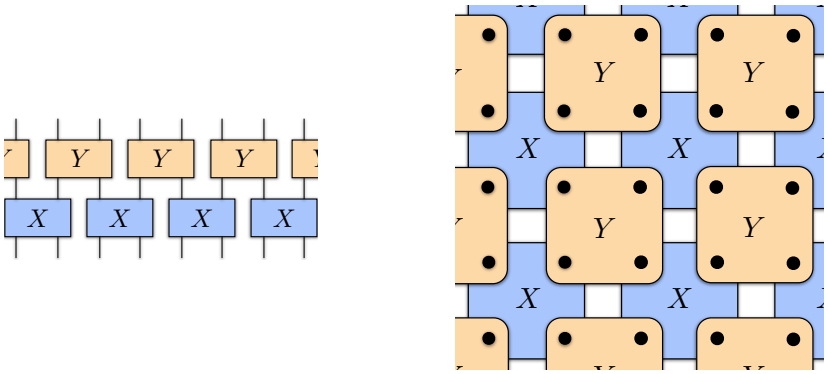


Figure 3.2: Examples of the Margolous block structure of Definition 3.4. In the left and right figures we see the one step evolution of a one- and a two-dimensional automaton decomposed in two layers of local transformations. Notice that the transformations in the same layer are identical.

Clearly an automaton of the Margolous type is invertible if and only if its block transformations are invertible (for example if and only if both X and Y are invertible in the examples of Fig. 3.2) and the bloch structure guarantees the structural reversibility of Definition 3.2. In [116] Kari proved that in one

and two dimensions any CA admits a Margolous block realization up to some shift-like transformations:

Theorem 3.5 (Kari 1996) *All the structurally reversible* CA*s in one and two dimensions admit a Margolous block realization with no more that two, respectively four, layers combined with shift-like transformations.*

Observe that in $d > 1$ two layers are in general not sufficient for locally implement a structurally reversible CA as shown by an explicit counterexample in [117], where a two-dimensional CA has been devised that does not admit a two-layer representation. This counterexample has been used in a quantum scenario [109] for deriving a *no-go theorem* for the two-layer local implementability of a d-dimensional QCA with $d > 1$. A general result on the block representation of a quantum automaton is presented in the next Section.

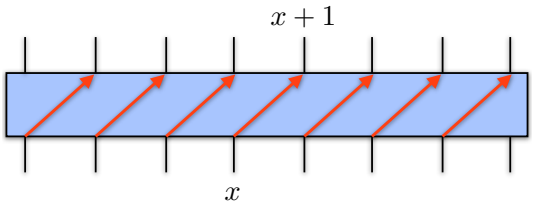


Figure 3.3: The right shift CA. This is an example of structurally reversible (see Definition 3.2) CA which is not localizable (see Definition 3.3). The same is true for the n-sites shifts. The n-sites shifts are also one of the main bricks in the realization Theorem 3.5 of Kari, providing the non-localizable component of a CA.

Remark 3.1 (Shift-like transformations) The Kari Theorem 3.5 can be considered the germ of the *index* theory briefly introduced in Section 3.4. The statement can be interpreted saying that the "non-localizable contribute" in a one- or two-dimensional CA is its shift-like component, while the remaining part is localizable in a Margolous block structure. It is easy to convince yourself that a shift-like transformation, which clearly is a structurally reversible CA, is not localizable and does not admit a Margolous block realization. Consider for example the simple case of the right (or equivalently the left) one-site shift in Fig. 3.3: it is a structurally reversible CA but clearly it cannot be decomposed in transformations acting on a block of cells, and the only way to implement it is to act on all systems in one shot. The index theory makes this intuition rigorous and quantifies the "amount of shift" in the evolution of a CA. As one can imagine the CA is localizable (without combining the Margolous blocks with shift-like transformations) if and only if that "amount is zero".

3.2 Quantum Cellular Automata

We review the definition of QCA whose main novelty with respect to a generic automaton is that cells are quantum systems and their evolution is unitary.

3.2.1 Algebraic Definition

The easiest way to define a QCA is to substitute in Definition 3.1 the cells at each site x with Hilbert spaces of some quantum system A, replacing the set of symbols with the set of possible states $\mathsf{St}(\mathrm{A})$. The difficulty here is that, working with an infinite lattice, the whole system of the automaton should be the infinite tensor product of Hilbert spaces $\bigotimes_x \mathcal{H}_x$ which is in general not a Hilbert space. This formal issue led to the algebraic approach which focus on the quantum observables rather than on states, working in the Heisenberg picture rather than in the Schrödinger one[1]. Thus we consider as usual an infinite d-dimensional cubic lattice $\mathbb{Z}^d$ of cells labelled by integer vectors $x \in \mathbb{Z}^n$. According to the original idea of CA all cells are identical and correspond to a finite dimensional quantum system. Let $\mathcal{H}$, with $\dim \mathcal{H} = n$, be the Hilbert space associated to each cell. The transition to the algebraic formalism is achieved associating to each $x \in \mathbb{Z}^n$ the observable algebra $\mathcal{A}_x$ also denoted cell structure. These algebras are all isomorphic to $\mathcal{M}_n(\mathbb{C})$, the algebra of $n \times n$ complex matrices. If we consider a finite subregion of the lattice $\Gamma \subset \mathbb{Z}^n$, the algebra of observables belonging to all cells in Γ is $\bigotimes_{x\in\Lambda} \mathcal{A}_x$ and will be denoted $\mathcal{A}(\Gamma)$. If $\Gamma_1 \subset \Gamma_2$, $\mathcal{A}(\Gamma_1)$ denotes the subalgebra of $\mathcal{A}(\Gamma_2)$ obtained by tensoring with unit operator in $\Gamma_2 \backslash \Gamma_1$. Notice that tensoring with the identity does not change the norm and we get a local *normed algebra of local observables*, whose completion is the *quasi-local algebra* $\mathcal{A}(\mathbb{Z}^n)$, namely the algebra of observables on all the cells of the QCA. In the following we will use the notation $\Gamma + x = \{y + x;\, y \in \Gamma\}$, or for two subregions $\Gamma_1 + \Gamma_2 = \{x_1 + x_2;\, x_j \in \Gamma_j\}$.

We are ready to provide the algebraic definition of QCA:

Definition 3.5 (Quantum Cellular Automaton) *A* QCA *having neighborhood scheme N_x (with $N = N_0$ and $N_x = x + N$) is a unital homeomorphism $T : \mathcal{A}(\mathbb{Z}^n) \to \mathcal{A}(\mathbb{Z}^n)$ of the quasi-local algebra satisfying the shift-invariance and the locality condition:*

(i) T commutes with all the lattice translations.
(ii) $\forall \Gamma \subset \mathbb{Z}^n$, $T(\mathcal{A}(\Gamma)) \subset \mathcal{A}(\Gamma + N)$.

The local rule T_x of the QCA *T is given by the restriction to the one-site algebra*

$$T_x := T|_{\mathcal{A}_x} : \mathcal{A}_x \to \mathcal{A}(N_x). \tag{3.2}$$

An immediate consequence of the QCA definition is a deep connection between the global T and the local T_x transition rules as stated in the following Lemma [44]:

[1] One could pursue the Hilbert spaces definition taking a restriction on the possible quantum configurations c (see Definition 3.1). Upon defining a quiescent state (a vacuum state), this means that only a finite number of cells can be in a state different from the quiescent one.

Lemma 3.1 (Local and global homeomorphisms) *The global T and the local T_x transition rules of a* QCA *satisfy the two properties:*

(i) The local rule T_x defines uniquely T.

(ii) An homeomorphism $T_x : \mathcal{A}_x \to \mathcal{A}(N_x)$ is a valid local transition rule of a QCA *if and only if $\forall x, y \in \mathbb{Z}^n$ such that $N_x \cap N_y \neq 0$, $T_x(\mathcal{A}_x)$ and $T_y(\mathcal{A}_y)$ commute element-wise.*

Proof. Let us start with point (ii). Consider a subregion of the lattice $\Gamma \in \mathbb{Z}^n$ and the product $\bigotimes_{x\in\Lambda} A_x = \prod_{x\in\Gamma} A_x$, $A_x \in \mathcal{A}_x$, where we have identified the one site algebras $\mathcal{A}_x$ with subalgebras of $\mathcal{A}(\Gamma)$ by tensoring with unit operators. Since T is an homomorphism we have

$$T\left(\bigotimes_{x\in\Lambda} A_x\right) = \prod_{x\in\Lambda} T_x(A_x). \tag{3.3}$$

Moreover T preserves commutation and the $T_x(A_x)$ must commute. Hence the commutativity condition is necessary. We can show that it is also sufficient. Indeed since every local observable can be expressed as a linear combination of tensor products of local-observables, and the products of commuting factors $T_x(A_x)$ is unambiguously defined, Eq. (3.3) defines a homomorphism on the quasi-local algebra. Finally, since we have given an explicit formula of T in terms of T_x, also (i) is proved. ■

It is clear from the Lemma that in order to check if some homomorphism T_x corresponds to a particular QCA we have to verify the commutation between the overlapping algebras under the action of T_x. Thus we just need to verify a finite number of equations. Indeed, according to translation invariance, we can (1) focus on a particular $T_x(\mathcal{A}_x)$, (2) consider all the translations of $T_x(\mathcal{A}_x)$ having a non trivial intersection with $T_x(\mathcal{A}_x)$ and finally (3) verify that $T_x(\mathcal{A}_x)$ commutes with all the selected translated algebras. For all practical purposes it is then possible to restrict from the infinite lattice Z^n to a *finite lattice with periodic boundary conditions* as in Fig. 3.4. Formally, this is equivalent to turning from an infinite lattice $\mathbb{Z}^d$ to a d-dimensional torus $\mathbb{T}^d$. It is worth mentioning that not all restrictions are allowed. Given a particular neighborhood scheme N_x we say that a periodic lattice is regular if it presents the same overlaps between translated algebras as the infinite lattice.

In Section 3.1.1 we highlighted some general questions about the reversibility of an automaton. We mentioned that for a CA the two notions of reversibility and structural reversibility given in Definition 3.2 are not equivalent. Moreover in case of reversibility there are no upper bound on the neighborhood scheme of the inverse automaton. It is striking that for a QCA the issues regarding reversibility are all trivialized [44] due to the unitarity of the evolution:

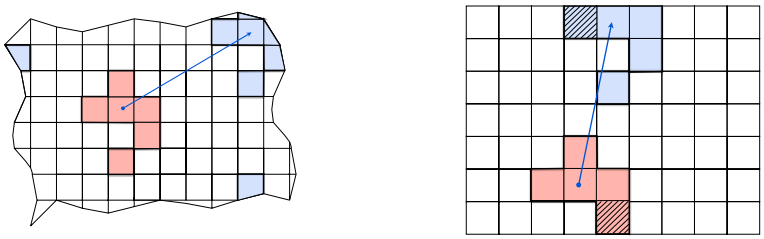


Figure 3.4: Examples of finite lattices with periodic boundary conditions. In the figures we show a particular neighborhood scheme N (in red) and its translated (in blue). The periodic lattice in the left figure is regular (as one can check) while the right one (the 7×9 torus) is not regular since the overlap shown in the figure between two translated copy of N cannot be achieved on the infinite lattice.

Lemma 3.2 (Structural reversibility) *Let T be a* QCA *with neighborhood scheme N_x. Then it is structurally reversible and the inverse* QCA *T^{-1} has neighborhood scheme $N_x^{-1} = x - N_x$.*

Proof. First notice that being T a unital homomorphism T^{-1} exists and is a unital homomorphism. Now consider $A_x \in \mathcal{A}_x$ and $B_y \in \mathcal{A}_y$ with $y \notin x - N_x$. This menas that $x \notin N_y$ and, given $T(B_y) \in \mathcal{A}(N_y)$, one has $[T(B_y), A_x] = 0$. Since T^{-1} is an homeomorphism and preserves commutativity we get $[T^{-1}(T(B_y),T^{-1}(A_x)] = [B_y, T^{-1}(A_x)] = 0$. Observing that B_y is as an arbitrary element of the full algebra $\mathcal{A}_y$ localized in y, $T^{-1}(A_x)$ must be the identity upon this node. The same can be said for any site y outside the set $x - N_x$ which concludes the proof. ■

3.3 Quantum Walks

Thinking about the discrete evolution of physical systems, the most natural example is certainly a particle moving on a lattice. A (classical) *Random Walk* is exactly the description of a particle which moves in discrete time steps and with certain probabilities from one lattice position to the neighboring lattice positions. A quantum version of such a Random Walk, denoted *Quantum Walk* (QW), was first introduced in [102] where measurements of the z-component of a spin-1/2 particle decide whether the particle moves right or left. Later the measurement was substituted by a unitary operator on the spin-1/2 quantum system, also denoted internal degree of freedom or *coin* system, with the QW representing a discrete unitary evolution of a particle state with the internal degree of freedom given by the spin. The formal definition of QW can be found in [57, 118] for the one-dimensional case, and in [119] for graphs of any dimension. See also [120] for a complete review (including walks with continuous time evolution not considered in the present context).

In the most general case the internal degree of freedom at site x can be represented by an Hilbert space $\mathcal{H}_x$ and the total Hilbert space of the system

is

$$\mathcal{H} = \bigoplus_{x \in \mathbb{Z}^d} \mathcal{H}_x \tag{3.4}$$

with the specification of the $\mathcal{H}_x$ representing the cell structure of the QW. Notice that, differently from the QCA case, here the Hilbert spaces are in direct sum rather than in tensor product. Usually QWs are defined in the general non translationally invariant case:

Definition 3.6 (Quantum Walks) *A* QW *with neighborhood scheme N_x is a unitary operator U such that*

$$U \left|\psi\right\rangle \in \bigoplus_{x \in N_x} \mathcal{H}_x, \qquad \forall \left|\psi\right\rangle \in \mathcal{H}_x. \tag{3.5}$$

Denoting by $U_{yx} : \mathcal{H}_x \to \mathcal{H}_y$ the block matrix corresponding to the direct sum in Eq. (3.4), *i.e.*

$$U \bigoplus_{x \in \mathbb{Z}^d} \left|\psi(x)\right\rangle = \bigoplus_{y} \sum_{x \in N_y} U_{yx} \left|\psi(x)\right\rangle, \tag{3.6}$$

for any x only finitely many y give non-zero summands.

In most of the applications it is considered the cell structure given by $\mathcal{H}_x = \mathcal{H}_0 = \mathbb{C}^n$ for all $x \in \mathbb{Z}^d$, as for QCAs. In this case the total Hilbert space satisfies the unitary equivalence

$$\mathcal{H} \cong \mathbb{C}^n \otimes l_2(\mathbb{Z}^d), \tag{3.7}$$

where $l_2(\mathbb{Z}^d)$ is the space of square summable complex functions on the lattice according to the direct sum rather than the tensor product of the one-cell spaces. The state $\left|\psi\right\rangle$ of a particle assigns a vector $(\psi_1(x), \dots \psi_n(x))^T \in \mathbb{C}^n$ to each point x on the lattice with the normalization condition $\sum_x \|\psi(x)\|^2 = \sum_x \sum_{a=1}^n |\psi_a(x)|^2 = 1$. In this scenario one frequently looks at translationally invariant walks, namely unitaries U commuting with the shift operators S_y

$$S_y(\left|\nu\right\rangle \left|x\right\rangle) = \left|\nu\right\rangle \left|x + y\right\rangle, \tag{3.8}$$

for any translation vector y. In case of translational invariance it is possible to study the QW in the momentum space via the Fourier transform

$$\left|\hat{\psi}(k)\right\rangle := \frac{1}{\sqrt{2\pi}} \sum_{x \in \mathbb{Z}} e^{-ikx} \left|\psi(x)\right\rangle, \quad k \in [-\pi, \pi]. \tag{3.9}$$

The spectral decomposition of the walk in the momentum space is given by

$$\hat{U}(k) = \sum_{a=1}^{n} e^{-i\omega_a(k)} \left|a\right\rangle_k \left\langle a\right|_k \tag{3.10}$$

with eigenvalues $e^{-i\omega_a(k)}$ and eigenvectors $|a\rangle_k$. Notice that the ω_j and the $\partial_k\omega_j$ are respectively the *dispersion relation* and the *group velocity* of the QW which are crucial in evaluating the dynamics of a state due to the iterated application of the walk. This analysis, both in the finite and in the asymptotical time regime, has been thoroughly carried out by many authors, see for example [104, 118, 121–124]. In the analysis of the asymptotic behavior it is usually applied the stationary phase approximation (see Appendix B) and in this presentation we will do the same for the Dirac case in Chapter 5.

3.3.1 An example: the Hadamard walk

The most popular QW considered in the literature is the one-dimensional Hadamard walk. The total system Hilbert space is $\mathcal{H} = \mathbb{C}^2 \otimes l_2(\mathbb{Z})$ for which the set $\{|i\rangle\,|x\rangle\,;\, i = r, l,\, x \in \mathbb{Z}\}$ forms an orthonormal basis. The walk is given by

$$U = (C_H \otimes I)S_H, \tag{3.11}$$

where the coin unitary matrix C_H[2] is given by

$$C_H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad \begin{aligned} (C_H \otimes I)\,|x\rangle\,|r\rangle &= \tfrac{1}{\sqrt{2}}\,|x\rangle\,(|r\rangle + |l\rangle), \\ (C_H \otimes I)\,|x\rangle\,|l\rangle &= \tfrac{1}{\sqrt{2}}\,|x\rangle\,(|r\rangle - |l\rangle), \end{aligned} \tag{3.12}$$

while the conditional shift S_H acts as follows on the basis vectors

$$S_H\,|r\rangle\,|x\rangle = |r\rangle\,|x+1\rangle\,, \qquad S_H\,|l\rangle\,|x\rangle = |l\rangle\,|x-1\rangle\,. \tag{3.13}$$

In the momentum space we have

$$\hat{U}(k) = \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} e^{-ik} & 0 \\ 0 & e^{ik} \end{pmatrix}, \tag{3.14}$$

and by diagonalizing this matrix we get the k-dependent eigenvalues $e^{-i\omega_\pm(k)}$ of the walk. For the Hadamard walk the dispersion relation ω ($\omega_\pm = \pm\omega$) and the group velocity $v(k) := \partial_k\omega(k)$ are

$$\omega(k) = \arccos\left(\tfrac{1}{\sqrt{2}}\sin k\right), \qquad v(k) = \frac{\cos k}{\sqrt{1 - \frac{1}{2}\sin^2 k}}, \tag{3.15}$$

(see the plots in Fig. 3.5).

Remark 3.2 (Internal degree of freedom factorization) Notice that the internal space $\mathbb{C}^n$ in Eq. (3.7) can either be considered as a coin (namely the

[2] Clearly the Hadamard walk can be generalized considering some $SU(2)$ coin operation $C_H = \begin{pmatrix} e^{i\alpha}n & e^{i\beta}m \\ -e^{-i\beta}m & e^{-i\alpha}n \end{pmatrix}$ with $n, m \in \mathbb{R}$ and $n^2 + m^2 = 1$.

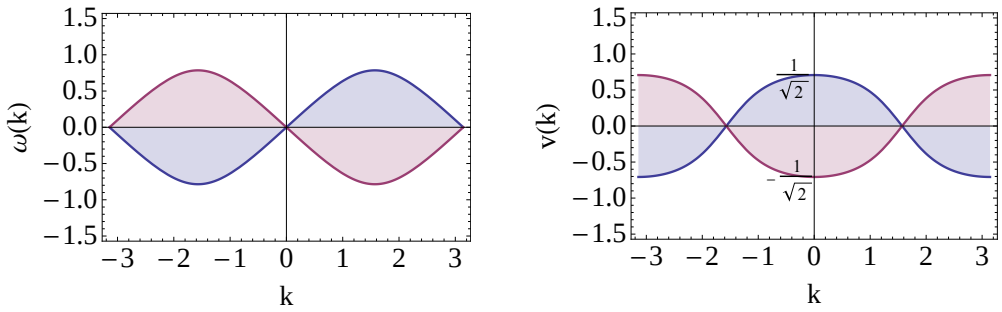


Figure 3.5: The Hadamard walk dispersion relation (left) and group velocity (right). Notice the maximum group velocity for $k = 0, \pm\pi$.

walk is given by alternating coin tosses and shifts, with the last ones depending on the coin) or as a generic internal degree of freedom of a walking particle. In the first case the QW $U(k)$ in the momentum space can be factorized in a k-dependent matrix and a k-independent coin toss C. In the second case we do not have such a factorization. An example of the first case is the above Hadamard walk (see Eq. (3.14) showing the factorization), while for an example of the second kind see the one-particle sector of the Dirac QCA in Section 4.2 (see Eq. (4.38)).

3.3.2 Comparing walks and automata

We mentioned that the term QCA has been used in the literature [43, 49, 125] for a unitary evolution of a particle on a discretized space, where the more appropriate term is QW.

As commented in [44], to see the connection between QCAs and QWs it is easier to consider their classical counterparts, namely Random Walks and CAs: a Random Walk can be seen as a special CA, where each cell can be either empty or occupied, started in a configuration with only one occupied cell. If we put arbitrarily many particles on the lattice, the dynamics is well-defined by the Random Walk as long as the particles do not collide. What is missing in the Random Walk case is the description of collisions which is instead included in the CA dynamics.

Consider now the QCAs and QWs scenario. For a QCA in order to speak about "particles on the lattice", we have to equip the single cell algebra $\mathcal{A}_x$ with a notion of vacuum state. Then we can define the quantity *particle number* [3] which ought to be conserved by the QCA whose dynamics restricted to the "one-particle" Hilbert space is a unitary evolution of the QW type. The one-site algebra can be of any finite dimension n with the internal degree of freedom of the particle represented by the internal Hilbert space of the QW (3.7) (actually we will see in Proposition 4.1 that particles without internal degree of freedom

[3] Notice that the particle number is an infinite sum of 0's and 1's, and the corresponding observable must be defined carefully.

admit only trivial evolutions—say the identity or the shift).

Thus we wonder whether it is possible to consider QWs, with any internal degree of freedom, as the one-particle sector of a QCAs allowing arbitrarily many particles. As pointed out in [44] this can be done in many ways. One possibility is to "second quantize" the QW, considering fields $\psi(x)$ in place of quantum states $|\psi(x)\rangle$ on the sites of the lattice. Regarding fileds as the generators of the local algebra $\mathcal{A}_x$ we have a QCA describing the many-particles evolution whose the one-particle component is the given QW. In this thesis we explore this scenario introducing the *Quantum Field Cellular Automaton*.

3.4 Local implementation for Quantum Cellular Automata and Quantum Walks

We know that an automaton operational description corresponds to locally implement its evolution (see Section 3.1.2). We also know that in general a reversible automaton is non-localizable and that its local implementation could require the aid of shifts which are non-localizable automata (see Theorem 3.5). In the quantum case the following Theorem [45] provides a bridge between the axiomatic algebraic Definition 3.5 and a concrete operational implementation of the QCA. The same results holds in the case of QWs as proved in [46]. The idea is to locally implement the evolution using additional ancillary systems, *i.e.* for any quantum system is introduced an identical copy of it at the same lattice site. Thus the single cell of the system is given by $\mathcal{A}_x \otimes \mathcal{A}_x$ in the automaton case and by $\mathcal{H}_x \oplus \mathcal{H}_x$ in the walk case, with the total system respectively identified with the quasi-local algebra $\mathcal{A}(\mathbb{Z}^d) \otimes \mathcal{A}(\mathbb{Z}^d)$ and with the Hilbert space $\mathcal{H} \oplus \mathcal{H}$. The theorem [45, 46] goes as follows:

Theorem 3.6 (Structure Theorem)

(i) Let T be a QCA *on $\mathcal{A}(\mathbb{Z}^d)$, then the* QCA *$T \otimes T^{-1}$ is locally implementable as the product of the local unitary elements*

$$T_x = (\mathrm{id}_{\mathcal{A}(\mathbb{Z}^d)} \otimes T)(E_x), \qquad T \otimes T^{-1} = \left(\prod_{x\in\mathbb{Z}^d} E_x\right)\left(\prod_{x\in\mathbb{Z}^d} T_x\right), \quad (3.16)$$

where id *is the identical map while $E_x \in \mathcal{A}_x \otimes \mathcal{A}_x$ is the unitary which swaps the two tensor factors.*

(ii) Let U be a QW *on $\mathcal{H}$, then the* QW *$U \oplus U^\dagger$ is locally implementable as the product of the local unitaries*

$$U_x = (U^\dagger \oplus I)E_x(U \oplus I), \qquad U \oplus U^\dagger = \left(\prod_{x\in\mathbb{Z}^d} E_x\right)\left(\prod_{x\in\mathbb{Z}^d} U_x\right), \quad (3.17)$$

where I is the identity operator while E_x is now the unitary operator which swaps the two summands in $\mathcal{H}_x \oplus \mathcal{H}_x$.

Proof. (i) Since the swaps E_x are commuting unitaries and $(\mathrm{id} \otimes T)$ is an homomorphism, also the T_x's in Eq. (3.16) are a family of commuting unitaries localized in $\mathcal{A}_x \otimes \mathcal{A}(N_x)$ (where N_x is the neighborhood scheme of T). Hence their product defines a QCA $\tilde{T}$ on $\mathcal{A}(\mathbb{Z}^d) \otimes \mathcal{A}(\mathbb{Z}^d)$ according to Lemma 3.1. Now for $A_x \in \mathcal{A}_x$ we have

$$\begin{aligned} T_x^\dagger (A_x \otimes I) T_x &= (\mathrm{id}_{\mathcal{A}(\mathbb{Z}^d)} \otimes T)(E_x)(A_x \otimes I)(\mathrm{id}_{\mathcal{A}(\mathbb{Z}^d)} \otimes T)(E_x) \\ &= (\mathrm{id}_{\mathcal{A}(\mathbb{Z}^d)} \otimes T)(E_x(A_x \otimes I)E_x) = (\mathrm{id}_{\mathcal{A}(\mathbb{Z}^d)} \otimes T)(I \otimes A_x) \\ &= I \otimes T(A_x). \end{aligned}$$

In the same way we see that if $A_y \in \mathcal{A}_y$ with $y \neq x$ it is $T_x^\dagger (A_y \otimes I) T_x = A_y \otimes I$ because A_y and E_x commute. Therefore we have $\tilde{T}(A \otimes I) = I \otimes T(A)$ and analogously we can prove that $\tilde{T}(I \otimes A) = T^{-1}(A) \otimes I$, showing that $\tilde{T} = T^{-1} \otimes T$.

(ii) Essentially the same idea works for QWs. The unitaries U_x's in Eq. (3.17) commute, because they are images of the commuting transformations E_x's under the same unitary conjugation, and are localized in a neighborhood of x. Their infinite product defines the unitary of a QW as discussed in Section 3.3 which is

$$\prod_x U_x = (U^\dagger \oplus I) E (U \oplus I) = E(I \oplus U^\dagger)(U \oplus I) = E(U \oplus U^\dagger), \tag{3.18}$$

where $E = \prod_x E_x$ is the global swapping. ■

This Structure Theorem tells us that the commuting unitaries defined in Eq. (3.16) (Eq. (3.17)), followed by a global swap operation between the original system and the ancillary system, implement locally the QCA (the QW) on one system and its inverse on the ancillary system. Thus, upon the introduction of the auxiliary systems, any QCA (QW) can be locally implemented. Notice that in the translationally invariant case the local unitaries in Eq. (3.16) (Eq. (3.17)) are all identical and the local implementation is compatible with the Margolous block structure in Definition 3.4. But how many layers of local blocks do we need? The answer to this question is in the following Corollary [45]:

Corollary 3.1 (n-layer block representation) *Any d-dimensional* QCA T *(*QW U*) admits a 2^d-layer block representation.*

Proof. At each site $x = (i_1, \ldots 1_d)$ the unitary T_x (U_x) of Theorem 3.6 is local to cells $N_x = \{i_1 + N_{x_1}\} \times \ldots \times \{i_d + N_{x_d}\}$. Notice that for an arbitrary neighborhood scheme $N_x = \{i_1 + N_{x_1}\} \times \ldots \times \{i_d + N_{x_d}\}$, it is always possible to group cells in into "supercells" obtaining the so called 1/2-neighborhood scheme $N_x = \{i_1, i_1 + 1\} \times \ldots \times \{i_d, i_d + 1\}$[4]. Thus, after the suitable grouping of cells, whenever $x = (i_1, \ldots 1_d)$ and $y = (j_1, \ldots j_d)$ are such that $|i_k - j_k| > 1$

[4]This grouping is taken very often in the literature on cellular automata. In the $\frac{1}{2}$-scheme to know the state of cell x at time $t+1$, only the states of cells x and $x+1$ at t need be known.

for any k, T_x (U_x) and T_y (U_y) can be performed in parallel. Therefore it is possible to apply simultaneously the $T_x = T_{(i_1,\dots 1_d)}$ ($U_x = U_{(i_1,\dots 1_d)}$) with even i_k's. Now, since any $y \in \mathbb{N}_x$ can be written uniquely as $x + z$ with $z \in \{0,1\}^d$, we need $|\{0,1\}^d| = 2^d$ layers to apply all the local unitaries $T_x = T_{(i_1,\dots 1_d)}$ ($U_x = U_{(i_1,\dots 1_d)}$). Notice that by shift-invariance all the local unitaries are identical accordingly to the Margolous block structure. ■

3.4.1 Index theory for walks and automata

The Structure Theorem 3.6 ensures the operational realizability of any QCA and QW adding ancillary systems. There is still an open question in the local implementability scenario:

- In which cases the discrete evolution can be locally implemented without supplementary systems?

The answer to this question stems in the aforementioned Kari Theorem 3.5 where, even though in the classical framework, he proved the local implementability of one- and two-dimensional automata using the Margolous block structure and some shifts. He realized that the non-localizable component of the evolution is in the shift. For example, while the one-dimensional shift S in Fig. 3.3 cannot be localized, according to Theorem 3.6 the QCA $S \otimes S^{-1}$ (or $S \oplus S^\dagger$ if a QW) can be locally implemented. It seems that the inverse automaton (walk) arises for "symmetrizing the information flow" in the total system.

In [46] this idea has been formalized in the one-dimensional case (the results work also for QCAs and QWs without translational invariance) in the so called *index theory*. When a one-dimensional quantum lattice system is subjected to one step of a reversible discrete-time dynamics, the amont of "quantum information" that enter any block of cells from the left, has to exit that block to the right. For QWs and QCAs this intuition can be made rigorous defining the *index*, a quantity that measures the "net flow of quantum information" through the system.

The index has many striking features whose analysis goes beyond the scope of this Chapter. Here we only present the theorem connecting the localizability of a QCA (QW) to its index. The theorem shows that the QCAs (QWs) with *trivial index* are precisely those which can be locally implemented without ancillary systems. In words, the index is trivial when the net flow of information is zero[5], otherwise it is non-trivial. Prototypes of systems with non-trivial

The name $\frac{1}{2}$-scheme is because the most natural way to represent such an automaton is to shift the cells by $\frac{1}{2}$ at each time step, so that visually the state of a cell depends only on the state of the two cells under it. We stress that this definition is not restrictive since by grouping cells any automaton (walk) with an arbitrary finite neighborhood N can be made into a $\frac{1}{2}$-neighborhood one.

[5]This does not mean that in localizable automata or walks the information cannot move, on the contrary it means that it is not forced to follow a preferred direction but can flow

index are shifts where the flow is clearly not balanced. Intuitively, denoting by S_n the shift on an automaton (walk) with n-dimensional cell algebra $\mathcal{M}_n$ (n-dimensional internal Hilbert space $\dim \mathcal{H}_x = n$), one has

$$\mathrm{ind}(S_n) = n, \tag{3.19}$$

and noticing that $S_n \otimes S_m = S_{nm}$ ($S_n \oplus S_m = S_{n+m}$) we see that the index of an automaton (walk) must be a group of numbers under multiplication (addition). Actually it turns out that for QCAs the index takes values in the positive rationals and the index group is isomorphic to $(\mathbb{Q}_+, \cdot)$, whereas for QWs the index is integer-valued with the index group isomorphic to $(\mathbb{Z}, +)$.

Here is the definition of the index for QCAs [46]:

Definition 3.7 (Index for Quantum Cellular Automata) *Given a one-dimensional* QCA T *with cell structure* $\mathcal{A}_x$*, by regrouping of cells we can assume that it has nearest neighborhood scheme* $N_x = \{x-1, x, x+1\}$*. Now consider any two neighboring cells* $\mathcal{A}_{2x} \otimes \mathcal{A}_{2x+1}$*. Clearly it is*

$$T(A_{2x} \otimes \mathcal{A}_{2x+1}) \subset (A_{2x-1} \otimes \mathcal{A}_{2x}) \otimes (A_{2x+1} \otimes \mathcal{A}_{2x+2}), \tag{3.20}$$

and we can consider the two algebras

$$\mathcal{R}_{2x} = \mathrm{Suppalg}(T(A_{2x} \otimes \mathcal{A}_{2x+1}), (A_{2x-1} \otimes \mathcal{A}_{2x})), \tag{3.21}$$
$$\mathcal{R}_{2x+1} = \mathrm{Suppalg}(T(A_{2x} \otimes \mathcal{A}_{2x+1}), (A_{2x+1} \otimes \mathcal{A}_{2x+2})), \tag{3.22}$$

where $\mathrm{Suppalg}(\mathcal{A}, \mathcal{B})$ *denotes the support algebra of* $\mathcal{A}$ *on* $\mathcal{B}$[6]*. The index of the* QCA *is then defined as*

$$\mathrm{ind}(T) = \frac{r(2x)}{n(2x)} = \frac{n(2x+1)}{r(2x+1)}, \tag{3.23}$$

where $n(y)$ *and* $r(y)$ *are the integer numbers such that* $\mathcal{A}_y \cong \mathcal{M}_{n(y)}$ *and* $\mathcal{R}_{n(y)} \cong \mathcal{M}_{r(y)}(\mathbb{C})$*. The value of the index is independent on the site* x *(even in the non translationally invariant case).*

This definition is very intuitive since the algebras $\mathcal{R}_y$ with even index become larger when information flows to the right, whereas the ones with odd index describe a flow to the left. Assuming identical cells carrying an n-dimensional algebra, for the right shift S_n it is $\mathcal{R}_{2x} = (A_{2x-1} \otimes \mathcal{A}_{2x})$, $\mathcal{R}_{2x+1} = \mathbb{C}I$, corresponding to $r(2x) = n^2$, $r(2x+1) = 1$, and hence $\mathrm{ind}(S_n) = n$, as in Eq. (3.19).

We can now define the index for QWs [46, 126]:

in any direction choosing suitable initial states. We will see this in the details for the Dirac automaton in Chapters 4 and 5.

[6]Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be finite dimensional C^* algebras, and $\mathcal{A} \subset \mathcal{B}_1 \otimes \mathcal{B}_2$. The *support algebra* of $\mathcal{A}$ on $\mathcal{B}_1$, denoted $\mathrm{Suppalg}(\mathcal{A}, \mathcal{B}_1)$, is the smallest C^*-subalgebra $\mathcal{B}_1' \subset \mathcal{B}_1$ such that $\mathcal{A} \subset \mathcal{B}_1' \otimes \mathcal{B}_2$. Such a subalgebra always exists.

Definition 3.8 (Index for Quantum Walks) *For a* QW U *having neighborhood scheme* N_x *the index is defined as*

$$\mathrm{ind}(U) = \sum_{x \geqslant 0, y \in N_x} (\mathrm{Tr}[U_{xy}^\dagger] U_{xy} - \mathrm{Tr}[U_{yx}^\dagger] U_{yx}). \tag{3.24}$$

Starting from this general formula one can compute the index for translationally invariant walks which is connected to the QW *unitary operator* (3.10) *in the momentum representation via*

$$\det U(k) = K e^{-ik\,\mathrm{ind}(U)}, \tag{3.25}$$

for some constant K*.*

From Eq. (3.24) it is easy to see that for the right shift S_1 it is $\mathrm{ind}(S_1) = 1$, according to Eq. (3.19). The corresponding unitary in the momentum space is simply $U(k) = e^{-ik}$ which again gives $\mathrm{ind}(S_1) = 1$ using Eq. (3.25).

Finally we state the theorem connecting the index with the QCAs and QWs local implementability:

Theorem 3.7 (Index Theorem)

(i) A QCA T *can be locally implemented (see Definition 3.1.2) without ancillary systems if and only if* $\mathrm{ind}(T) = 1$*. In this case it can be realized with just two layers of local operatons. If* T *is regrouped in nearest neighbor form, the local operations can be taken to couple pairs of nearest neighbors only.*

(ii) A QW *is locally implementable (see Definition 3.1.2) without ancillary systems if and only if* $\mathrm{ind}(U) = 0$*. In this case it can be written as a product of just two partitioned unitaries, namely it can be realized with just two of layers of local operatons. If* U *is regrouped in nearest neighbor form, then the local operations can be chosen to couple only pairs of nearest neighbors.*

See [46] for the proof.

CHAPTER 4

The Quantum Field Cellular Automaton: the Dirac automaton

The discret time lattice quantum systems introduced in Chapter 3 embody the elementary operational framework for the description of physical dynamics as processing of quantum information. The finite dimension of the quantum systems at the sites of the lattice and the locality of the evolution allow the full computability of the theory. Moreover the computation is homogeneous due to translational invariance and the physical rule is independent on the region of the causal network where it is observed.

After the birth of QCAs some authors considered the possibility of applying the automaton paradigm to the elementary constituents of QFT, *i.e.* to the quantum fields. The word "automaton" first appeared in relation to relativistic field-theory in the pioneering work of Bialynicki-Birula [49], where the automaton describes a discretization of the Weyl and Dirac differential equations. Here the automaton is a unitary matrix representing an updating rule of classical fields evaluated on a lattice of cells. The possibility of using automata for describing the evolution of classical relativistic fields also emerged in the context of *lattice-gas* simulations, especially in the seminal work of Meyer [50], where a notion of "field automaton" first appeared, and in the papers of Yepez [51].

However, in all the above cases the automaton is regarded as an approximation of the underlying continuous dynamics which has to be recovered in the continuum limit. This is also the spirit of the *Lattice Gauge Theory* ap-

proach [52] where the dynamics of fields is approximated on a discrete set of points in order to provide phenomenological predictions in the non perturbative regime of QFT, *e.g.* in quantum *chromodynamics* [53]. These approaches are not focused on the foundations of QFT but on developing new technics for exploiting the impressive predicting power of the theory.

A reconsideration of the mathematical structure of QFT is in the *Algebraic Quantum Field Theory* of Refs. [31, 32] where physical systems are described via nets of algebras, namely via a mapping from finite spacetime regions to the algebras of local observables. Physically the most important notion is the *principle of locality* [33] and the only physical observables are the ones connected with finite regions of spacetime. Global observables—like the total charge or the total energy momentum vector—refer to infinite spacetime regions and lie outside of the picture. Moreover there is a constraint on the observables of the local algebras: the local algebra of observables on a spacetime region Γ must commute with the observables of any algebra on a spacetime region Γ' that is space-like separated from Γ. This principle resembles the Einstein causality and it is the main relativistic ingredient of algebraic QFT.

The QCA approach to the foundations of QFT can provide a bridge between the rigour of the algebraic QFT and the possibility of having a discrete more fundamental description of physical dynamics with QFT emerging at a larger scale. Besides the theoretical interest, the automaton approach provides also a genuinely new way of performing QFT calculations on a lattice as a valid alternative to the existing discretized field models. As we will see the main difference is that in the automaton case it is the unitary evolution to be discretized instead of the Hamiltonian.

In this Chapter, after introducing the notion of *Quantum Field Cellular Automaton*, we derive the simplest one-dimensional field automaton, here denoted Dirac QCA, which is covariant with respect to the symmetries of the causal network. Assuming that the distance between sites on the lattice is small, we show how the automaton recovers precisely the Dirac dynamics in the limit of small masses and momenta (a rigorous comparing between the Dirac automaton and the usual Dirac dynamics will be the subject of Chapter 6). Finally we discuss the difference between the automaton model and the usual discretization of QFT in Lattice Gauge Theory or lattice gasses. A first consequence of the new approach is that the Dirac automaton in one dimension does not suffer the Fermion doubling (see for example [127]), namely the existence of states with momentum different from zero corresponding to a minimum of the energy.

The content of this Chapter is based on Ref. [59]. In the last Section it is presented the informational derivation of the Dirac automaton in three space-dimensions recently appeared in [60].

4.1 The one-dimensional Quantum Field Cellular Automaton

A one-dimensional QCA describes the discrete time local evolution of quantum systems on the one-dimensional lattice $\mathbb{Z}$ (see Fig. 4.1) as extensively discussed in Chapter 3. In our framework, any site $x \in \mathbb{Z}$ corresponds to a (Bosonic or Fermionic) quantum field operator $\psi(x)$ located at the same position. If the field has any internal degree of freedom we define

$$\{\psi_a(x)\}_{a\in\mathsf{A}}, \tag{4.1}$$

with $\mathsf{A} = \{1, \ldots, \Lambda\}$ a finite set, the generators of the field observables local algebra $\mathcal{F}_x$. Clearly the Bosonic (Fermionic) field operators satisfy suitable commutation (anticommutation) rules

$$[\psi_a(x), \psi_b(x)]_\pm = [\psi_a^\dagger(x), \psi_b^\dagger(y)]_\pm = 0, \qquad [\psi_a^\dagger(x), \psi_b(y)]_\pm = \delta_{xy}\delta_{ab}. \tag{4.2}$$

The field automaton implements a unitary local evolution of the fields on the lattice. In the usual QFT the field evolves by a unitary operator V

$$\psi(x, t+1) = V^\dagger \psi(x, t) V, \tag{4.3}$$

where the action of V is linear in the field, namely

$$\psi_a(x, t+1) = \sum_{y\in\mathbb{Z},\, b\in\mathsf{A}} U_{xy}^{ab}\, \psi_b(y, t) \tag{4.4}$$

for some complex coefficients U_{xy}^{ab}. Due to the linearity of the map $V^\dagger \cdot V$ in Eq. (4.3) we can describe the one step evolution of the field by a unitary matrix as follows. Upon introducing the vector field ψ

$$\psi := \begin{pmatrix} \cdots \\ \psi(x) \\ \psi(x+1) \\ \cdots \end{pmatrix}, \quad \psi(x) := \begin{pmatrix} \psi_1(x) \\ \cdots \\ \psi_{|\mathsf{A}|}(x) \end{pmatrix}, \tag{4.5}$$

where each $\psi(x)$ is also a vector with Λ components corresponding to the internal degrees of freedom of the field, we have the equality

$$\psi(t+1) = V^\dagger \psi(t) V = U\psi(t), \tag{4.6}$$

where U is the unitary matrix $UU^\dagger = I = U^\dagger U$ having entries U_{xy}^{ab} according to Eq. (4.4).

The map U in Eq. (4.6) represents the Quantum Field Cellular Automaton whose "cell structure" is naturally identified with the field local algebras $\{\mathcal{F}_x\}_{x\in\mathbb{Z}}$. The locality of the evolution requires that $\psi(x, t+1)$ must be a linear

combination of the field on few neighboring sites $N_x \subset \mathbb{Z}$ at time step t, that is

$$\psi_a(x, t+1) = \sum_{y\in\mathbb{Z},\, b\in\mathsf{A}} U^{ab}_{xy}\, \psi_b(y,t), \qquad U^{ab}_{xy} = 0 \quad \forall y \notin N_x. \tag{4.7}$$

In general the local structure N_x of the automaton can vary from site to site and defines the neighborhood scheme of the automaton.

We can now give a formal definition of a Quantum field Cellular Automaton:

Definition 4.1 (Quantum Field Cellular Automaton) *Consider a field ψ with $\Lambda < \infty$ internal degrees of freedom and denote by $\mathcal{F}_x$ the algebra of local observables generated by the field operators $\{\psi^a_x\}_{a\in\mathsf{A}}$, $\mathsf{A} = \{1, \dots, \Lambda\}$. A one-dimensional Quantum Field Cellular Automaton with cell structure $\mathcal{F}_x$ and neighborhood scheme N_x is a unitary operator U such that*

$$\begin{aligned} U &: \mathcal{F}_x \to \mathcal{F}_{N_x} \\ &\psi_a(x, t+1) \mapsto \sum_{y\in N_x, b\in\mathsf{A}} U^{ab}_{xy} \psi_b(y,t). \end{aligned} \tag{4.8}$$

As we will see in the following, in a physical scenario the translational invariance implies the same structure all over the lattice as in the usual QCA theory (see Definition 3.5). Moreover, without loss of generality, we will consider automata with *nearest neighborhood scheme*, namely $N_x = \{x-1, x, x+1\}$ (see

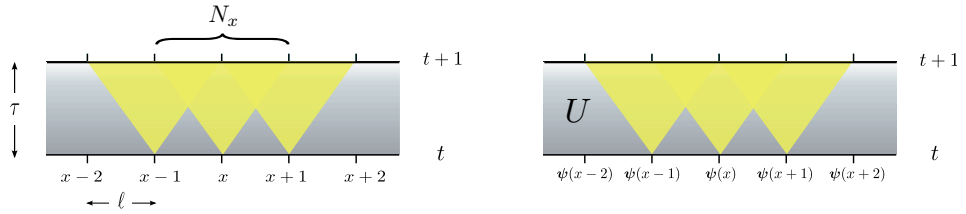


Figure 4.1: **Left figure:** A 1d Cellular Automaton. In the classical case any site corresponds to a finite set of possible classical states while in the quantum case it corresponds to a finite dimensional quantum system. The automaton is a local rule evolving the state of the lattice at time t to the state of the lattice at time $t+1$: the state in position x at time $t+1$ depends only on the states of few neighboring sites N_x at time t. In general the local structure N_x of the automaton can vary from site to site and defines the neighborhood scheme of the automaton. However, in a physical scenario the translation invariance implies the same structure all over the lattice. The example in the figure corresponds to the nearest neighborhood scheme, namely the state at site x at time $t+1$ depends only on the states of the closest neighbors $N_x = \{x-1, x, x+1\}$ at time t. The locality of the automaton suffices to guarantee a bound on the maximal speed of propagation: the causal speed $c = \ell/\tau = 1$ is equal to "one-event-per-step". **Right figure:** Generic 1d Quantum Field Cellular Automaton U. Each site of the lattice corresponds to a quantum field $\psi(x)$ located in the same position.

also Fig. 4.1). The nearest neighborhood interaction is not an assumption by itself, since it is always possible to reduce to such a case by grouping a periodic pattern of the network into a single node of the automaton.

In Section 3.3.2 we pointed out that a QCA restricted to the one-particle sector corresponds to a QW. However, when the evolution is linear as in the quantum field case (4.8), the one-particle dynamics fully specifies the automaton and the two notions of QCA and QW coincide. Accordingly, we can observe that the filed QCA (4.8) corresponds to substitute quantum fields ψ to one-particle states $|\psi\rangle$ in the QW expression (3.6). This corresponds to the "second quantization" of the QW discussed in Section 3.3.2.

Remark 4.1 (Fermionic algebra of local observables) The class of admissible matrices U for a field automaton depends on the field local algebra $\mathcal{F}_x$ according to the condition in Eq. (4.8). At this stage the commutative or anticommutative nature of the field becomes relevant. However all the results regarding the structure of a QCA or QW unitary evolution presented in Chapter 3, hold for the unitary Quantum Field Cellular Automaton. In case of anticommuting fields the local algebra of observables $\mathcal{F}_x$ can be transformed into a commuting quantum algebra $\mathcal{A}_x$ via the Jordan-Wigner isomorphism [128]. Hence the evolution of Fermions on a lattice can be regarded as the evolution of a lattice of qubits, extending to the Fermionic case the QCAs features. In particular the Structure Theorem 3.6 holds in the Fermionic case (see Ref. [129]).

4.1.1 Translational invariance and covariance of the field automaton

According to Eq. (4.7) an automaton U with nearest neighborhood interactions has the band diagonal form

$$U = \begin{pmatrix} \ddots & & & & & & \\ & U_{0-1} & U_{00} & U_{01} & & & \\ & & U_{10} & U_{11} & U_{12} & & \\ & & & & & \ddots & \end{pmatrix}, \tag{4.9}$$

where any U_{xy} is an $\Lambda \times \Lambda$ matrix with Λ the number of the field internal degrees of freedom. From the physical point of view the interesting automata should respect the symmetries of the underlying lattice. In particular they must be *translational invariant* according to the following definition

Definition 4.2 (Translational invariance) *A Quantum Field Cellular Automaton U is translationally invariant if it commutes with the shift operator, namely*

$$[U, S] = 0, \quad S : \psi(x) \to \psi(x+1). \tag{4.10}$$

In this case the structure of the automaton further simplifies as follows

$$U = \begin{pmatrix} \ddots & & & & & \\ & U_{-1} & U_0 & U_1 & & \\ & & U_{-1} & U_0 & U_1 & \\ & & & & & \ddots \end{pmatrix}. \tag{4.11}$$

For convenience, as usual in the literature [57, 103, 104, 130, 131] (see Section 3.3), we will study the dynamics of the field automaton in the momentum representation. For the field operator we have

$$\psi(k) := \frac{1}{\sqrt{2\pi}} \sum_{x\in\mathbb{Z}} e^{-ikx}\psi(x), \quad k \in [-\pi, \pi], \tag{4.12}$$

where with little abuse of notation we utilize the variable name k to denote the Fourier transform of any function of x. Notice that the automaton model is naturally band-limited $k \in [-\pi, \pi]$ and periodic in momenta due to the discreteness of the lattice. Correspondingly, for the unitary matrix U in Eq. (4.6) we have

$$U = \int_{-\pi}^{\pi} \mathrm{d}k\ U(k) \otimes |k\rangle\langle k|. \tag{4.13}$$

We will see that the features of a translationally invariant QCA crucially depend on its momentum eigenvalues and for most of the thesis we will work in the momentum representation.

According to Eq. (4.11) the automaton in momentum space (4.38) assumes the very simple form

$$U(k) = \sum_{x\in\{-1,0,1\}} e^{-ikx} U_x, \tag{4.14}$$

and the unitary operator U can be written as

$$U = \sum_{x\in\{-1,0,1\}} U_x \otimes S_x \qquad S_{-1} = S^\dagger, \quad S_1 = S, \quad S_0 = I. \tag{4.15}$$

In the one-particle sector this is a translationally invariant QW (see Section 3.3) on the Hilbert space

$$\mathbb{C}^\Lambda \otimes l_2(\mathbb{Z}), \tag{4.16}$$

with $\mathbb{C}^\Lambda$ the particle internal Hilbert space.

In general the lattice of an automaton is endowed with a discrete group of symmetries. The one-dimensional lattice $\mathbb{Z}$ only exhibits parity symmetry, corresponding to the lattice reflection with respect to some site (we also have the time reversal symmetry if we consider the causal network given by the automaton evolution). We can ask the automaton to be *covariant* with respect to the lattice symmetries according to the following definition:

Definition 4.3 (Covariance) *Let G be the group of symmetries of the lattice $\mathbb{Z}$. The elements of G are linear functions $g : \mathbb{Z} \to \mathbb{Z}$ that can be represented on the automaton field local algebra $\mathcal{F}_x$ as automorphisms*

$$O_g : \mathcal{F}_x \to \mathcal{F}_{g(x)}. \tag{4.17}$$

We say that a one-dimensional Quantum Field Cellular Automaton U acts covariantly with respect to G if

$$U = \sum_{x\in\{-1,0,1\}} O_g U_x O_g^\dagger \otimes S_{g(x)}, \qquad \text{for any} \quad g \in G. \tag{4.18}$$

A general result of the field automata theory is that all the translational invariant automata with $\Lambda = 1$ are trivial (an automaton is said to be trivial if it is just a shift).

Proposition 4.1 (Triviality of $\Lambda = 1$ fields) *Any one-dimensional translationally invariant Quantum Field Cellular Automaton with $\Lambda = 1$ is trivial.*

Proof. We know that a translationally invariant automaton must be as in (4.14). Since $\Lambda = 1$ the matrices U_{-1}, U_1 and U_0 are just complex numbers and from the unitarity of $U(k)$ one has

$$\begin{aligned} &U_0 = e^{i\theta}, U_{-1} = U_1 = 0, \\ &U_{-1} = e^{i\theta}, U_0 = U_1 = 0, \qquad \theta \in [0, 2\pi], \\ &U_1 = e^{i\theta}, U_0 = U_{-1} = 0, \end{aligned} \tag{4.19}$$

which, modulo a global phase, correspond respectively to the identical (I), the right shift (S), and the left shift ($S^\dagger$) automaton. ■

According to the last Proposition an automaton for the free scalar field is necessarily trivial and, as we will see in Section 4.2, only the identical one is also covariant.

4.1.2 Local implementability of the field automaton

The locality of the field automaton does not ensure the possibility of locally implement it, and a counterexample is provided by the shift field automaton for the same reasons discussed in Sections 3.1.2 and 3.4. Hence local implementability represents a constraint on the field automaton whose unitary time step must always admit a decomposition in a number of elementary gates involving only neighboring systems (see Fig. 4.2).

Theorem 3.6 shows that adding ancillary systems, any d-dimensional field automaton can be locally implemented using 2^d layers of quantum gates. However, in a physical scenario the ancillary systems introduced in the localization deserve an interpretation, *i.e.* they have to correspond to some physical systems.

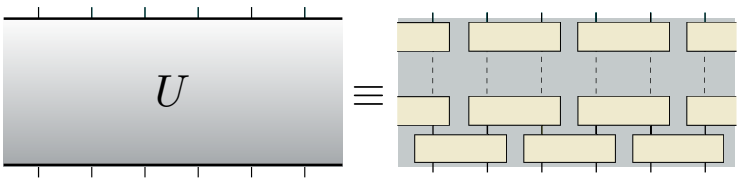


Figure 4.2: Local implementation of the generic one-dimensional Quantum Field Cellular Automaton U. The problem of localizability has been introduced in Section 3.1.2. We know that any one-dimensional automaton can be localized using only two layers of identical gates as in the Margolous block structure of Definition 3.4.

The index theory presented in Section 3.4 gives necessary and sufficient conditions for deciding the field automaton localizability without invoking ancillary systems. Since the one-particle sector of a translationally invariant field automaton U coincides with a translationally invariant QW (see Eq. (4.16)), the index Theorem 3.7 gives the prescription

$$U \quad \text{locally implementable} \quad \Leftrightarrow \quad \det(U(k)) = K, \tag{4.20}$$

for some constant K, where $U(k)$ is the field automaton in the momentum representation (4.13). Moreover, in the nearest neighborhood scheme considered here, any localizable one-dimensional field automaton can be realized with two layers of local operations coupling only pairs of nearest neighbors (thus only two layers of gates in Fig. 4.2 are needed).

The covariant Dirac automaton derived in the following will automatically satisfy (4.20) and no ancillary systems are needed for its operational realization.

4.2 Derivation of the one-dimensional Dirac automaton

Consider a field ψ having Λ internal degrees of freedom. According to Eq. (4.15) the most general 1d translationally invariant QCA with nearest neighborhood interaction is given by

$$U = R \otimes S + L \otimes S^\dagger + M \otimes I, \tag{4.21}$$

where S is the shift operator $S\psi(x) := \psi(x+1)$, and $R = U_1, L = U_{-1}$ and $M = U_0$ are $\Lambda \times \Lambda$ matrices which do not depend on the site. The unitarity of U implies

$$\begin{aligned} &RR^\dagger + LL^\dagger + MM^\dagger = I\,, \\ &MR^\dagger + LM^\dagger = 0\,, \qquad LR^\dagger = 0\,. \end{aligned} \tag{4.22}$$

We now derive the Dirac automaton as the minimal dimension field QCA satisfying the following requirements

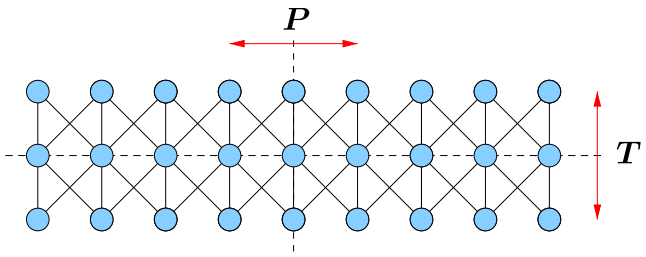


Figure 4.3: Schematic of the three time steps causal network corresponding to a one-dimensional automaton with nearest neighborhood interaction. The topology of the network is left invariant by the mappings $t \mapsto -t$ and $x \mapsto -x$ and the dynamics of the automaton is assumed to be time reversal (T) and parity (P) invariant (see Eqs. (4.23) and (4.24)).

(i) Unitarity of the evolution.
(ii) Locality of the evolution.
(iii) Homogeneity of the interaction topology.
(iv) Covariance under time-reversal $t \mapsto -t$.
(v) Covariance under parity $x \mapsto -x$.
(vi) Minimal dimension for a non-trivial evolution.

The first three assumptions are already contained in the definition itself of one-dimensional translationally invariant QCA with nearest neighborhood scheme. Assumptions (iv) and (v) correspond to the covariance of the automaton with respect to the symmetries of the causal network (in Fig. (4.3)). The more constraining assumption is thus the minimality one (vi) which consists in the minimal number of field internal degrees of freedom Λ for a non trivial evolution.

According to Definition 4.3 assumptions (iv) and (v) can be translated in the following equations

$$(T \otimes I)U(T^\dagger \otimes I) = U^\dagger, \tag{4.23}$$
$$(P \otimes I)U(P^\dagger \otimes I) = R \otimes S^\dagger + L \otimes S + M \otimes I, \tag{4.24}$$

where T is an anti-unitary representation of the operator associated to the time reversal transformation, and P a unitary representation of the operator associated to the lattice reflection. Notice that $(P \otimes I)U(P^\dagger \otimes I)$ is the same as U with S exchanged with $S^\dagger$ according to the reflection of the lattice $\mathbb{Z}$. For $\Lambda = 1$ the only translational invariant QCA satisfying parity invariance is the identical one $U = I$ which is trivial. Next, we have the case $\Lambda = 2$. Eq. (4.24) shows that R and L are unitarily equivalent and from Eq (4.22) it follows that they are both rank one. Thus we can choose the basis where

$$R = \begin{pmatrix} a_1 & a_2 \\ 0 & 0 \end{pmatrix}, \quad \psi(x) := \begin{pmatrix} \psi_r(x) \\ \psi_l(x) \end{pmatrix}, \tag{4.25}$$

naming the two components of the field ψ_r and ψ_l *right* and *left* modes. We now require P and T to be represented in such basis as

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad T = \Xi \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{4.26}$$

where Ξ is the anti-unitary operator denoting complex conjugation in the representation (4.25). With this choice of basis the parity invariance (4.24), which implies

$$PRP^\dagger = L, \qquad PMP^\dagger = M, \tag{4.27}$$

gives

$$R = \begin{pmatrix} a_1 & a_2 \\ 0 & 0 \end{pmatrix}, \quad L = \begin{pmatrix} 0 & 0 \\ a_2 & a_1 \end{pmatrix}, \quad M = \begin{pmatrix} a_3 & a_4 \\ a_4 & a_3 \end{pmatrix}, \tag{4.28}$$

with $a_i \in \mathbb{C}$ for $i = 1, \ldots, 4$. From the time-reversal invariance (4.23), that is

$$TRT^\dagger = L^\dagger, \qquad TMT^\dagger = M^\dagger, \tag{4.29}$$

it follows $a_2 = a_3 = 0$. Finally, using the unitarity of U (4.22) we get

$$|a_1|^2 + |a_4|^2 = 1 \qquad \Re(a_1 \bar{a}_4) = 0, \tag{4.30}$$

which, up to a global phase, gives the unique automaton

$$U = \begin{pmatrix} nS & -im \\ -im & nS^\dagger \end{pmatrix}, \quad n^2 + m^2 = 1\,. \tag{4.31}$$

The constants n and m in the last equation can be chosen positive (the relative sign corresponds to take the left-mode redefined with a minus sign). The unitarity constraint $n^2 + m^2 = 1$ in Eq. (4.31) forces the parameter m to be bounded between 0 and 1

$$n^2 + m^2 \Rightarrow m \in [0, 1]. \tag{4.32}$$

As we will see in the following, m plays the role of an adimensional inertial mass, and the inertial mass of the field is bounded from above [36] [1].

The automaton for $m = 0$ corresponds to the Weyl equation, and will be refereed as Weyl automaton

$$U_W = \begin{pmatrix} S & 0 \\ 0 & S^\dagger \end{pmatrix}. \tag{4.33}$$

The Dirac QCA in Eq. (4.31) introduces a coupling between the two components of the field slowing down the corresponding propagation. Physically the coupling corresponds to the rest frame mass of the field itself.

[1] If we assume that the small scale of the automaton is the Planck one, for the digital-analog conversion from the automaton to the usual dimensional Dirac equation we take the Planck length ℓ_P and the Planck time τ_P as conversion factors for space and time respectively, whereas the mass conversion factor is taken equal to the Planck mass m_P. The maximal speed of local-state propagation in the dimensional case is then $c = \ell_P/\tau_P$, corresponding to the speed of light. ℓ_P, τ_P, and m_P define the measure for dimensions $[L]$ $[T]$ $[M]$, and from them one can derive the other constants, e.g. $\hbar = m_P \ell_P c$.

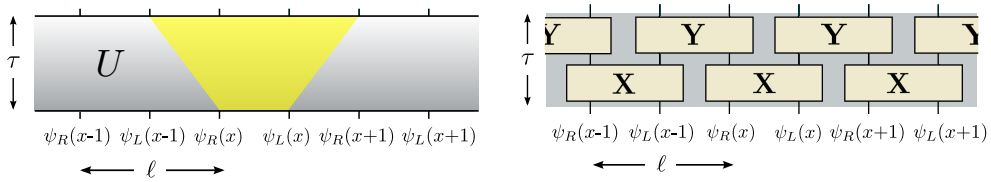


Figure 4.4: Local implementation of the Dirac automaton U in Eq. (4.31) with $X = -i\sigma_1$, $Y = mI + in\sigma_1$. The field operator at site x interacts with the field $\psi(x \pm 1)$ at neighboring sites. In the case of the Dirac automaton the field operator has two components, the right ψ_r and the left ψ_l mode of the field.

Notice that in the derivation of the automaton (4.31) the parity symmetry (assumption (v)) implies a minimal internal dimension $\Lambda = 2$ for a non identical evolution. This means that it is not possible to consider an automaton having just an internal degree of freedom–say a scalar field. As a byproduct of the parity assumption we also have the localizability of the automaton. Indeed in the Dirac case the condition (4.20) is trivially satisfied as one can easily check using the automaton in the momentum space (4.38)

$$\det\left(U(k)\right) = n^2 + m^2 = 1. \tag{4.34}$$

Moreover, as a consequence of the Index Theorem 3.7 any one-dimensional localizable field automaton can be locally implemented using two layers of quantum gates. For the Dirac automaton (4.31) we have the local implementation shown in Fig. 4.4 and, as one can easily check, X and Y result to be

$$X = -i\sigma_1, \qquad Y = mI + in\sigma_1. \tag{4.35}$$

Clearly also the Weyl automaton is localizable since it is just a particular case of the Dirac one with $m = 0$. In the Weyl case (4.33) the local gates X and Y in Fig. 4.4 are simply

$$X_W = Y_W = \sigma_1, \tag{4.36}$$

which correspond to swap the left and the right modes of the field. Thus if we prepare a state localized in x and in one of the two modes, namely $|x\rangle \otimes (1, 0)$ or $|x\rangle \otimes (0, 1)$, it moves at the causal speed in the in the right or left direction.

In Figs. 4.5 and 4.6 we give computer evaluations of the Dirac automaton evolution. These behaviors will be derived analytically in Chapter 5. Fig. 4.5 shows the evolution for localized states which are not described in QFT. We see that for small m the state splits in its left and right components which evolve at speed close to the causal one $c = 1$. For higher m the state spreads in region narrower than the causal cone. Fig. 4.5 shows two evolutions for smooth Gaussian states. We see that the states remain localized in a finite region evolving with a certain drift and diffusion coefficient. In the second figure we also see the *Zitterbewegung* oscillation of the mean position.

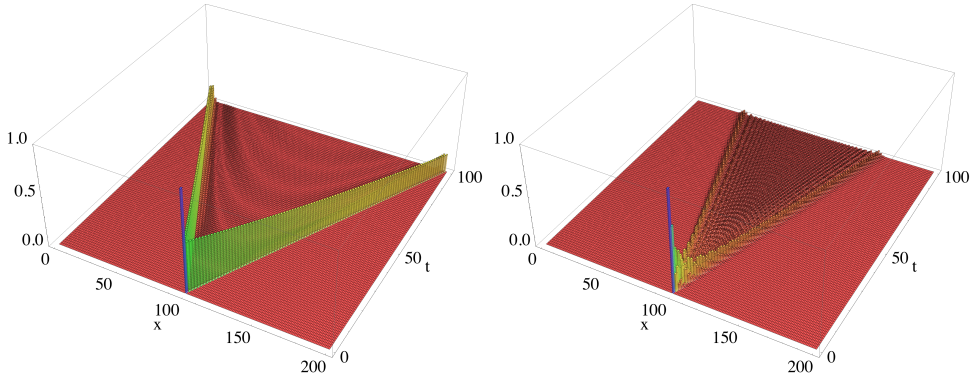


Figure 4.5: The automaton (4.31) evolution for localized states. Automaton ($m = 0.1$ on the left and $m = 0.8$ on the right) evolution for the localized sate $|100\rangle \otimes \frac{1}{\sqrt{2}}(1,1)$ showing the spread in the causal cone.

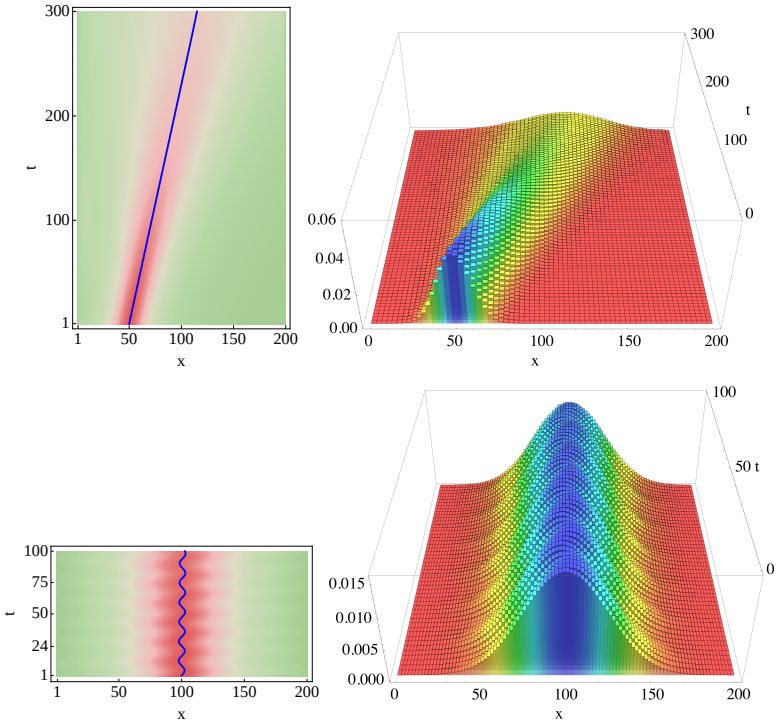


Figure 4.6: The automaton (4.31) evolution for smooth states. **Top figure:** automaton ($m = 0.4$) evolution for the smooth state given by $\sum_x g(x)e^{ik_0x}\,|x\rangle \otimes |+\rangle_{k_0}$ with $g(x)$ Gaussian with mean value $x_0 = 50$ and width $\hat{\sigma} = 10$. $|\pm\rangle_k$ are the positive/negative-frequency one-particle eigenstates (4.52) of the Dirac automaton with momentum $k_0 = 0.1$. The drift $v = 0.22$ and diffusion coefficient $D = 2.30$ are evaluated in Eqs. (5.16,5.17). The blue line in the left figure is the mean position evolution. **Bottom figure:** automaton ($m = 0.2$) evolution for a smooth state having both a particle and an antiparticle component $\sum_x g(x)e^{ik_0x}\,|x\rangle \otimes \frac{1}{\sqrt{2}}(|+\rangle_k + |-\rangle_{k_0})$, with the Gaussian $g(x)$ having mean $x_0 = 100$ and width $\hat{\sigma} = 25$, $k_0 = 0$. Notice the Zitterbewegung of the mean position.

4.3 The Dirac equation in the large-scale limit

In this Section we show how the Dirac QCA (4.31) recovers, for low momenta and masses, the dynamics of the Dirac equation

$$i\hbar\partial_t\psi(x,t) = \begin{pmatrix} -i\hbar c\partial_x & mc^2 \\ mc^2 & i\hbar c\partial_x \end{pmatrix}\psi(x,t). \tag{4.37}$$

According to the convention (4.12) for the Fourier transform, the automaton (4.31) in the momentum representation is[2]

$$U = \int_{\text{-}\pi}^{\pi} \mathrm{d}k\ U(k)\otimes|k\rangle\langle k|, \quad U(k) = \begin{pmatrix} ne^{ik} & -im \\ -im & ne^{-ik} \end{pmatrix}. \tag{4.38}$$

By taking a real power of the unitary matrix, we define an abstract Hamiltonian that describes the automaton evolution for continuous times, interpolating between time steps, namely

$$U^t = \exp(-iHt). \tag{4.39}$$

Upon diagonalizing the matrix $U(k)$ in Eq. (4.38), one obtains

$$H = \int_{-\pi}^{\pi} \mathrm{d}k\ H(k)\otimes|k\rangle\langle k|, \quad H(k) = \frac{\omega}{\sin(\omega)}\begin{pmatrix} -n\sin(k) & m \\ m & n\sin(k) \end{pmatrix}, \tag{4.40}$$

where ω is the dispersion relation of the automaton

$$\omega \coloneqq \omega(k) = \arccos(\sqrt{1-m^2}\cos(k)). \tag{4.41}$$

Notice that in the Weyl case of Eq. (4.33), namely for $m = 0$, the Hamiltonian and the dispersion relation are

$$U_W(k) = e^{-iH_W(k)} = \begin{pmatrix} e^{ik} & 0 \\ 0 & e^{-ik} \end{pmatrix},\ H_W(k) = \begin{pmatrix} -k & 0 \\ 0 & k \end{pmatrix} \quad \omega_W(k) = k. \tag{4.42}$$

It is worth mentioning that the Hamiltonians H and H_W are unphysical, as is unphysical the continuous process between two following time steps. Indeed, one could easily see that the Hamiltonian (4.40) involves interactions between all systems, including those very far apart.

We now expand $H(k)$ near $k = 0$ and $m = 0$, in order to recover the Dirac equation for the quantum field, along with the first corrections. One has

$$H(k) = \sum_{i,j=0}^{\infty} \frac{1}{i!j!}\left.\frac{\partial H(k)}{\partial k^i \partial m^j}\right|_{k,m=0} k^i m^j \tag{4.43}$$

[2]Notice that the Dirac automaton unitary $U(k)$ in the momentum space, which in the one-particle sector corresponds to a QW on $\mathbb{C}^2\otimes l_2(\mathbb{Z})$, does not factorize in a k-dependent component and an internal coin unitary matrix (see Remark 3.2).

and collecting all terms with the same overall order in m and k, we find

$$H(k) = H_D(k) + \frac{m}{3}\begin{pmatrix} mk & \frac{1}{2}(k^2+m^2) \\ \frac{1}{2}(k^2+m^2) & -mk \end{pmatrix} + O(m^p k^q), \qquad p+q=5, \tag{4.44}$$

where $H_D(k)$ is the Dirac Hamiltonian of Eq. (4.37) in the momentum representation and in Plack units ($\hbar = c = 1$)

$$H_D(k) = \begin{pmatrix} -k & m \\ m & k \end{pmatrix}. \tag{4.45}$$

The Hamiltonian (4.45) allows us to identify the parameters k and m of the automaton with the momentum and the mass of the Dirac field, respectively[3].

In Fig. 4.7 the automaton dispersion relation ω is compared with the Dirac one

$$\omega_D := \omega_D(k) = \sqrt{k^2+m^2}. \tag{4.46}$$

The dispersion relations show an overlap for k around $k = 0$ and small m, corresponding to the typical particle physics regime. Analytically, the leading term correction to the Dirac dispersion is given by

$$\omega = \omega_D\left(1 - \frac{m^2}{6}\frac{k^2-m^2}{k^2+m^2}\right). \tag{4.47}$$

Notice that the smallest non zero value of the mass $m = 0.3$ reported in Fig. 4.7 is huge compared to any known particle mass, and still is possible to find a sector in the momentum space around $k = 0$ where the automaton dispersion relation well overlaps the Dirac one. This means that for m sufficiently small it is possible to fix a cutoff in the momenta $|k| \leqslant \bar{k} < \pi$ such that in that regime the Dirac automaton and the usual Dirac evolution are very similar.

Intuitively taking a cutoff for the momenta corresponds to taking a *large-scale limit* of the automaton, namely to look at the automaton at larger scale in the position space. In Chapter 6 we will define the large-scale limit rigorously but we can already point out the conceptual difference with the *continuum limit*

[3]If we assume the automaton to act at the Planck scale the approximation (4.44) is well justified since the typical rest masses and momenta of particle physics experiments are many order of magnitude smaller than the Planck mass and the Planck momentum. In this case the maximal mass should coincide with the Planck one $m = 1 = m_P$ which in the usual units is $m_P = 2.17651(13) \times 10^8\text{kg} = 1.2209 \times 10^{28}\text{eV}/c^2$. A particle having rest mass 100 GeV/c^2 corresponds, in the automaton Planck units, to $m \approx 10^{-17}$ (the masses of an electron and a proton are as small as $m_e \approx 10^{-23}$ and $m_p \approx 10^{-19}$). Similarly a momentum of 10^{20}eV/c, the highest ever observed on the earth for Ultra-High Energy Cosmic Rays (UHECRs), see for example [132], corresponds in Planck units to $k_{CR} \approx 10^{-8}$, while a momentum of 1Tev, the order of magnitude of momenta in high energy physics experiments, *e.g.* at LHC, corresponds to $k_{LHC} \approx 10^{-16}$. The smallness of the automaton parameters in any physical accessible scenario completely justify the power expansion in Eq. (4.44).

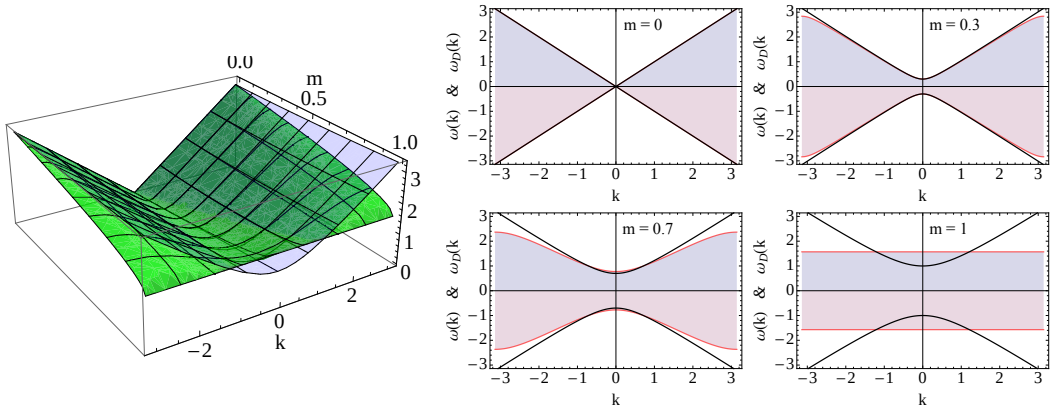


Figure 4.7: Comparison between the dispersion relation $\omega(k)$ of the Dirac automaton and the the Dirac equation one $\omega_D(k)$, respectively in Eqs. (4.41) and (4.46). In the left figure the dispersion relation is plotted versus the adimensional mass $m \in [0, 1]$ and momentum $k \in [-\pi, \pi]$. The green surface represent the automaton, whereas the blue one corresponds to Dirac. In the right figures $\omega(k)$ is plotted versus k for four values of m (the red line corresponds to the automaton, whereas the black one is the usual Dirac). We can see that the two dispersion relations coincide for small masses and momenta, and the larger is the mass the smaller is the overlap region around $k = 0$. Notice that in the left figure we have plotted $\pm\omega$ and $\pm\omega_D$ since, as in the Dirac equation, both positive and negative frequencies are allowed (see Section 4.3.2).

studied by some other authors—see for example Bialynicki-Birula in Ref. [49] and Meyer in Ref. [50]. Indeed, while in the large-scale limit the granularity of spacetime is fixed and the limit consists in considering typical states having momentum lower then some $\bar{k}$, the continuum limit is achieved for infinitesimal spatial period ℓ and time step τ:

$$\begin{aligned} &\text{large-scale limit:} \quad \ell = l_P,\ \tau = t_P \quad \text{and} \quad |k| \leqslant \bar{k} < \pi, \\ &\text{continuum limit:} \quad \ell \to 0,\ \tau \to 0. \end{aligned}$$

Remark 4.2 (Violation of Lorentz covariance) The QCA dispersion relation (4.41) is an evidence of Lorentz covariance violation at the automaton scale, hypothetically the Planck one. We will discuss this aspect in Chapter 7 where we will show the compatibility between the automaton dispersion relation and a model of deformed relativity of the kind considered in [24,25,62].

4.3.1 Avoiding the Fermion doubling

Here we want to stress the difference between the discrete evolution of the automaton and the usual discretization of QFT considered for example in the lattice gases models or in Lattice Gauge Theory. While in the automaton case it is the unitary to be discretized, in the literature it is usually the Hamiltonian, or the corresponding differential equation, which is approximated by its finite-difference counterpart.

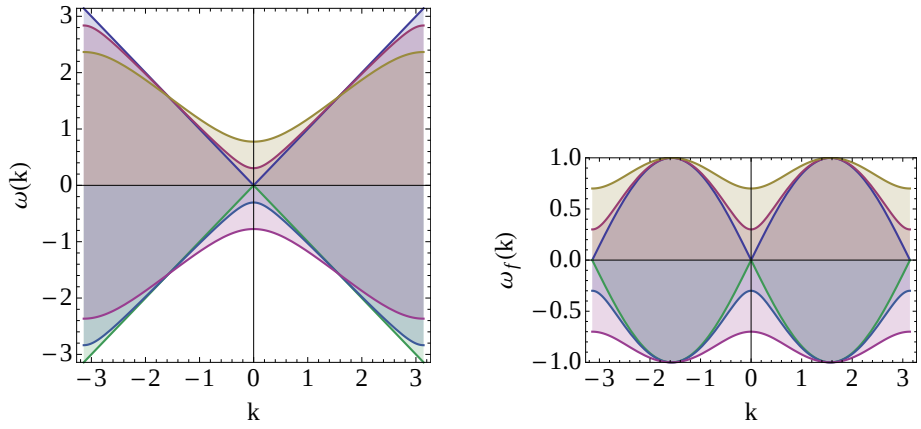


Figure 4.8: We show the difference between the dispersion relation of the Dirac automaton and the one corresponding to the finite-difference Dirac equation (the three plots in each figure correspond to different values of the mass m). On the left we have the automaton dispersion relation ω of Eq. (4.41) which is a monotone function of $|k|$. On the contrary the finite-difference dispersion relation ω_f (4.51) is not monotone in $|k|$ and allows states with $k \neq 0$ corresponding to a minimum of the energy (Fermion doubling).

The two different point of views lead to completely different ways of expressing "energy" versus the wave-vector k, namely to different dispersion relations.

On one side we have the discrete automaton of Eq. (4.31) whose dynamics can be recovered using the Hamiltonian $H = -i \log U$ (4.40) and the dispersion relation $\omega(k)$ (4.41). Despite the Hamiltonian is unphysical in between two discrete time steps, at the discrete points (x, t) of the causal network, where the automaton is defined, it provides exactly the automaton evolution.

On the other side we can replace the Dirac differential equation in Eq. (4.37) with its finite-difference version defining the finite-difference time and space derivatives as follows

$$\hat{\partial}_t = \tfrac{1}{2}(U - U^\dagger), \qquad \hat{\partial}_x = \tfrac{1}{2}(S - S^\dagger). \tag{4.48}$$

Accordingly Eq. (4.37) becomes

$$i\hat{\partial}_t \psi(x,t) = H_f\, \psi(x,t), \tag{4.49}$$

in terms of the finite-difference Hamiltonian [36]

$$H_f(k) = \tfrac{1}{2i}[U(k) - U^\dagger(k)] = \begin{pmatrix} -n\sin(k) & m \\ m & n\sin(k) \end{pmatrix}, \tag{4.50}$$

whose dispersion relation is

$$\omega_f(k) = \sin\omega(k). \tag{4.51}$$

As pointed out by many authors, *e.g.* by Susskind in Ref. [127], the finite difference differential equation leads to the so called Fermion doubling, namely

the existence of states with $k \neq 0$ corresponding to a minimum of the energy. Indeed the dispersion relation ω_f in (4.51) has more than one minimum as one can check in the right Fig. 4.8. Differently the dispersion relation ω in Eq. (4.41) is a monotone function of $|k|$ as one can verify looking at the left Fig. 4.8. Therefore in the "automaton like" discretization there is no Fermion doubling, namely no state other than for $k = 0$ corresponds to a minimum of the "energy" $\omega(k)$. For dimension greater than one, the dispersion relation can be as well made monotonic continuous by exploiting the multi-valued nature of the dispersion relation (see Ref. [49]).

4.3.2 Eigenstates of the Dirac automaton

The eigenvalues and the eigenvectors of the unitary matrix $U(k)$ in Eq. (4.38) are given by

$$u_k(s) = e^{-is\omega}, \quad |s\rangle_k := \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{1 - sv} \\ s\sqrt{1 + sv} \end{bmatrix}, \quad s = \pm, \tag{4.52}$$

in terms of the automaton dispersion relation (4.41) and group velocity

$$v := \frac{\partial \omega(k)}{\partial k} = \sqrt{\frac{1 - m^2}{1 + m^2 \cot^2(k)}}, \tag{4.53}$$

which is plotted in Fig. (4.9). In analogy with the Dirac theory the $s = +1$ eigenvalues in Eq. (4.52) correspond to positive-energy particle states, whereas the negative $s = -1$ eigenvalues correspond to negative-energy antiparticle states. Accordingly the group velocities for positive- and negative-energy states have the same module and opposite sign as one can check in Fig. (4.9). The most general state $|\psi\rangle$ is thus a superposition of a positive- and a negative-energy state, *i.e.* $|\psi_+\rangle + |\psi_-\rangle$, and typical aspects of the Dirac field dynamics, such as the *Zitterbewegung* (see the right Fig. 4.6) and the *Klein paradox*, are also dynamical feature of the Dirac automaton as shown in Chapter 5.

The QCA automaton U generally operates on the vector field ψ which describes an arbitrary number of particles. The vacuum state for the automaton is defined as the state $|\Omega\rangle$ such that

$$\psi_s(k) |\Omega\rangle = 0 \qquad \forall s = \pm, \quad \forall k \in [-\pi, \pi]. \tag{4.54}$$

Up to now, we have not specified the nature (Fermionic or Bosonic) of the field, and indeed the same automaton could be used to describe both cases (along with anyons and parastatistics). Even if in one dimension the spin of the field is not a relevant feature, we will focus for simplicity on the Fermionic case of anticommuting field. Clearly a N-particle state can be obtained by acting with the field operator on the vacuum as follows

$$|N, \boldsymbol{k}, \boldsymbol{s}\rangle = \left(\prod_{i=1}^{N} \psi^\dagger_{s_i}(k_i) \right) |\Omega\rangle \,. \tag{4.55}$$

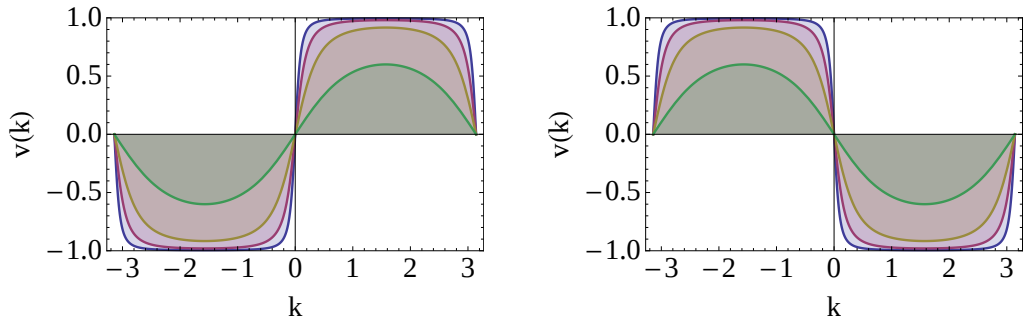


Figure 4.9: The group velocity (4.53) of the Dirac automaton for different values of the mass $m = 0.1, 0.2, 0.4, 0.8$. On the left (right) the group velocity corresponding to positive (negative) eigenvalues ω.

Specifically, for the $N = 1$ particle eigenstates of U, we write

$$\psi_s^\dagger(k)\,|\Omega\rangle = |s\rangle_k\,|k\rangle\,, \tag{4.56}$$

whereas for $N = 2$ we have

$$\psi_{s_1}^\dagger(k_1)\psi_{s_2}^\dagger(k_2)\,|\Omega\rangle = |s_1\rangle_{k_1}\,|s_2\rangle_{k_2}\,|k_1, k_2\rangle\,, \tag{4.57}$$

where $|k_1, k_2\rangle = -\,|k_2, k_1\rangle$, and so forth for $N > 2$. The corresponding eigenvalues of the logarithm of U are

$$\omega(N, \boldsymbol{k}, \boldsymbol{s}) = \sum_{i=1}^{N} s_i\,\omega(k_i, m). \tag{4.58}$$

4.4 Derivation of the three-dimensional Dirac automaton

Even though this thesis is entirely devoted to the one space-dimensional case, it is worth mentioning that G. M. D'Ariano and P. Perinotti have recently derived [60] from informational principles the automaton describing the Dirac equation in $d = 1, 2, 3$ space-dimensions. Clearly for $d = 1$ they obtain the same automaton (4.31) derived in Section 4.2. Since the result in [60] is not part of the present thesis we do not enter in the details of the derivation.

The informational principles in [60] are essentially the same considered in Section 4.2, namely (i) unitarity of the evolution (ii) locality of the evolution (iii) homogeneity of the interaction topology (iv) isotropy (or covariance) with respect to the discrete topology (v) and the minimal dimension for a non-trivial evolution. However, while in one dimension the topology of the homogeneous causal network was necessarily the one in Fig. 4.3 (with spatial-lattice $\mathbb{Z}$), in higher dimension there is a variety of inequivalent homogeneous spatial-lattices which in principle could be compatible with the aforementioned

assumptions. The authors have then considered the problem from a very general point of view, taking a denumerable set X of interacting quantum systems, with $N_x \subset X$ the subset of systems interacting with $x \in X$[4] . The locality (ii) and the homogeneity (iii) requirements amount to ask that N_x is finite and with the same cardinality N for any $x \in X$. Hence the structure of the connections between systems can be regarded as the application of generators $y \in N = N_+ \cup N_- \cup \{e\}$ (where N_- is the set of inverses of the elements in N_+, according to the isotropy assumption (iv), while e is the identity element) of a discrete group which allows to move from an element $x \in X$ to another element $x' = yx \in X$, identifying the set X with the group itself. The suitable mathematical description of this structure, as pointed out in [60], is the Cayley graph $\mathrm{Cay}(X, N)$ of the group X, with the quantum systems at the nodes of the graph and the links representing the quantum systems interactions.

The general QCA for a field ψ having Λ internal degrees of freedom, is then a unitary (assumption (i)) operator

$$U = \sum_{y \in N} U_y \otimes S_y \tag{4.59}$$

on $\mathbb{C}^\Lambda \otimes l_2(X)$, where S is an irreducible representation of the group X over the Hilbert space $l_2(X)$, and N is the finite set of generators of X[5]. The covariance (iv) with respect to the discrete topology is satisfied if the QCA is covariant under a representation O of some group G which acts transitively on the generators of X, namely[6]

$$U = \sum_{y \in N} U_y \otimes S_y = \sum_{y \in N} O_g U_h O_g^\dagger \otimes S_{g(y)}. \tag{4.60}$$

After developing the above general construction, the authors of [60] restrict the infinite groups X to those having a Cayley graph that is *quasi-isometrically embeddable* [133] in the Euclidean space $\mathbb{R}^3$, among which the only BCC (body centered cubic) Bravais lattice is compatible with the unitarity constraint (i). Then it is shown that the unique (modulo CP symmetry) automaton satisfying all the requirements has minimal dimension $\Lambda = 2$ and recovers the Weyl evolution in the large-scale limit $|k| \ll 1$, where $k = (k_x, k_y, k_z)^T$ is the the wave-vector in the first Brillouin zone of the BCC lattice.

The Dirac QCA in $d = 3$ is finally obtained coupling two Weyl automata (thus describing a field with $\Lambda = 4$ internal degrees of freedom) in the only possible ways consistent with the locality (ii) and unitarity (i) assumptions. This results in only two admissible QCAs which are related each other by a

[4]Notice that X corresponds to the automaton spatial-lattice while N_x is the equivalent of the neighborhood scheme.

[5]The S_y with $y \in N$, correspond to the usual shifts on the lattice, as in Eq. (4.15) for the one-dimensional field automaton.

[6]This condition correspond to the one-dimensional field automaton covariance of Definition (4.3).

CPT symmetry and which recover the Dirac evolution in the large-scale limit of small momenta and masses $|k|, m \ll 1$. The dispersion relations of the two Dirac QCAs are

$$\omega^{\pm}(k) = \arccos(\sqrt{1-m^2}(c_x c_y c_z \pm s_x s_y s_z)), \tag{4.61}$$

with $c_\alpha = \cos(k_\alpha/\sqrt{3})$, $s_\alpha = \sin(k_\alpha/\sqrt{3})$, for $\alpha = x, y, z$, and $k_x = \frac{1}{2}(k_1 + k_2)$, $k_y = \frac{1}{2}(k_1 + k_3)$, $k_z = \frac{1}{2}(k_1 + k_4)$.

CHAPTER 5

The Dirac automaton in one dimension: free evolution and scattering from potential

In this Chapter we analyze in the details the one-particle sector of the Dirac automaton derived in Chapter 4. We show that that particle states on the automaton are "smooth-states" peaked around some momentum eigenstate of the QCA. We present an analytical approximation method for evaluating the automaton evolution for one-particle states deriving a momentum-dependent differential equation for the automaton.

Then we analyse in the details dynamical quantities as the particle position, momentum and velocity, along with their evolution both in the free case and in the presence of a potential. We show that typical features of the Dirac quantum field evolution–as *Zitterbewegung* and *Klein paradox*–are recovered from the quantum information processing of the QCA.

Recently we have witnessed a renewed interest in the Dirac equation features in solid state and atomic physics, which provide physical hardwares to simulate the Dirac dynamics. Zitterbewegung can be seen in the response of electrons to external fields [134] and can appear for nonrelativistic particles in a crystal [135–137], quasiparticles in superconductors [138] and systems with spin-orbit coupling [139, 140]. Proving that the oscillation behavior is not unique to Dirac electrons, but rather is a generic feature of spinor systems with linear dispersion relations, these works opened the way for possible simulations of Zitterbewegung using for example trapped ions [141, 142], two-band

crystalline structure such as graphene [143, 144] or semiconductors [145–149], ultra cold atoms [150], and finally photonic crystals [151].

The Klein paradox (tunneling of relativistic particles), which provides insight in the mechanics of relativistic particles propagating through potential barriers and in vacuum polarization effects, has been a focus in the topic of graphene as a simulator for the Dirac equation, as in Refs. [152, 153], and in trapped ions, as in [142]. Recently also microfabricated optical waveguide circuits have become an alternative physical simulator for particle dynamics [154].

In Section 5.1 we derive a momentum-dependent differential equation describing the automaton dynamics for one-particle states. We show that it recovers the usual Dirac and Schrödinger evolutions in the appropriate regimes. The analytic approximation is compared with exact simulations, and the leading order corrections to the Dirac equation is evaluated. In Section 5.2 we present the evolution of position and momentum operators for the automaton, showing the Zitterbewegung behavior given by the interference between positive and negative frequencies. In Section 5.3 we modify the Dirac QCA in order to add a potential in the free evolution, and show the automaton dynamics in the presence of a barrier for one-particle states.

The content of this Chapter is based on Refs. [59] and [61].

5.1 Dispersive differential equation describing the Dirac automaton evolution

Here we describe analytically the evolution of the Dirac QCA for suitable one-particle states. These are states smoothly peaked around some one-particle momentum eigenstate (4.56) of the Dirac automaton and then localized in a finite region of the lattice. Their evolution will be well described by a differential equation whose coefficients depend on the mass m and on the momentum k_0 of the state. We will denote this states *smooth-states*:

Definition 5.1 (Smooth-states) *A smooth-state for the Dirac automaton in Eq.* (4.31) *is a state smoothly peaked around some k_0, namely*

$$|\psi(0)\rangle = \int_{-\pi}^{\pi} \frac{\mathrm{d}k}{\sqrt{2\pi}}\, g_{k_0}(k,0)\, |s\rangle_k\, |k\rangle\,, \qquad s=\pm, \tag{5.1}$$

where $g(k,0) \in C_0^\infty[-\pi,\pi]$ is a smooth function satisfying the bound

$$\frac{1}{2\pi}\int_{k_0-\sigma}^{k_0+\sigma} \mathrm{d}k\ |g_{k_0}(k,0)|^2 \geqslant 1-\epsilon, \quad \sigma,\epsilon>0, \tag{5.2}$$

and the two-component vector $|s\rangle_k$ is defined in Eq. (4.52).

Remark 5.1 In the following it will be convenient to work with the continuous time t, interpolating exactly the discrete automaton evolution U^t. Therefore

we will take x, t to be real-valued continuous variable by extending the Fourier transform

$$|\psi(x,t)\rangle = \int_{-\pi}^{\pi} \frac{\mathrm{d}k}{\sqrt{2\pi}}\, e^{ikx} U^t(k)\, |\psi(k,0)\rangle \tag{5.3}$$

to real x, t. Since the automaton is band limited in momenta $0 \leqslant k \leqslant \pi$, then the continuous function $\psi(x,t)$ is completely defined by its value on the discrete points (x,t) of the automaton causal network. The possibility of sampling a band-limited function is stated in the Nyquist-Shannon *Sampling Theorem* (see Appendix A). However, all numerical results will be given only for discrete t, namely for repeated applications of the automaton unitary U in Eq. (4.31), and for discrete lattice sites x.

In the following Proposition we derive the differential equation governing the evolution for smooth-states:

Proposition 5.1 (Dispersive differential equation) *Consider the evolution of the Dirac automaton* (4.31) *on a smooth-state* $|\psi(0)\rangle$ *(see Definition 5.1) peaked around* $|s\rangle_{k_0} |k_0\rangle$ *and having width* σ*. Then for any positive integer* p*, the state at time* t *is given by*

$$|\psi(x,t)\rangle = e^{i(k_0 x - \omega_0 t)} |\tilde{\phi}(x,t)\rangle - \epsilon - \gamma \sigma^{p+1} t - \mathcal{O}(\sigma^{p+3}) t \tag{5.4}$$

where $|\tilde{\phi}(x,t)\rangle$ *is solution of the following differential equation*

$$i\partial_t |\tilde{\phi}(x,t)\rangle = s \sum_{n=1}^{p} \frac{(i)^n \omega_0^{(n)}}{n!} \frac{\partial^n}{\partial x^n} |\tilde{\phi}(x,t)\rangle, \tag{5.5}$$

with

$$\gamma := \frac{\omega_0^{(p+1)}}{2\pi} \int_{k_0-\sigma}^{k_0+\sigma} \mathrm{d}k\ |g_{k_0}(k,0)|^2, \qquad \omega_0^{(n)} := \left.\frac{\partial^n \omega(k)}{\partial k^n}\right|_{k=k_0}, \tag{5.6}$$

and $\omega(k)$ *the Dirac automaton dispersion relation* (4.41)*.*

Proof. First we notice that at time t the state (5.1) in the momentum representation is simply

$$|\psi(k,t)\rangle = e^{-is\omega t} |\psi(k,0)\rangle = e^{-i\omega t} g_{k_0}(k,0) |s\rangle_k . \tag{5.7}$$

At the same time t, the state in the position representation can then be written as

$$|\psi(t)\rangle = \sum_x |\psi(x,t)\rangle |x\rangle , \qquad |\psi(x,t)\rangle := e^{i(k_0 x - s\,\omega_0 t)} |\phi(x,t)\rangle , \tag{5.8}$$

$$|\phi(x,t)\rangle := \int_{-\pi}^{\pi} \frac{\mathrm{d}k}{\sqrt{2\pi}}\, e^{i(Kx - s\Omega(k,m)t)} g_{k_0}(k,0) |s\rangle_k , \tag{5.9}$$

where we posed

$$\begin{aligned} &K = k - k_0, \\ &\Omega(k) = \omega(k) - \omega_0, \quad \omega_0 = \omega(k_0). \end{aligned} \tag{5.10}$$

As stated in Remark 5.1 it is convenient to take x, t to be real-valued continuous variables by extending the Fourier transform in Eq. (5.9) to real x, t.

Now we derive the integral in Eq. (5.9) with respect to t and expand Ω vs k around k_0. Then, taking the resulting derivatives with respect to x out of the integral (using the dominated derivative theorem), we obtain the following differential equation whose coefficients depend on the momentum k_0 and on the mass m

$$i\partial_t|\phi(x,t)\rangle = s\sum_{n=1}^{\infty}\frac{(i)^n\omega_0^{(n)}}{n!}\frac{\partial^n}{\partial x^n}|\phi(x,t)\rangle, \quad \omega_0^{(n)} = \left.\frac{\partial^n\omega(k)}{\partial k^n}\right|_{k=k_0}. \tag{5.11}$$

If we truncate the expansion on the RHS of Eq. (5.11) at the pth order and denote by $|\tilde{\phi}(x,t)\rangle$ the solution of the corresponding truncated differential equation with the identification of the initial condition $|\tilde{\phi}(x,0)\rangle = |\phi(x,0)\rangle$, we get the approximate state (5.8) at time t

$$|\tilde{\psi}(x,t)\rangle = e^{i(k_0x - s\,\omega_0 t)}|\tilde{\phi}(x,t)\rangle. \tag{5.12}$$

If the initial smooth-state satisfies Eq. (5.2), then the accuracy of the approximation (5.12) in terms of the parameters σ and ϵ is evaluated by the overlap between the states (5.8) and (5.12), that is

$$\begin{aligned} |\langle\tilde{\psi}(t)|\psi(t)\rangle| &= \left|\int_{-\pi}^{\pi}\frac{\mathrm{d}k}{2\pi}\,e^{-i(\omega_{k_0}^{(p+1)}k^{p+1}+\mathcal{O}(k^{p+2}))t}|g_{k_0}(k,0)|^2\right| \geqslant \\ &\geqslant \left|\frac{1}{2\pi}\int_{k_0-\sigma}^{k_0+\sigma}\mathrm{d}k\;e^{-i(\omega_{k_0}^{(p+1)}k^{p+1}+\mathcal{O}(k^{p+2}))t}|g_{k_0}(k,0)|^2\right| + \\ &\qquad - \left|\frac{1}{2\pi}\int_{|k-k_0|\geqslant\sigma}\mathrm{d}k\;e^{-i(\omega_{k_0}^{(p+1)}k^{p+1}+\mathcal{O}(k^{p+2}))t}|g_{k_0}(k,0)|^2\right| \geqslant \\ &\geqslant \left|1 - it\frac{\omega_{k_0}^{(p+1)}\sigma^{p+1}}{2\pi}\int_{k_0-\sigma}^{k_0+\sigma}\mathrm{d}k\;|g_{k_0}(k,0)|^2 - \mathcal{O}(\sigma^{p+3})t\right| - \epsilon \geqslant \\ &\geqslant 1 - \epsilon - \gamma\sigma^{p+1}t - \mathcal{O}(\sigma^{p+3})t \end{aligned} \tag{5.13}$$

where

$$\gamma = \frac{\omega_0^{(p+1)}}{2\pi}\int_{k_0-\sigma}^{k_0+\sigma}\mathrm{d}k\;|g_{k_0}(k,0)|^2.$$

Therefore the exact state $|\psi(x,t)\rangle$ at time t can be approximated by(5.12) with the accuracy given by the overlap (5.13), which proves the thesis (5.4). ■

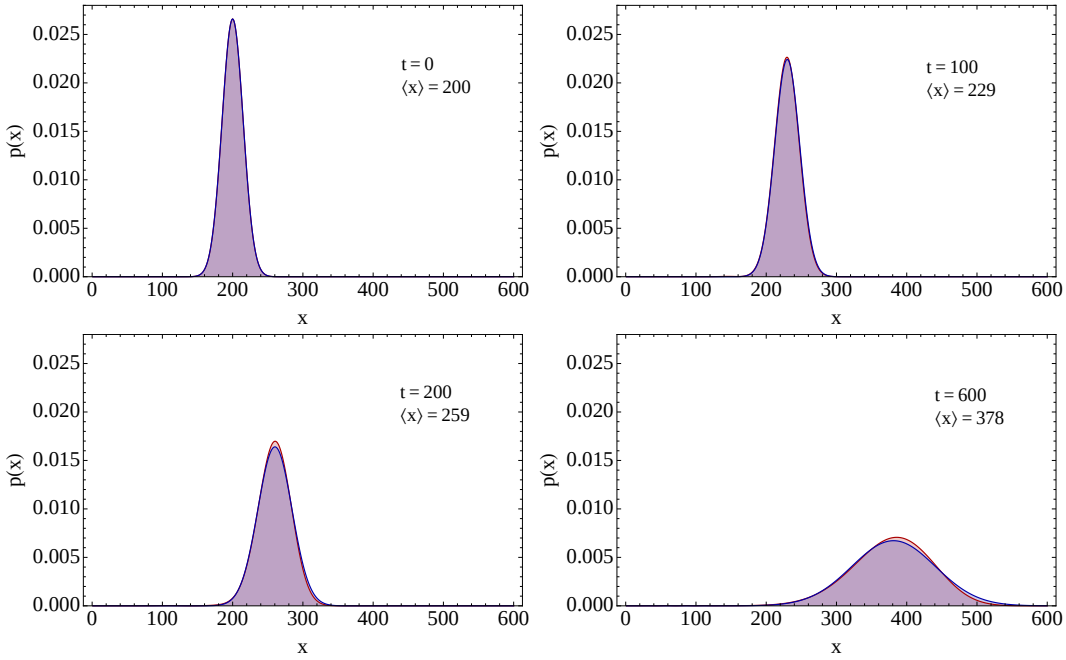


Figure 5.1: Test of the approximation (5.14) of the ($m = 0.4$) Dirac automaton (4.31) evolution. The initial smooth-state (5.1) is a Gaussian peaked around the momentum $k_0 = 0.1$ and having width in position $\hat{\sigma} = \sigma^{-1} = 10$. We show a four times $t = 0$, 100, 200, 600 comparison between the automaton probability distribution $|\psi(x,t)|^2$ (in red) and the approximate solution (5.15) $|\tilde{\psi}(x,t)|^2$ (in blue). The drift velocity and the diffusion coefficient are respectively $v = 0.22$ and $D = 2.30$.

In the following we will be interested in the second order approximation of Proposition 5.1, namely

$$|\psi(x,t)\rangle = e^{i(k_0 x - \omega_0 t)}|\tilde{\phi}(x,t)\rangle - \epsilon - \gamma\sigma^3 t - \mathcal{O}(\sigma^5)t, \tag{5.14}$$

with $|\tilde{\phi}(x,t)\rangle$ satisfying the second order diffusive differential equation

$$i\partial_t|\tilde{\phi}(x,t)\rangle = s\left(iv\frac{\partial}{\partial x} - \frac{1}{2}D\frac{\partial^2}{\partial x^2}\right)|\tilde{\phi}(x,t)\rangle, \tag{5.15}$$

with the drift v and the diffusion D coefficients depending on k and m as follows

$$v := \omega_0^{(1)} = \sqrt{\frac{1-m^2}{1+m^2\cot^2(k_0)}}, \tag{5.16}$$

$$D := \omega_0^{(2)} = \frac{\sqrt{1-m^2}m^2\cos(k_0)}{(\sin^2(k_0)+m^2\cos^2(k_0))^{\frac{3}{2}}}. \tag{5.17}$$

We can test the accuracy of the approximation (5.14) by comparing $|\tilde{\psi}(x,t)\rangle$ with the exact $|\psi(x,t)\rangle$ given by the automaton simulation.

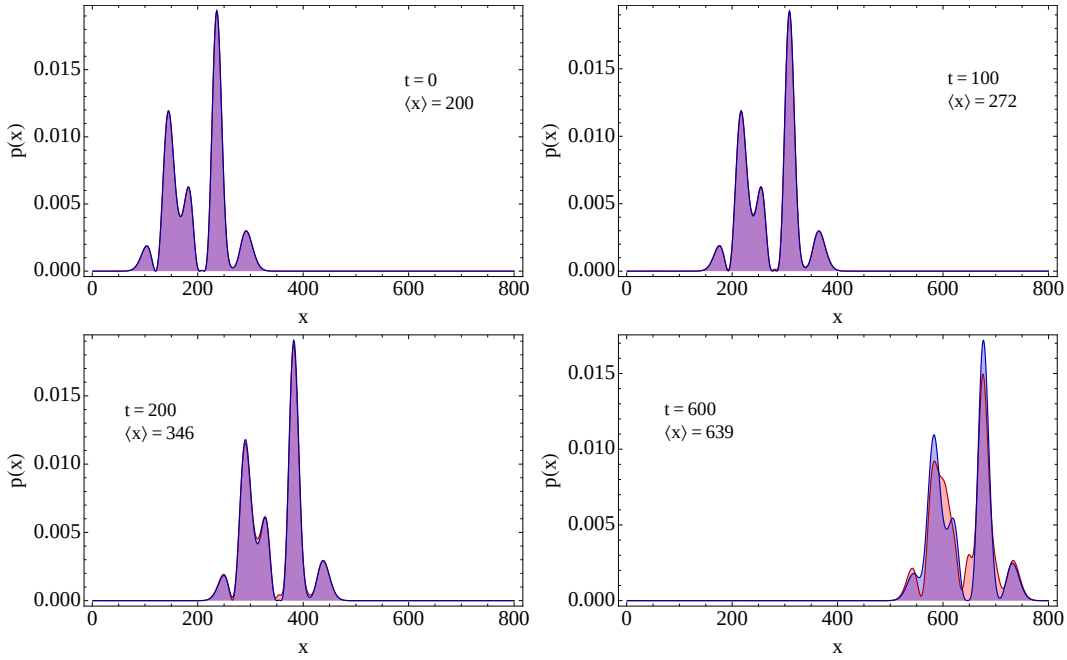


Figure 5.2: Test of the approximation (5.14) of the Dirac automaton (4.31) evolution. The smooth-state (5.1) is a superposition of Hermite functions (the polynomials $H_j(x)$ multiplied by the Gaussian) peaked around momentum $k_0 = 3\pi/10$, specifically $|\psi(x,0)\rangle = \mathcal{A}e^{ik_0 x}\sum_{j\in\mathbb{N}} c_j e^{-x^2/4\hat{\sigma}^2} H_j(x/2\hat{\sigma})|+\rangle_{k0}$ where $\hat{\sigma} = \sigma^{-1} = 20$ is the position variance corresponding to momentum variance σ, and the nonvanishing terms are $c_0 = \sqrt{1/3}$, $c_2 = \sqrt{4/9}$, $c_7 = \sqrt{2/9}$. The automaton mass is $m = 0.6$. In the picture we show a comparison at four different times $t = 0, 100, 200, 600$ between the automaton probability distribution $|\psi(x,t)|^2$ (in red) and the solution of the differential equation (5.15) $|\tilde{\psi}(x,t)|^2$ (in blue). The drift and the diffusion coefficients are respectively $v = 0.73$ and $D = 0.31$. The mean position moves at the group velocity given by the drift coefficient v. The approximation remains accurate even for position spread $\hat{\sigma} = 20$ "Planck" lengths. According to Eq. (5.14) one has significant deviations for $t \approx \gamma\sigma^3$, which is $t = 600$ in the present case. However, a reasonable spread $\hat{\sigma}$ in a typical particle physics scenario is the Fermi length $\hat{\sigma} \approx 10^{20}$, that would need a time t comparable to many universe life-times to introduce a significant error. The ϵ error in Eq. (5.14) can be taken very small by considering $n\sigma$ instead of σ in Eq. (5.2). For Gaussian states it is enough to consider 3σ to get $\epsilon \approx 10^{-3}$.

In Fig. 5.1 it is $m = 0.4$ and the initial smooth-state is a Gaussian peaked around the momentum $k_0 = 0.1$ and having width σ

$$|\psi(x,0)\rangle = Ae^{ik_0 x}e^{-x^2/4\hat{\sigma}^2}\,|+\rangle_{k_0}\,, \qquad \hat{\sigma} = \sigma^{-1} = 10, \tag{5.18}$$

where $\hat{\sigma}$ is the width in position and A represents the normalization factor. The differential equation in Eq. (5.15) can be easily solved for Gaussian states

and the approximate solution (5.14) at time t is

$$|\tilde{\psi}(x,t)\rangle = A' \exp\left(\frac{-(x+vt)^2 - i\omega}{4\hat{\sigma}(t)}\right)|+\rangle_{k_0}, \qquad \hat{\sigma}(t) = \hat{\sigma}\sqrt{1+i\frac{D}{2\hat{\sigma}^2}t}, \tag{5.19}$$

where the drift and diffusion coefficients are respectively $v = 0.22$ and $D = 2.30$.

In Fig. 5.2 it is $m = 0.6$ and the initial smooth-state (5.1) is a superposition of Hermite functions (the Hermite polynomials $H_j(x)$ multiplied by a Gaussian) peaked around the momentum $k_0 = 3\pi/10$ and having width σ

$$|\psi(x,0)\rangle = Ae^{ik_0x}\sum_{j\in\mathbb{N}} c_j e^{-x^2/4\hat{\sigma}^2} H_j(x/2\hat{\sigma})\,|+\rangle_{k_0} \qquad \hat{\sigma} = \sigma^{-1} = 20, \tag{5.20}$$

(see the Figure for the details).

One can notice how the approximation remains accurate even for states having a position width of few sites on the lattice (few "Planck" lengths). In Figs. 5.1 and 5.2 we have respectively $\hat{\sigma} = 10$ and $\hat{\sigma} = 20$. For example, according to Eq. (5.14), in the simulation of Fig. 5.2 one has a significant deviation for $t \approx \gamma\sigma^3 = 600$ while the error ϵ can be taken arbitrary small by considering $n\sigma$ instead of σ in Eq. (5.2) (for Gaussian states it is enough to consider 3σ to get $\epsilon \approx 10^{-3}$).

On the other hand, in a more typical scenario it is reasonable to assume the state width $\hat{\sigma}$ of the order of one Fermi, as in a typical particle physics context. In this case the time t needed for a significant departure would be comparable to many universe life-times.

5.1.1 First order correction to the Dirac and Scrödinger free evolution

In the relativistic regime, namely for

$$k, m \ll 1, \qquad k/m \gg 1, \tag{5.21}$$

the momentum-dependent differential equation (5.15) approaches the Dirac equation. The leading order and the corrections to the drift and diffusion coefficients introduced by the automaton evolution are

$$v = v_D\left(1 - \frac{1}{3}m^2 + \frac{1}{6}\frac{m^2k^2}{k^2+m^2}\right), \qquad v_D := \frac{k}{\sqrt{k^2+m^2}}, \tag{5.22}$$

$$D = D_D\left(1 + \frac{1}{3}m^2k^2 - \frac{1}{2}\frac{m^2k^4}{k^2+m^2}\right), \qquad D_D := \frac{m^2}{\sqrt{(k^2+m^2)^3}}. \tag{5.23}$$

The leading order terms correspond to the Dirac equation drift v_D and diffusion D_D coefficients.

In the non relativistic regime, namely

$$k, m \ll 1, \qquad k/m \ll 1, \tag{5.24}$$

the usual Schrödinger drift and diffusion coefficients are recovered with the following corrections

$$v = V_S \left(1 + \frac{1}{3}m^2\right), \qquad v_S \coloneqq \frac{k}{m}, \tag{5.25}$$

$$D = D_S \left(1 + \frac{5}{6}k^2\right), \qquad D_S \coloneqq \frac{1}{m}. \tag{5.26}$$

Notice that the leading terms are the usual group velocity v_S and diffusion coefficient D_S of the Schrödinger equation.

5.2 Evolution of Position and momentum operators

In this Section we show how the Dirac automaton reproduces a typical feature of the one-particle Dirac dynamics, namely the Zitterbewegung.

The Zitterbewegung was first recognized by Schrödinger in 1930 [155] who noticed that in the Dirac equation describing the free relativistic electron the velocity operator does not commute with the Dirac Hamiltonian. It follows that in the evolution of the position operator, in addition to classical motion, it appears a very fast periodic oscillation with frequency $2mc^2$ and amplitude equal to the Compton wavelength $\hbar/mc$, with m the mass of the relativistic particle. This jittering motion, first encountered in the Dirac theory of the electron, was then shown [134] to arise from the interference of states corresponding to the positive and negative energies resulting from the Dirac equation, with the trembling disappearing for long time evolutions [156]. The Zitterbewegung oscillation cannot be directly observed by current experimental techniques for a Dirac electron since its amplitude results to be very small $\approx 10^{-12}$ m. However, it can be seen in a number of solid-state, atomic-physics, photonic-cristal and optical waveguide simulators, as quoted in the introduction to the Chapter.

5.2.1 The automaton position and momentum operators

Up to now we have considered only smooth-states 5.1 whose Dirac automaton evolution is well described by the approximate differential equation derived in Proposition 5.1. On the other hand in the automaton framework are allowed states very far apart from the smooth-ones and in the limit one can also consider perfectly localized states:

$$|\psi\rangle = |\zeta\rangle\,|x\rangle \qquad x \in \mathbb{Z}, \qquad \zeta = c_r\,|r\rangle + c_l\,|l\rangle\,,\ |c_r|^2 + |c_l|^2 = 1, \tag{5.27}$$

where $|r\rangle$ and $|l\rangle$ denote the $\mathbb{C}^2$ basis vectors corresponding to the Dirac field representation in Eq. (4.25).

The position operator X providing the representation $|x\rangle$, i.e. such that $X\,|\eta\rangle\,|x\rangle = x\,|\eta\rangle\,|x\rangle$, is defined as follows

$$X = \sum_{x\in\mathbb{Z}} x(I \otimes |x\rangle\langle x|). \tag{5.28}$$

Accordingly the average position for an arbitrary one-particle state

$$|\psi\rangle = \sum_{x\in\mathbb{Z},\,\eta=r,l} g_\eta(x)\,|\eta\rangle\,|x\rangle\,, \tag{5.29}$$

is given by

$$\langle\psi|\,X\,|\psi\rangle\,. \tag{5.30}$$

The definition of the mechanical momentum would need an interacting theory allowing momentum exchange between different particles. However, in Chapter 4 and Section 5.1 we have seen that for small k and m, the wave-vector k (namely the conjugated variable of x via the Fourier transform) corresponds to the Dirac particle momentum. Moreover the momentum operator should correspond to the generator of translations over the discrete one-dimensional lattice. Therefore, as conjugated momentum we take the following operator

$$P = \int_{-\pi}^{\pi} \frac{\mathrm{d}k}{2\pi}\, k(I \otimes |k\rangle\langle k|). \tag{5.31}$$

We can now compute the commutator between X and P as defined in Eqs. (5.28) and (5.31), that is

$$\begin{aligned}[X,P] &= \sum_x \int_{-\pi}^{\pi} \frac{\mathrm{d}k}{2\pi}\ \sum_y xk|x\rangle\langle y|e^{-ik(x-y)} - \int_{-\pi}^{\pi} \frac{\mathrm{d}k}{2\pi}\ \sum_{z\nu} zk|\nu\rangle\langle z|e^{ik(z-\nu)} \\ &= \sum_{xy} \int_{-\pi}^{\pi} \frac{\mathrm{d}k}{2\pi}\ (x-y)k|x\rangle\langle y|e^{-ik(x-y)},\end{aligned} \tag{5.32}$$

where in the second equality it was possible to interchange the sum and the integral according to the Fubini Theorem. Now integrating by parts we get

$$\langle\psi|\,[X,P]\,|\psi\rangle = i - \frac{i}{2}\sum_{\eta=r,l}\left(|\hat{g}_\eta(\pi)|^2 + \hat{g}_\eta(-\pi)|^2\right), \tag{5.33}$$

where $|\psi\rangle$ is the generic state of Eq. (5.29) and $\hat{g}(k)$ is the discrete Fourier transform of $g(x)$. Eq. (5.33) differs from the usual canonical commutation relation $[X,P] = i$ by a boundary term. This is in agreement with the existence of perfectly localizable states on the automaton (5.27),

$$|\zeta\rangle\,|x\rangle = \sum_{y\in\mathbb{Z},\,\eta=r,l} g_\eta(y)\,|\eta\rangle\,|y\rangle\,, \qquad g_\eta(x) = c_\eta\delta_{xy}, \tag{5.34}$$

for which $\hat{g}_\eta(\pi) = \hat{g}_\eta(-\pi) = c_\eta$ and the expectation value of the commutator (5.33) vanishes. From now on we will consider states having negligible boundary term in (5.33).

In Section 4.3.2 we have derived the momentum eigenstates (4.52) of the Dirac automaton

$$U(k)\left|s\right\rangle_k = e^{-isH(k)}\left|s\right\rangle_k = e^{-is\omega(k)}\left|s\right\rangle_k, \qquad s = \pm \tag{5.35}$$

where we remind the Hamiltonian (4.40) to be

$$H(k) = \omega\left(\left|+\right\rangle_k\left\langle+\right|_k - \left|-\right\rangle_k\left\langle-\right|_k\right) = \mathrm{sinc}^{-1}\omega(-n\sin k\,\sigma_3 + m\sigma_1). \tag{5.36}$$

We observed that the positive ($s = 1$) and negative ($s = -1$) eigenstates correspond to the particle and antiparticle states of the automaton. As in QFT, one can define the *Newton-Wigner* position operator X_{NW} which does not mix positive and negative eigenstates. In order to do that we introduce the operator U_{FW} providing the *Foldy-Wouthuysen* representation of the Dirac automaton in which the Hamiltonian $H(k)$ is diagonal

$$U_{FW}(k) : \{\left|r\right\rangle, \left|l\right\rangle\} \to \{\left|+\right\rangle_k, \left|-\right\rangle_k\}, \qquad U_{FW}^{-1}(k)H(k)U_{FW}(k) = \omega\sigma_3. \tag{5.37}$$

By direct computation we find that

$$U_{FW} = \int_{-\pi}^{\pi} \mathrm{d}k\; U_{FW}(k) \otimes \left|k\right\rangle\left\langle k\right| \tag{5.38}$$

$$U_{FW}(k) = U_{FW}^{-1}(k) = \tfrac{1}{\sqrt{2}}(\sqrt{1-v}\sigma_3 + \sqrt{1+v}\sigma_1), \tag{5.39}$$

and the Newton-Wigner rotated position operator is thus defined as

$$X_{NW} = U_{FW}XU_{FW}. \tag{5.40}$$

5.2.2 The Zitterbewegung

In the following analytical derivation we work with continuous x and t (see Remark 5.1). However, as in the previous Section, the numerical results will be given only for discrete t, namely for repeated applications of the automaton unitary operator (4.31).

The evolution of the position operator (5.28)

$$X(t) = U^{-t}XU^t \tag{5.41}$$

can be easily computed via the velocity and the acceleration operators

$$V(t) = \int_{-\pi}^{\pi} \mathrm{d}k\; V(k,t)|k\rangle\langle k|, \quad A(t) = \int_{-\pi}^{\pi} \mathrm{d}k\; A(k,t)|k\rangle\langle k|,$$

derived by the commutator with the automaton Hamiltonian (5.36)

$$V(t) = i[H, X(t)], \qquad A(t) = i[H, V(t)]. \tag{5.42}$$

From an easy computation it follows

$$V(k) = \frac{\sin(\omega) - \omega\cos(\omega)}{\sin(\omega)\omega} vH(k) - \frac{\omega}{\sin(\omega)} n\cos(k)\sigma_z, \tag{5.43}$$

$$A(k) = -\frac{2\omega^2}{\sin^2(\omega)} nm\cos(k)\sigma_y, \tag{5.44}$$

and the analytical expression of $X(t)$ is obtained by doubly integrating the acceleration operator $A(t)$. This is eased by the anticommutation between H and A which gives

$$A(k,t) = e^{2iHt}A(k). \tag{5.45}$$

Therefore, integrating the first time we get

$$V(k,t) = \hat{V}(k) + V_Z(k,t), \tag{5.46}$$

$$\hat{V}(k) = -v^2\sigma_z + v\sqrt{1-v^2}\sigma_x, \qquad V_Z(k,t) = \frac{1}{2i}H^{-1}(k)A(k,t), \tag{5.47}$$

with $H^{-1}(k) = \omega^{-2}\,H(k)$, and integrating again one has

$$X(t) = X(0) + \hat{V}t + X_Z(t) - X_Z(0), \tag{5.48}$$

$$X_Z(k,t) = -\frac{1}{4}H^{-2}(k)A(k,t). \tag{5.49}$$

The operator $\hat{V}$ in Eqs. (5.46) and (5.48) is the classical component of the velocity operator which, in the Hamiltonian diagonal basis (5.37), is proportional to the group velocity $v(k)$

$$\hat{V}(k) = v(k)\sigma_z. \tag{5.50}$$

In addition to the classical contribution $\hat{V}t$, we see that the position operator (5.48) presents, as in the usual Dirac theory, a time dependent component $X_Z(t)$ and a constant shift term $X_Z(0)$. Again, it the Hamiltonian (5.37) diagonal basis it is

$$X_Z(t) = \int_{-\pi}^{\pi} \mathrm{d}k\; e^{2i\omega(k)\sigma_z t} Z_X(k) \otimes |k\rangle\langle k|, \tag{5.51}$$

$$X_Z(k) = z(k)\sigma_y, \qquad z(k) = \frac{m\cos\omega(k)}{2\sin^2\omega(k)}. \tag{5.52}$$

Consider now a generic one-particle state and separate its positive and a negative energy components as follows

$$|\psi\rangle = |\psi_+\rangle + |\psi_-\rangle\,. \tag{5.53}$$

Since in the basis $\{|+\rangle\,, |-\rangle\}$ it is $\hat{V} \propto \sigma_z$ (5.50) and $X_Z \propto \sigma_y$ (5.52) we have

$${}_k\langle \mp|\,\hat{V}(k)\,|\pm\rangle_k = 0, \qquad {}_k\langle \pm|\,X_Z(k)\,|\pm\rangle_k = 0, \tag{5.54}$$

and the position operator $X(t)$ (5.48) mean value for the generic state (5.53) can always be written as

$$\langle\psi|\, X(t)\,|\psi\rangle = x^+_\psi(t) + x^-_\psi(t) + x^{\text{int}}_\psi(t) \tag{5.55}$$

$$x^\pm_\psi(t) \coloneqq \langle\psi_\pm|\, X(0) + \hat{V}t\,|\psi_\pm\rangle \tag{5.56}$$

$$x^{\text{int}}_\psi(t) \coloneqq 2\Re[\langle\psi_+|\, X(0) - X_Z(0) + X_Z(t)\,|\psi_-\rangle], \tag{5.57}$$

with $\Re$ denoting the real part.

The first two terms $x^\pm_\psi(t)$ simply correspond to the "classical" evolution of the particle and antiparticle components of the initial state $|\psi\rangle$, which evolve independently according to the classical component (5.50) of the velocity operator. The interference between positive and negative frequencies is responsible for the term $x^{\text{int}}_\psi(t)$ in Eq. (5.57). Obviously in case of $|\psi\rangle$ having only positive or negative component, the interference disappears. The additional term $x^{\text{int}}_\psi(t)$ consists of two contributes: a constant shift and a time dependent term. In the following Proposition we show that the time dependent contribution is an oscillating term that for $t \to \infty$ goes to 0 as $1/\sqrt{t}$, and whose amplitude is bounded by $1/m$—say by the Compton wavelength $\hbar/mc$ in the usual dimensional units. We will also provide a bound for the magnitude of the constant shift term. These results show that $x^{\text{int}}_\psi(t)$ is the automaton analogue of the so called Zitterbewegung:

Proposition 5.2 (Zitterbewegung) *Let $|\psi\rangle$ be a generic one-particle state of the form*

$$\begin{aligned} &|\psi\rangle = |\psi_+\rangle + |\psi_-\rangle\,, \\ &|\psi_\pm\rangle = \int_{-\pi}^{\pi} \frac{\mathrm{d}k}{\sqrt{2\pi}}\, g_\pm(k)\,|\pm\rangle_k\,|k\rangle\,, \qquad g_\pm(k) \in C^\infty_0[-\pi,\pi], \end{aligned} \tag{5.58}$$

with $\langle\psi|\, X(t)\,|\psi\rangle = x^+_\psi(t) + x^-_\psi(t) + x^{\text{int}}_\psi(t)$ the mean position value at time t given in Eq. (5.55)*. The time dependent component of $x^{\text{int}}_\psi(t)$ has the following asymptotic behavior and bounded amplitude*

$$2\Re[\langle\psi_+|\, X_Z(t)\,|\psi_-\rangle] \xrightarrow{t\gg 0} \frac{1}{\sqrt{t}}, \qquad |2\Re[\langle\psi_+|\, X_Z(t)\,|\psi_-\rangle]| \leqslant \frac{1}{m}. \tag{5.59}$$

The constant shift term in $x^{\text{int}}_\psi(t)$ has magnitude bounded as follows

$$|2\Re\,\langle\psi_+|\, X(0) - X_Z(0)\,|\psi_-\rangle\,| \leqslant \frac{1}{m} + \frac{2}{m^2}. \tag{5.60}$$

Proof. Given the state in Eq. (5.58) and using Eqs. (5.51) and (5.52) one has

$$2\Re[\langle\psi_+|\, X_Z(t)\,|\psi_-\rangle] = \int_{-\pi}^{\pi} \frac{\mathrm{d}k}{\pi}\, z(k)\Re\big[ig^*_+(k)g_-(k)e^{2i\omega(k)t}\big],$$

which shows the oscillating nature of the term $2\Re[\langle\psi_+|\, X_Z(t)\,|\psi_-\rangle]$.

One can notice that $z(k) \in L_1[-\pi, \pi]$ for any $m \neq 0$ and $X_Z(k)$ is a bounded operator, hence according to the Riemann-Lebesgue Lemma (see Appendix B) it is

$$\lim_{t\to\infty} 2\Re[\langle\psi_+|\, X_Z(t)\, |\psi_-\rangle] = 0. \tag{5.61}$$

However, we are interested in the analytical dependence on t in the asymptotical limit. Since, for any $m \neq 0$, $\omega(k)$ has three stationary points in $k = 0, \pm\pi$ ($\omega^{(1)}(0) = \omega^{(1)}(\pm\pi) = 0$ and $\omega^{(1)}(k) \neq 0$ elsewhere in the closed interval $k \in [-\pi, \pi]$, with $\omega^{(2)}(0) \neq 0, \omega^{(2)}(\pm\pi) \neq 0$), the stationary phase approximation (see Appendix B) gives

$$2\Re[\langle\psi_+|\, Z_X(t)\, |\psi_-\rangle] \xrightarrow{t\gg 0} \sum_{k=0,\pm\pi} z(k)\Re\left[ig_+^*(k)g_-(k)e^{2i\omega(k)t}\sqrt{\frac{i}{\pi\omega^{(2)}(k)t}}\right], \tag{5.62}$$

showing the first part of (5.59), namely that the term $2\Re[\langle\psi_+|\, X_Z(t)\, |\psi_-\rangle]$ goes to 0 as $1/\sqrt{t}$[1]. The upper bound (5.59) on the oscillation amplitude can be easily evaluated by the following elementary chain of inequalities

$$\begin{aligned} |2\Re[\langle\psi_+|\, X_Z(t)\, |\psi_-\rangle]| &\leqslant 2|X_Z(t)| \leqslant 2|X_Z(0)| \leqslant \\ &\leqslant 2\max_{k\in[-\pi,\pi]} |z(k)| = 2z(0) = \frac{\sqrt{1-m^2}}{m} \leqslant \frac{1}{m}. \end{aligned} \tag{5.63}$$

Similarly one can evaluate an upper bound (5.60) on the constant shift term. First we notice that

$$|2\Re[\langle\psi_+|\, X(0) - X_Z(0)\, |\psi_-\rangle]| \leqslant 2|\,\langle\psi_+|\, X(0)\, |\psi_-\rangle\,| + 2|X_Z(0)|. \tag{5.64}$$

Now defining the $C_0^\infty[-\pi, \pi]$ test function $\varphi(k, k') = g_+^*(k)g_-(k')\,\langle +|_k\, |-\rangle_{k'}$, we have

$$\begin{aligned} |\,\langle\psi_+|\, X(0)\, |\psi_-\rangle\,| &= \left|\left\langle \frac{\mathrm{d}\delta(k-k')}{\mathrm{d}(k-k')}\middle|\varphi(k,k')\right\rangle\right| = \left|\left\langle \delta(k-k')\middle|\frac{\mathrm{d}\varphi(k,k')}{\mathrm{d}(k-k')}\right\rangle\right| = \\ &= \left|\int_{-\pi}^{\pi} \frac{\mathrm{d}k\,\mathrm{d}k'}{2\pi}\,\delta(k-k')g_+^*(k)g_-(k')\frac{\mathrm{d}}{\mathrm{d}(k-k')}\,\langle +|_k\, |-\rangle_{k'}\right| = \\ &= \left|\int_{-\pi}^{\pi} \frac{\mathrm{d}k}{2\pi}\, g_+^*(k)g_-(k)f(k)\right| \leqslant \\ &\leqslant \max_{k\in[-\pi,\pi]} |f(k)| = f(0) = \frac{\sqrt{1-m^2}}{m^2} \leqslant \frac{1}{m^2} \\ f(k) &:= \frac{n}{\sin^2\omega}. \end{aligned} \tag{5.65}$$

[1]Notice that this asymptotic beahaviour could be different in the presence of a potential as shown in Refs. [157,158], where a revival of the Zitterbewegung occurs for a Dirac oscillator.

Finally, exploiting Eqs. (5.65) and (5.63) in Eq. (5.64) we get the upper bound (5.60). ■

Proposition 5.2 allows to state that $x_\psi^{\text{int}}(t)$ in Eq. (5.57) is the automaton analogue of the Zitterbewegung. The results of this Section are in agreement with the ones for the Hadamard walk in Ref. [159].

In Fig. 5.4 we compare our analytical results with some simulations of the automaton evolution. In the simulations we have considered states (5.58) with particle and antiparticle components smoothly peaked around some momentum eigenstate (see Definition 5.1), namely

$$c_+ |\psi_+\rangle + c_- |\psi_-\rangle\,, \quad |\psi_\pm\rangle = \int \frac{\mathrm{d}k}{\sqrt{2\pi}}\, g_{k_0}(k)\, |\pm\rangle_k\, |k\rangle\,, \qquad |c_+|^2 + |c_-|^2 = 1, \tag{5.66}$$

where g_{k_0} is a Gaussian peaked around k_0 with width σ. An easy computation shows that for these states the shift contribution reduces to

$$2\Re[\langle\psi|\, X(0) + Z_X(0)\, |\psi\rangle] = \Im(c_+^* c_-)/(2\pi) \int_{-\pi}^{\pi} \mathrm{d}k\ |g_{k_0}(k)|^2 z(k), \tag{5.67}$$

with the function $z(k) = m \cos\omega(k)/\sin^2\omega(k)$ bounded again by the Compton wavelength $1/m$. From proposition 5.2 we know that in the asymptotic limit $t \gg 0$ the oscillation has frequency $\omega(0)/\pi$ and amplitude damps from $z(0)$ to zero. On the other hand for small times t it becomes relevant the frequency $\omega(k_0)/\pi$ and the amplitude $z(k_0)$, corresponding to the momentum k_0 of the state. In any case the functions z and ω/π determine the oscillation amplitude and frequency of the position expectation value (see also Fig. 5.3). For example, when the wave-packets are both peaked around $k = 0$, as in the first two cases of Fig. 5.4, the damping of the oscillation amplitude can be observed for times much longer than the ones we could consider in our simulations. Indeed, since the wave-packet is sharp in $k = 0$, the asymptotic approximation of Eq. (5.62) results to be accurate only for very large values of t.

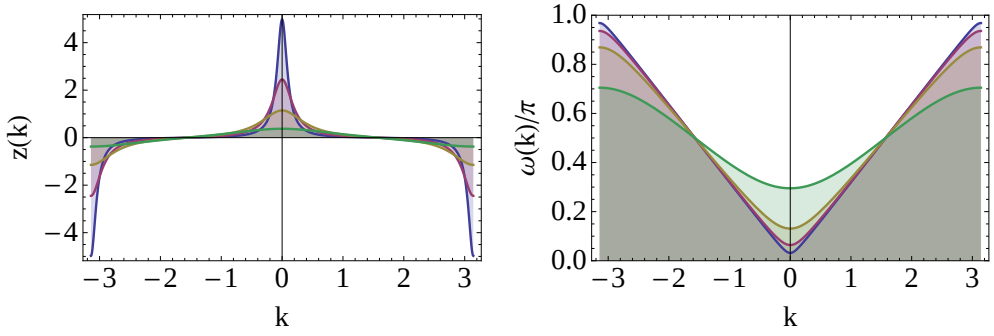


Figure 5.3: Plots of $z(k)$ (left) and $\omega(k)/\pi$ (right) related to the oscillation amplitude and frequency of the position expectation value in Eq. (5.48). In both cases the plots are reported for different values of the mass ($m = 0.1, 0.2, 0.4, 0.8$ from the top in the left figure and from the bottom in the right figure).

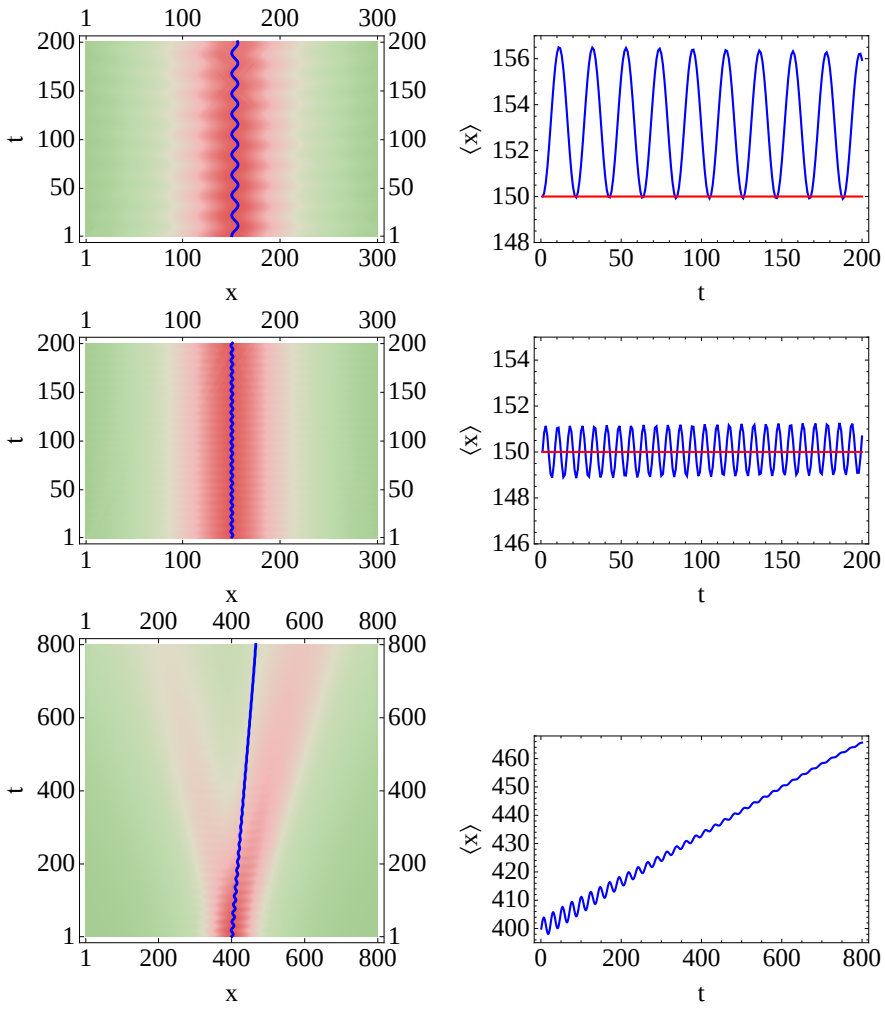


Figure 5.4: Automaton evolution of a state as in Eq. (5.66) showing the Zitterbewegung of the position mean value. **Top:** $m = 0.15$, $c_+ = 1/\sqrt{2}$, $c_- = i/\sqrt{2}$, $k_0 = 0$, and $\sigma = 40^{-1}$. The calculated shift and oscillation frequency are respectively $\langle\psi|\, X(0)+Z_X(0)\,|\psi\rangle = 3.2$ and $\omega(0)/\pi = 0.05$, compatibly with the simulation. **Middle:** $m = 0.15$, $c_+ = 1/\sqrt{2}$, $c_- = 1/\sqrt{2}$, $k_0 = 0$, $\sigma = 40^{-1}$. The calculated shift and oscillation frequency are 0 and 0.13, respectively. **Bottom:** $m = 0.13$, $c_+ = \sqrt{2/3}$, $c_- = 1/\sqrt{3}$, $k_0 = 10^{-2}\pi$, $\sigma = 40^{-1}$. In this case the particle and antiparticle contributions are unbalanced and the drift velocity of the mean position is $\langle\psi_+|\,\hat{V}\,|\psi_+\rangle + \langle\psi_-|\,\hat{V}\,|\psi_-\rangle = (|c_+|^2 - |c_-|^2)v(k_0) = 0.08$, corresponding to a mean position $x^+_\psi(800)+x^-_\psi(800) = 464$ at $t = 800$ (see Eq. (5.56)). Notice that for $t \to \infty$ the term $2\Re[\langle\psi_+|\, Z_X(t)\,|\psi_-\rangle$, which is responsible of the oscillation, goes to 0.

Remark 5.2 (Newton-Wigner position operator evolution) As in the usual QFT the Newton-Wigner position operator (5.40) does not suffer the jittering of the mean position even for states having both a particle and an antiparticle component. Indeed, in this case, the velocity operator is

$$V(t) = [H, X_{NW}(t)], \qquad V(k) = \hat{V}(k), \tag{5.68}$$

corresponding to the classical component (5.50) of the velocity operator in Eq. (5.46) and leading to a null acceleration

$$A(t) = [H, \hat{V}(t)] = 0. \tag{5.69}$$

By integrating (5.68) we see that the position operator $X_{NW}(t)$ is simply

$$X_{NW}(t) = X_{NW}(0) + \hat{V}t. \tag{5.70}$$

5.3 Scattering with a square potential barrier

In this Section we add a position-dependent potential to the Dirac automaton (4.31) evolution and study the scattering with a square potential barrier. We will derive explicitly the transmission T and reflection R coefficients as functions of the energy and mass of the incident wave-packet and of the potential barrier's height. We will find a general behavior independently on the energy and mass of the incident particle: increasing the value of the potential barrier beyond a certain threshold a transmitted wave reappears and the reflection coefficient starts decreasing. The width of the $R = 1$ region is an increasing function of the mass which is proportional to the gap between the positive- and the negative-frequency eigenvalues of the automaton unitary evolution.

The Dirac unitary matrix (4.31) can be written as follows

$$U := \sum_x \begin{pmatrix} n|x-1\rangle\langle x| & -im|x\rangle\langle x| \\ -im|x\rangle\langle x| & n|x+1\rangle\langle x| \end{pmatrix}, \tag{5.71}$$

describing a *Quantum Walk* on the Hilbert space $\mathbb{C}^2 \otimes l_2(\mathbb{Z})$ as noticed in Section 4.16. Since the "energy" of the Dirac field is represented by the automaton dispersion relation ω, introducing a potential in the evolution simply means to modify the dispersion relation of the automaton. In the free case the dispersion relation ω appears in the phases $e^{\pm i\omega}$ corresponding to the Dirac automaton eigenvalues (see Section 4.3.2), and we can introduce a potential as an additional phase $e^{i\phi}$. Hence, for a generic potential $\phi(x)$, the unitary evolution (5.71) becomes

$$U_\phi := \sum_x e^{-i\phi(x)} \begin{pmatrix} n|x-1\rangle\langle x| & -im|x\rangle\langle x| \\ -im|x\rangle\langle x| & n|x+1\rangle\langle x| \end{pmatrix}, \tag{5.72}$$

as also proposed in Refs. [125, 159].

Here we analyze the simple case

$$\phi(x) := \phi\,\theta(x) \qquad \theta(x) = \begin{cases} 0 & x < 0, \\ 1 & x \geqslant 0, \end{cases} \tag{5.73}$$

($\theta(x)$ is the Heaviside step function) that is a potential step which is 0 for $x < 0$ (region I) and has a constant value $\phi \in [0, 2\pi]$ for $x \geqslant 0$ (region II) as illustrated in the left Fig. 5.5.

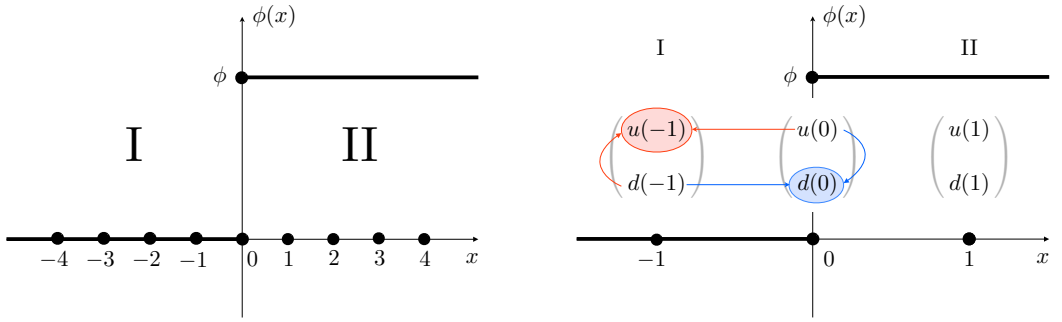


Figure 5.5: **Left Figure:** Schematic of the square potential in Eq. (5.73) . **Right Figure:** Schematic of the boundary conditions at the junction between the regions I and II where the square potential (5.73) turns on.

A plane wave eigenstate of the Dirac QCA without the potential barrier has the form

$$|\Phi^{(\omega)}\rangle = |s\rangle_k |k\rangle = e^{ikx} |s\rangle_k |x\rangle , \qquad U|\Phi^{(\omega)}\rangle = e^{-is\omega(k)}|\Phi^{(\omega)}\rangle, \tag{5.74}$$

where the vectors

$$|s\rangle_k = \begin{pmatrix} u_k^s \\ d_k^s \end{pmatrix} \qquad u_k^s = 2^{-1/2}\sqrt{1-sv}, \quad d_k^s = 2^{-1/2}s\sqrt{1+sv}, \tag{5.75}$$

are the Dirac QCA momentum eigenstates $U(k)\,|s\rangle_k = e^{-is\omega}\,|s\rangle_k$ derived in Section 4.3.2.

When the potential barrier (5.73) is turned on, the plane wave eigenstates of the automaton U_ϕ are no longer of the simple form (5.74). Let us study the eigenstates of U_ϕ of the form

$$|\Phi^{(\omega)}\rangle = \Pi_{\rm I}\,|+\rangle_k |k\rangle + \Pi_{\rm I}\beta_k\,|+\rangle_{-k} |k\rangle + \gamma_k \Pi_{\rm II}\,|+\rangle_{k'} |k'\rangle , \tag{5.76}$$

$$\Pi_{\rm I} := \sum_{x<0} I \otimes |x\rangle\langle x|, \qquad \Pi_{\rm II} := \sum_{x\geqslant 0} I \otimes |x\rangle\langle x|, \tag{5.77}$$

where β_k, γ_k and k' are functions of k. The condition that $|\Phi_k^{(\omega)}\rangle$ is an eigenstate of U_ϕ, *i.e.* $U_\phi|\Phi_k^{(\omega)}\rangle = e^{-i\omega(k)}$, implies

$$\omega(k') = \omega(k) - \phi. \tag{5.78}$$

Physically in region I we have an incident wave and a reflected wave

$$|+\rangle_k |k\rangle + \beta_k |+\rangle_{-k} |k\rangle , \tag{5.79}$$

which must have the same momentum k with opposite sign. In region II we have the transmitted wave

$$\gamma_k |+\rangle_{k'} |k'\rangle , \tag{5.80}$$

whose momentum k' depends on the height of the potential barrier according to Eq. (5.78).

The derivation of the coefficients β_k, γ_k requires the boundary conditions at site $x = 0$ separating the regions I and II. When the potential barrier (5.73) is turned on any plane wave eigenstate $\left|\Phi^{(\omega)}\right\rangle$ of the automaton U_ϕ (5.72) having energy ω must satisfy the following boundary conditions (see also right Fig. 5.5)

$$e^{-i\omega}u(-1) = ne{-}i\phi u(0) - imd(\omega)(-1), \tag{5.81}$$
$$e^{-i\omega}d(0) = nd(-1) - ime^{-i\phi}u(0), \tag{5.82}$$

as one can easily check applying the automaton (5.72) to the wave $\left|\Phi^{(\omega)}\right\rangle$. Using the expression (5.76) for the eigenstate $\left|\Phi^{(\omega)}\right\rangle$, the boundary conditions (5.81,5.82) become

$$e^{-i\omega(k)}(e^{-ik}u_k^+ + \beta_k e^{ik}u_{-k}^+) = ne^{-i\phi}\gamma_k u_{k'}^+ - im(e^{-ik}d_k^+ + \beta_k e^{-ik}d_{-k}^+), \tag{5.83}$$
$$e^{-i\omega(k)}\gamma_k d_{k'}^+ = n(e^{-ik}d_k^+ + \beta_k e^{ik}d_{-k}^+) - ime^{-i\phi}\gamma_k u_{k'}^+. \tag{5.84}$$

Now we can use the identity $U(k)\left|+\right\rangle = e^{-i\omega}\left|+\right\rangle$ (see Eq. (4.38) for $U(k)$) to simplify (5.83,5.84) as follows

$$u_k^+ + \beta_k u_{-k}^+ = e^{-i\phi}\gamma_k u_{k'}^+, \tag{5.85}$$
$$e^{-ik}d_k^+ + \beta_k e^{ik}d_{-k}^+ = e^{-i\phi}e^{-ik'}\gamma_k d_{k'}^+, \tag{5.86}$$

and by an easy manipulation we get

$$\beta_k = \frac{e^{-ik}\sqrt{(1+v)(1-v')} - e^{-ik'}\sqrt{(1-v)(1+v')}}{-e^{ik}\sqrt{(1-v)(1-v')} + e^{-ik'}\sqrt{(1+v)(1+v')}},$$
$$\gamma_k = \frac{2e^{i\xi}(v\cos k - i\sin k)}{-e^{ik}\sqrt{(1-v)(1-v'))} + e^{-ik'}\sqrt{(1+v)(1+v')}},$$

with $v := v(k)$ and $v' := v(k')$ the group velocities of the incident and transmitted wave.

The states in Eq. (5.76) are the equivalent of the plane waves for the free case and are not normalized. Let us consider smooth-states (see Definition 5.1) given by the superposition

$$\left|\Psi(0)\right\rangle := \int \frac{\mathrm{d}k}{\sqrt{2\pi}}\, g_{k_0}(k)\left|\Phi_k\right\rangle,$$

where $g_{k_0}(k)$ is a function in $C_0^\infty[-\pi,\pi]$ and is assumed to be smoothly peaked around k_0. The state at time t is then

$$\left|\Psi(t)\right\rangle := \int \frac{\mathrm{d}k}{\sqrt{2\pi}}\, g_{k_0}(k)e^{-i\omega(k)t}\left|\Phi_k\right\rangle,$$

and one can verify that for $t \ll 0$ the state is negligible in region II while the only appreciable contribution in region I comes from the term $e^{ik_0 x}$, describing a wave-packet that moves at the group velocity $v(k_0)$ and hits the barrier form the left. When $t \gg 0$ the state can be approximated by a superposition of a reflected and a transmitted wave-packet as follows

$$|\Psi(t)\rangle \xrightarrow{t\gg 0} \beta(k_0) \int \frac{\mathrm{d}k}{\sqrt{2\pi}}\, g_{k_0}(k) e^{-i\omega(k)t} |+\rangle_{-k} |k\rangle + \\ + \tilde{\gamma}(k_0) e^{-i\phi t} \int \frac{\mathrm{d}k}{\sqrt{2\pi}}\, \tilde{g}_{k_0'}(k') e^{-i\omega(k')t} |+\rangle_{k'} |k'\rangle ,$$

where we defined

$$k_0' \text{ s.t. } \omega(k_0') = \omega(k_0) - \phi, \\ \tilde{\gamma}(k_0) \coloneqq \gamma(k_0) \sqrt{\frac{v(k_0')}{v(k_0)}}, \qquad \tilde{g}_{k_0'}(k') \coloneqq \sqrt{\frac{v(k_0')}{v(k_0)}} g_{k_0'}(k')$$

(one can check $\int \frac{\mathrm{d}k}{\sqrt{2\pi}} |\tilde{g}_{k_0'}(k')|^2 = 1$), whose group velocities are $-v(k_0)$ for the reflected wave-packet and $v(k_0')$ for the transmitted one (see Fig. 5.7).

We denote by R (reflection coefficient) and T (trasmission coefficient) the probability of finding the particle respectively in the reflected and transmitted wave-packet, namely

$$R = |\beta(k_0)|^2 = \frac{1 - vv' - \cos(k-k')\sqrt{(1-v^2)(1-v'^2)}}{1 + vv' - \cos(k+k')\sqrt{(1-v^2)(1-v'^2)}}, \tag{5.87}$$

$$T = |\tilde{\gamma}(k_0)|^2 = \frac{2(v^2\cos^2(k) + \sin^2(k))}{1 + vv' - \cos(k+k')\sqrt{(1-v^2)(1-v'^2)}}, \tag{5.88}$$

$$R + T = 1, \tag{5.89}$$

where $R + T = 1$ verifies the consistency of the result.

For $k \ll m \ll 1$ (see the non relativistic regime of the Dirac QCA in Eq. (5.25)) we recover the usual reflection and transmission coefficient for the Schröedinger equation with a step potential barrier.

5.3.1 The Klein paradox

In Fig. 5.6 we plot the reflection coefficient R as a function of ϕ and k for different values of the mass m. Clearly when the height of the potential barrier is $\phi = 0$ we have $R = 0$ and, increasing ϕ while keeping k fixed, the value increases up to $R = 1$

$$\phi = 0 \Rightarrow R = 0,\ T = 1, \tag{5.90}$$

$$0 < \phi < \omega(k) - \arccos(n) \Rightarrow R > 0,\ T > 0, \tag{5.91}$$

$$\phi = \omega(k) - \arccos(n) \Rightarrow R = 1,\ T = 0. \tag{5.92}$$

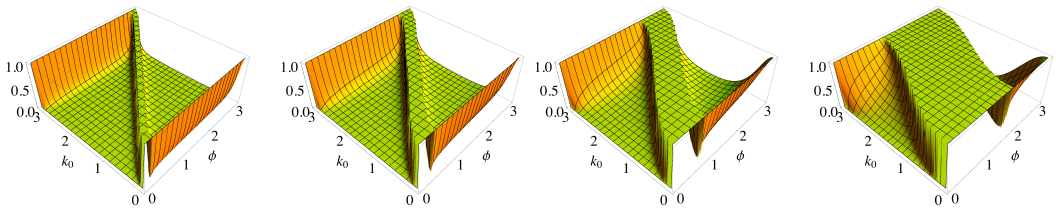


Figure 5.6: Reflection coefficient as a function of the potential barrier height ϕ and of the momentum k of the incident particle state. From the left to the right the reflection coefficient is depicted for different values of the mass: $m = 0.1,\ 0.2,\ 0.4,\ 0.8$.

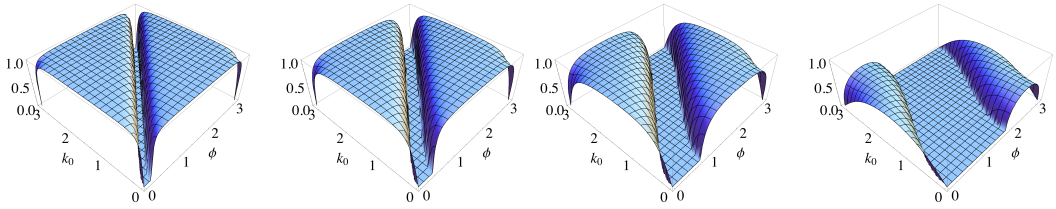


Figure 5.7: Group velocity of the transmitted wave-packet as a function of the potential barrier height ϕ and of the momentum k of the incident particle state. From the left to the right the transmitted group velocity for different values of the mass: $m = 0.1,\ 0.2,\ 0.4,\ 0.8$.

One notice that when $\omega(k) - \arccos(n) < \phi < \omega(k) + \arccos(n)$ Eq. (5.78) has solution for imaginary k' which implies an exponential damping of the transmitted wave corresponding to a pure reflection

$$\omega(k) - \arccos(n) < \phi < \omega(k) + \arccos(n) \Rightarrow R = 1,\ T = 0. \tag{5.93}$$

By further increasing the value of ϕ beyond the threshold $\omega(k) + \arccos(n)$ Eq. (5.78) has solution for real k' and negative $\omega(k')$, and then a transmitted wave reappears with the reflection coefficient decreasing

$$\phi > \omega(k) + \arccos(n) \Rightarrow R > 0,\ T > 0. \tag{5.94}$$

The revival of the transmitted wave for an high enough potential barrier is the so called "Klein paradox" which is originated by the presence of positive- and negative-frequency eigenvalues of the unitary evolution. Usually it is interpreted as the generation of an antiparticle with negative momentum which moves in the direction of the transmitted wave. In the automaton context the interpretation is very simple since Eq. (5.78) has solution for real $k' < 0$ and negative $\omega(k') < 0$, which gives rise to a wave having group velocity $v(k') > 0$ (remember that the group velocity corresponding to negative $\omega(k)$ is positive when $k < 0$ as shown in Fig. 4.9).

The width of the $R = 1$ region (see Fig 5.6) is an increasing function of the mass and is equal to $2\arccos(n)$, namely

$$-\arccos(n) < \phi < \arccos(n), \tag{5.95}$$

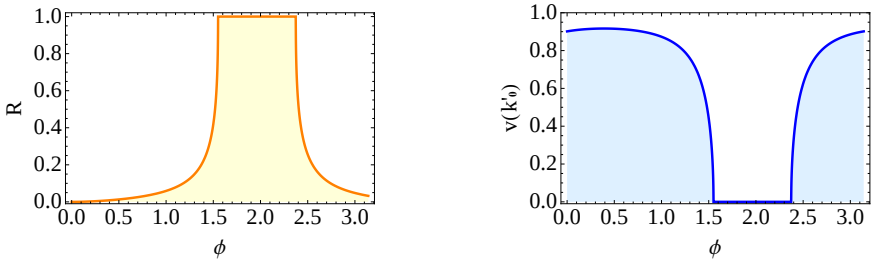


Figure 5.8: Reflection coefficient (left) and transmitted-wave group velocity (right) for $m = 0.4$ and momentum of the incident particle $k_0 = 2$ as functions of the potential barrier height ϕ (section of plots in Figs. 5.6 and 5.7 for $m = 0.4$, $k_0 = 2$).

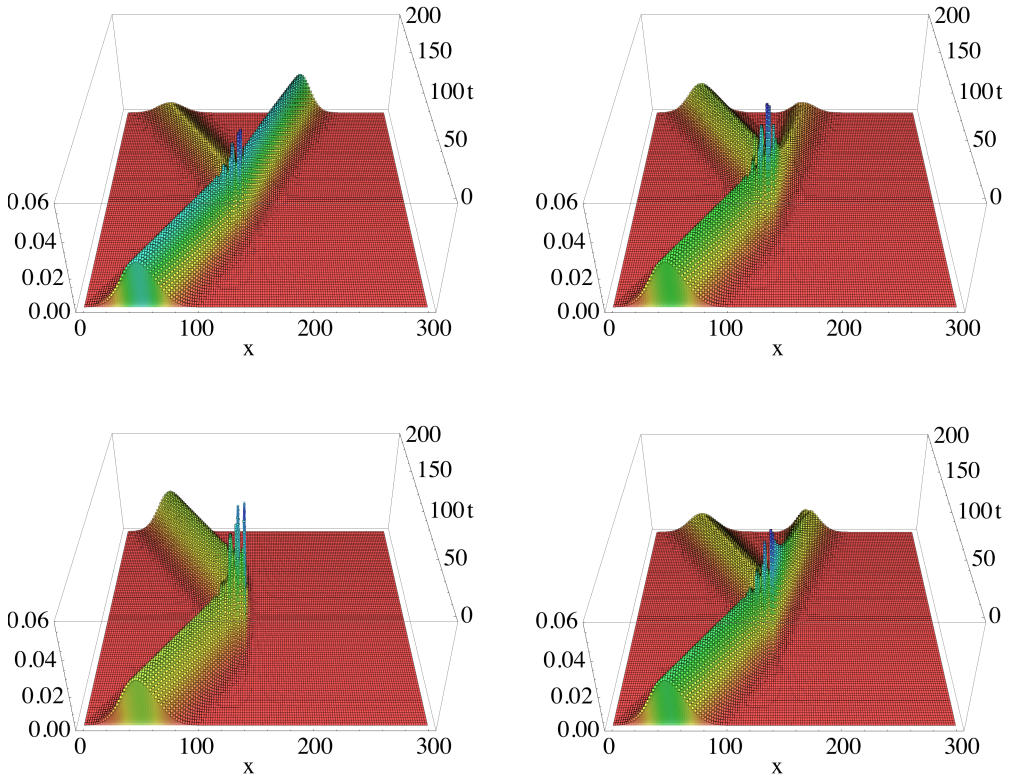


Figure 5.9: Simulations of the Dirac automaton evolution with a square potential barrier. Here the automaton mass is $m = 0.2$ while the barrier turns on at $x = 140$. In the simulation the incident state is a smooth-state of the form $|\psi(0)\rangle = \int \frac{\mathrm{d}k}{\sqrt{2\pi}}\, g_{k_0}(k)\, |+\rangle_k$ peaked around the positive-energy eigenstate $|+\rangle_{k_0}$ with $k_0 = 2$ and with g_{k_0} a Gaussian having width $\sigma = 15^{-1}$. The incident group velocity is $v(k_0) = 0.90$. The simulation is run for four increasing values of the potential ϕ. **Top-Left:** Potential barrier height $\phi = 1.42$, reflection coefficient $R = 0.25$, velocity of the transmitted particle $v(k_0') = 0.63$. **Top-Right:** $\phi = 1.55$, $R = 0.75$, $v(k_0') = 0.1$. **Bottom-Left:** $\phi = 2$, $R = 0.1$, $v(k_0') = 0$. **Bottom-Right:** $\phi = 2.4$, $R = 0.50$, $v(k_0') = 0.33$.

which is the gap between positive and negative-frequency solutions (see Fig. 4.7).

In Fig. 5.8 we plot the reflection coefficient R and the transmitted-wave velocity group $v(k_0')$ as a function of the potential barrier height ϕ with the incident wave-packet having $k_0 = 2$ and $m = 0.4$. From the figure it is clear that after a plateau with $R = 1$ the reflection coefficient starts decreasing for higher potentials. Finally In Fig. 5.9 we show the scattering simulation for four increasing values of the potential, say $\phi = 1.42, 1.55, 2, 2.4$ (see the caption to figure for the details).

CHAPTER 6

Discriminating the Dirac automaton and the usual Dirac evolutions

A major aim of the automaton approach to QFT is the connection to the phenomenology, namely to some experiment that could discriminate between the predictions of the automaton model and the usual QFT ones.

In Chapter 4 we have shown how the Dirac automaton recovers the Dirac equation in the large-scale limit of small masses and momenta. Here we make this notion more rigorous comparing the Dirac unitary operator with the automaton one in a quantum informational scenario, evaluating the probability of discriminating between the two evolutions in a generic experiment. The discrimination probability will be given in terms of the parameters of the experiment, namely the number of particles involved, their masses and their momenta. Finally we discuss possible ways of testing of the theory along with effects of the Planck scale discreteness on wave-packets fly-times.

The content of this Chapter is based on Ref. [59].

6.1 Theoretically optimal discrimination

The problem here is to experimentally discriminating between two different dynamics, the Dirac automaton and the usual Dirac evolutions. In quantum informational language this is a problem of discriminating optimally between

two different unitaries and is usually refereed to as "discrimination between two black boxes".

6.1.1 Discrimination between two black boxes

We can consider the following scenario. An experimentalist is given a black box that can be either the box A or the box D with equal probability $\frac{1}{2}$, and he is asked to guess which box. Without loss of generality we can include the processing of the experimental data in the experiment itself, therefore the box discrimination is a two-outcome experiment, with outcomes A and D: if the outcome is A the experimentalist will guess A and if the outcome is D he will guess D. In general the correct answer can be given with some probability and the challenge for the expermentalist is to design the experiment which minimizes the probablity of error (wrong guess)

$$p_e = \tfrac{1}{2}\left[p(A|D) + p(D|A)\right], \tag{6.1}$$

where $p(X|Y)$ is the probability of getting outcome X when the black box is Y. We want now to minimize p_e over all possible experiments. It is clear from this scenario that the minimum over all the possible experiments of the error probability is a well defined measure of how much the models A and D are far apart.

The optimal discrimination between the two black boxes is a special case of discrimination between quantum channels (see see Section 2.2.1 for the definition of quantum channel) , namely the discrimination between the two unitary evolutions described by U_A and U_D.

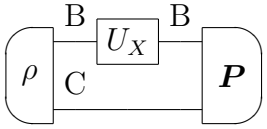
Let us begin our analysis by describing the most general experiment which discriminates between two unitary evolutions that act on a quantum system B with Hilbert space $\mathcal{H}_{\mathrm{B}}$. The experimental procedure can be divided into three steps:

(i) Prepare a quantum state ρ possibly entangled with an ancillary system C with Hilber space $\mathcal{H}_{\mathrm{C}}$.

(ii) Apply the unknown unitary evolution U_X ($X = A, D$) to system $\mathcal{H}_{\mathrm{B}}$ (the ancillary system evolves with the identity map).

(iii) Perform a two-outcome measurement on system $\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{C}}$: the two outcomes A and D correspond to the two possible evolutions. The measurement can be described by a POVM $\boldsymbol{P} = \{P_A, P_D\}$, where P_A and P_D are positive operators on $\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{C}}$ which satisfy $P_A + P_D = I$, I denoting the identity on $\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{C}}$ (see Section 2.2.1 for the definition of POVM).

We can pictorially represent the above procedure as follows:



. (6.2)

The probability of error in Eq. (6.1) becomes then

$$p_e = \frac{1}{2}\,\mathrm{Tr}[P_A(U_D\rho U_D{}^\dagger - U_A\rho U_A{}^\dagger)], \tag{6.3}$$

where we have used the trivial identity $P_D = I - P_A$.

Minimizing expression (6.3) over all possible experiments becomes a minimization over the set of the POVMs and the set of the available states. According to the old Helstrom's result [160] the minimization over the POVM set gives

$$\inf_{0\leqslant P_A\leqslant I} p_e = \frac{1}{2} - \frac{1}{2}||U_D\rho U_D{}^\dagger - U_A\rho U_A{}^\dagger)||_1 \tag{6.4}$$

where $||\cdot||_1$ denotes the trace norm[1]. Clearly in order to have the minimal error probability p_e we have now to maximize the term $||U_D\rho U_D{}^\dagger - U_A\rho U_A{}^\dagger)||_1$ in Eq. (6.4) over all states ρ. Generally, the optimal discrimination between quantum channels needs entangled states [161], namely states ρ_{BC} (see the circuit in Eq. (6.2)) where the system B is entangled with the ancilla C . On the other hand for the case of two unitaries with a single copy of each box, the optimal error probability $\bar{p}_e$ is simply given by

$$\bar{p}_e = \tfrac{1}{2} - \tfrac{1}{2}\|U_A - U_D\| = \tfrac{1}{2} - \tfrac{1}{2}\sup_{\rho\in\mathsf{St}(\mathcal{H}_{\mathrm{B}})} \|U_A\rho U_A^\dagger - U_D\rho U_D^\dagger\|_1, \tag{6.5}$$

where the supremum is taken only on the system B.

6.1.2 The Dirac black boxes case

We are interested in discriminating between the unitary evolution of the Dirac automaton and the unitary evolution of the Dirac equation. Clearly the evolution will involve an arbitrary state of the Dirac field ψ as discussed in Chapter 4. Since the discrimination experiment can have a generic duration t, the operators to be discriminated are the unitaries at time t. By direct computation using the Hamiltonians in Eqs. (4.38) and (4.45) we find

$$U^t = e^{-iH(k)t} = \begin{pmatrix} \cos(\omega t) + i\frac{\sin(\omega t)}{\omega}a & -ib\frac{\sin(\omega t)}{\omega} \\ -ib\frac{\sin(\omega t)}{\omega} & \cos(\omega t) - i\frac{\sin(\omega t)}{\omega}a \end{pmatrix} \tag{6.6}$$
$$a := \frac{\omega}{\sin(\omega)} n\sin(k) \qquad b := \frac{\omega}{\sin(\omega)} m$$

$$U_D^t = e^{-iH_D(k)t} = \begin{pmatrix} \cos(\omega_D t) + i\frac{\sin(\omega_D t)}{\omega_{\mathrm{D}}}k & -im\frac{\sin(\omega_D t)}{\omega_D} \\ -im\frac{\sin(\omega_D t)}{\omega_D} & \cos(\omega_D t) - i\frac{\sin(\omega_D t)}{\omega_D}k \end{pmatrix} \tag{6.7}$$

[1]This is the operational norm defined in Eq. (2.16) for a general probabilistic theory which in the QT case becomes the trace norm of (2.26).

where ω and ω_D are the automaton (4.41) and the usual Dirac (4.46) dispersion relations.

We have now to evaluate the minimal probability of error in Eq. (6.5). The supremum is taken over all possible experiments involving the two unitary evolutions (6.6) and (6.7), namely over all possible states $\rho \in \mathsf{St}(\mathcal{H}_{\mathrm{B}})$ of the Dirac field ψ. These states are characterized by three parameters:

(i) The number N of particles involved in the experiment.

(ii) The momenta k of the particles involved in the experiment.

(iii) The masses m of the particles involved in the experiment.

For large dimension for the Hilbert space $\mathcal{H}_{\mathrm{B}}$ and with no restriction on the states of the Dirac field, the probability (6.5) would be vanishingly small. However, in a typical physical scenario the available states for the experimentalist are limited, namely the particles involved in the experimental set up are limited both in number and in momentum. We thus consider the physical bounds on the number of particles $N \leqslant \bar{N} < \infty$ and their momenta $k \leqslant \bar{k} < \pi$ according to the following definition:

Definition 6.1 (Set of physical restricted states $\mathcal{T}_{\bar{k},\bar{N}}$) *By $\mathcal{T}_{\bar{k},\bar{N}}$ we denote the set of physically restricted states $\rho \in \mathsf{St}(\mathcal{H}_{\mathrm{B}})$ such that*

$$\mathrm{Tr}[\rho N_{\bar{k}}] = \mathrm{Tr}[\rho P_{\bar{N}}] = 0, \tag{6.8}$$

where $P_{\bar{N}}$ is the projector on the $N > \bar{N}$-particles sector and $N_{\bar{k}}$ is the operator that counts the number of particles with momentum $|k| > \bar{k}$, i.e $N_{\bar{k}} = \int_{|k|>\bar{k}} \mathrm{d}k\ \psi^{\dagger}(k)\psi(k)$.

Within the restricted set of sates the optimal probability of error in Eq. (6.9) becomes

$$\bar{p}_e = \tfrac{1}{2} - \tfrac{1}{2} \sup_{\rho \in \mathcal{T}_{\bar{k},\bar{N}}} \|U^t \rho U^{t\dagger} - U_D^t \rho U_D^{t\,\dagger}\|_1. \tag{6.9}$$

The explicit computation of $\bar{p}_e$ cannot be done analyticly. However, we are able to derive an analytic lower bound for Eq. (6.9) which is the main result of this Section.

Proposition 6.1 (Lower bound for the probability of error (6.9)) *The optimal probability of error (6.9) in discriminating between the Dirac automaton (6.6) and the Dirac equation (6.7) unitary evolutions can be lower bounded as follows*

$$\bar{p}_e \geqslant \tfrac{1}{2} - \tfrac{1}{2}\sqrt{1 - \cos^2(g(\bar{k}, m, \bar{N}, t))}, \tag{6.10}$$

where

$$\begin{aligned} g(\bar{k}, m, \bar{N}, t) &:= \bar{N} \arccos\left(\cos(\bar{\alpha} t) - \bar{\beta}\right), \\ \bar{\alpha} &:= \max_{k \in \{0,\bar{k}\}} |\omega_D - \omega|, \\ \bar{\beta} &:= \max_{k \in \{0,\bar{k}\}} \left| \frac{1}{2}\left(1 - v v_D - \sqrt{(1-v^2)(1-v_D^2)}\right)\right|, \end{aligned} \tag{6.11}$$

with ω and ω_D the automaton (4.41) *and the Dirac* (4.46) *dispersion relation, and $v = \partial_k \omega$ and $v_D = \partial_k \omega_D$ the automaton* (5.16) *and the Dirac* (5.22) *group velocity.*

We stress that the bound (6.10) is very general and can be trivially computed for any kind of experiment having a certain duration t and involving $\bar{N}$ particles having mass m and momentum $\bar{k}$. The proof of the lower bound (6.10) goes through three Lemmas and one Proposition. The reader not interested in the technical details of the proof can directly skip to the next Section.

Lemma 6.1 *Let $U_D^t(k)$ and $U^t(k)$ be defined according to Eqs.* (6.6) *and* (6.7)*. Upon defining $V(k,t) = U_D^t(k) U^{t\dagger}(k)$, let $e^{i\mu(k,m,t)}$ be an eigenvalue of $V(k,t)$. Then the following bound holds:*

$$\cos(\mu(k,m,t)) \geqslant \cos(\alpha t) - \beta \tag{6.12}$$

where

$$\begin{aligned} \alpha(k,m) &:= \omega_{\mathrm{D}} - \omega \\ \beta(k,m) &:= \frac{1}{2}\left(1 - v v_{\mathrm{D}} - \sqrt{(1-v^2)(1-v_{\mathrm{D}}^2)}\right). \end{aligned} \tag{6.13}$$

Proof. Since both $U_D^t(k)$ and $U^{t\dagger}(k)$ are $SU(2)$ matrices, we have that $V(k,t)$ is an $SU(2)$ matrix and its eigenvalues must be of the form $e^{i\mu(k,m,t)}$ and $e^{-i\mu(k,m,t)}$. This implies the equality $\cos(\mu(k,m,t)) = \frac{1}{2}\mathrm{Tr}[V(k,t)]$ which by direct computation gives

$$\cos(\mu(k,m,t)) = \left(1 - \frac{\beta}{2}\right)\cos(\alpha t) + \frac{\beta}{2}\cos(\gamma t) \tag{6.14}$$

where α and β are defined as in Eq. (6.13) and $\gamma := \omega + \omega_{\mathrm{D}}$. Finally, from Eq. (6.14) one can easily find the bound $\cos(\mu(k,m,t)) \geqslant \cos(\alpha t) - \beta$ ■

The second Lemma shows the monotonicity of the two functions α, β in Lemma 6.1:

Lemma 6.2 *Let $\alpha(k,m)$ and $\beta(k,m)$ be defined as in Eq.* (6.13) *and $0 \leqslant \bar{k} < \pi$. Then we have*

$$\begin{aligned} \bar{\alpha} &:= \max_{k\in[-\bar{k},\bar{k}]} |\alpha| = \max_{k\in\{0,\bar{k}\}} |\alpha| \\ \bar{\beta} &:= \max_{k\in[-\bar{k},\bar{k}]} |\beta| = \max_{k\in\{0,\bar{k}\}} |\beta| \end{aligned} \qquad \forall m \in [0,1]\,. \tag{6.15}$$

Proof. Since both ω and ω_{D} are even function of k, from Eq. (6.13) we have that also α and β are even function of k. For this reason we can restrict to $k \in [0, \bar{k}]$. The equality (6.15) can be proved by showing that α and β are nondecreasing functions of k for $k \in [0, \bar{k}]$.

Since $\partial_k \alpha = v_{\mathrm{D}} - v$, if $v_{\mathrm{D}}^2 - v^2 \geqslant 0$ for $k \in [0, \pi)$ then $\partial_k \alpha \geqslant 0$ in the same interval. By direct computation one can verify that

$$(v_{\mathrm{D}})^2 - (v)^2 = \frac{x(k,m)}{y(k,m)} \tag{6.16}$$

$$x(k,m) := k^2 - \sin^2(k)(1-m^2) \tag{6.17}$$

$$y(k,m) := (k^2+m^2)(\sin^2(k) + m^2\cos^2(k))\,. \tag{6.18}$$

Clearly we have $y(k,m) \geqslant 0$ and since $k \geqslant \sin(k)$ for $0 \leqslant k < \pi$, the thesis is proved for the function α.

Again the monotonicity of β for $k \in [0, \pi)$ follows from $\partial_k \beta \geqslant 0$ in the same interval. By elementary computation we have

$$\partial_k \beta = x(k,m)y(k,m)z(k,m), \tag{6.19}$$

$$x(k,m) := \frac{m^2}{\omega_{\mathrm{D}} \sin^2(\omega)}, \tag{6.20}$$

$$y(k,m) := (n\sin(k) - k), \tag{6.21}$$

$$z(k,m) := \frac{n\cos(k)}{\sin^2(\omega)} - \frac{1}{\omega_{\mathrm{D}}^2}. \tag{6.22}$$

Clearly $x(k,m)y(k,m) \leqslant 0$ for $k \in [0,\pi)$ and we just have to verify that $z(k,m) \leqslant 0$ in that interval, namely

$$m^2\cos^2(k) + \sin^2(k) - n\cos(k)\omega_{\mathrm{D}}^2 \geqslant 0\,. \tag{6.23}$$

The last equation is trivially satisfied for $k \in [\pi/2, \pi]$ therefore we restrict to $k \in [0, \pi/2]$. This allows to divide the left side of Eq. (6.23) by $\cos(k)$ achieving

$$m^2\cos(k) + \frac{\sin^2(k)}{\cos(k)} - n\omega_{\mathrm{D}}^2 \geqslant 0 \tag{6.24}$$

which is satisfied if

$$w(k,m) := m^2\cos(k) + \sin^2(k) - n\omega_{\mathrm{D}}^2 \geqslant 0\,. \tag{6.25}$$

It is easy to see that, for any $m \in [0,1]$, we have $\left(\partial_k^{(i)} w(k,m)\right)_{k=0} = 0$ for $i = 0, 1$, while $\partial_k^{(2)} f(k,m) \geqslant 0$ for any $k \in [0, \pi/2]$, which gives the monotonicity of β. ■

Lemma 6.3 *Let $0 \leqslant \bar{k} < \pi$, $\bar{N}$ be a positive integer number, and $\bar{\alpha}$, $\bar{\beta}$ be defined as in Eq. (6.15). If $\bar{\beta} \leqslant 1 - \cos(\frac{\pi}{2\bar{N}})$ and $t \leqslant f(\bar{k}, m, \bar{N})$, where*

$$f(\bar{k}, m, \bar{N}) := \frac{\arccos(\cos\left(\frac{\pi}{2\bar{N}}\right) + \bar{\beta})}{\bar{\alpha}}, \tag{6.26}$$

then

$$\bar{N}\mu(k,m,t)) \leqslant g(\bar{k}, m, \bar{N}, t) \leqslant \frac{\pi}{2} \tag{6.27}$$

where $g(\bar{k}, m, \bar{N}, t) := \bar{N}\arccos\left(\cos(\bar{\alpha}t) - \bar{\beta}\right)$.

Proof. The conditions $t \leqslant f(\bar{k}, m, \bar{N})$ and $\bar{\beta} \leqslant 1 - \cos(\frac{\pi}{2\bar{N}})$ imply

$$0 \leqslant \bar{\alpha}t \leqslant \arccos\left(\cos\left(\frac{\pi}{2\bar{N}}\right) + \bar{\beta}\right) \Rightarrow 1 \geqslant \cos(\bar{\alpha}t) - \bar{\beta} \geqslant \cos\left(\frac{\pi}{2\bar{N}}\right) \Rightarrow$$
$$\Rightarrow \cos(\alpha t) - \beta \geqslant \cos(\bar{\alpha}t) - \bar{\beta} \geqslant \cos\left(\frac{\pi}{2\bar{N}}\right). \quad (6.28)$$

By exploiting the bound (6.12) into Eq. (6.28) we have

$$\begin{aligned} &\cos(\mu(k, m, t)) \geqslant \cos(\bar{\alpha}t) - \bar{\beta} \geqslant \cos\left(\frac{\pi}{2\bar{N}}\right) \Rightarrow \\ &\Rightarrow \bar{N}\mu(k, m, t) \leqslant \bar{N} \arccos\left(\cos(\bar{\alpha}t) - \bar{\beta}\right) \leqslant \frac{\pi}{2}. \end{aligned} \quad (6.29)$$

■

We are now ready to prove the bound (6.10)

Proposition 6.2 *Let U^t and U_D^t be the unitary evolutions given by the Dirac automaton and by the Dirac equation respectively. If the hypothesis of Lemma 6.3 hold then we have*

$$\sup_{\rho \in \mathcal{T}_{\bar{k},\bar{N}}} ||(U^t \rho U^{t\dagger} - U_D^t \rho {U_D^t}^\dagger)||_1 \leqslant \sqrt{1 - \cos^2(g(\bar{k}, m, \bar{N}, t))}. \quad (6.30)$$

Proof. First we notice that thanks to the convexity of the trace distance we can, without loss of generality, consider ρ to be pure. If ρ is a pure state $|\chi\rangle\langle\chi|$ the trace distance becomes $\sqrt{1 - |\langle\chi| U^t U_{\mathrm{D}}^{t\dagger} |\chi\rangle|^2} = \sqrt{1 - |\langle\chi| V(t) |\chi\rangle|^2}$. If we expand $|\chi\rangle$ on a basis of eigenstates of V, i.e. $|\chi\rangle = \sum_{N,\boldsymbol{k},\boldsymbol{s}} \sqrt{p_{N,\boldsymbol{k},\boldsymbol{s}}} |N, \boldsymbol{k}, \boldsymbol{s}\rangle$, we have

$$\begin{aligned} |\langle\chi| V(t) |\chi\rangle| &= \left| \sum_{N,\boldsymbol{k},\boldsymbol{s}} p_{N,\boldsymbol{k},\boldsymbol{s}} \exp\left(i \sum_{j=0}^{N} s_j \mu(k_j, m, t)\right) \right| \geqslant \\ &\geqslant \left| \sum_{N,\boldsymbol{k},\boldsymbol{s}} p_{N,\boldsymbol{k},\boldsymbol{s}} \cos\left(\sum_{j=0}^{\bar{N}} s_j \mu(k_j, m, t)\right) \right|. \end{aligned} \quad (6.31)$$

Combining the bound (6.27) and Eq. (6.31) we have

$$\left| \sum_{N,\boldsymbol{k},\boldsymbol{s}} p_{N,\boldsymbol{k},\boldsymbol{s}} \cos\left(\sum_{j=0}^{\bar{N}} s_j \mu(k_j, m, t)\right) \right|^2 \geqslant \cos^2(g(\bar{k}, m, \bar{N}, t))$$

which finally implies

$$\sqrt{1 - |\langle\chi| V(t) |\chi\rangle|^2} \leqslant \sqrt{1 - \cos^2(g(\bar{k}, m, \bar{N}, t))}.$$

■

Inserting the bound (6.30) into Eq. (6.9) we finally have the bound (6.10).

6.1.3 Discrimination in the regime of small masses and momenta

As already observed the bound in Eq. (6.10) can be computed for any experiment involving the free evolution of the Dirac field. The bound is useful both from the theoretical and from the phenomenological point of view: it allows to verify the theoretical discriminability between the automaton and the Dirac evolutions and it gives insights on the suitable experimental scenario for seeking eventual violations of the Dirac dynamics. On the other hand the analytical expression of (6.10) is very complicated and one cannot easily judge the order of magnitude of the error probability just looking at the formula.

A simplified version of the bound in the typical particle physics regime, namely $k, m \ll 1$, is obtained by expanding in series the function g in Eq. (6.11) near $m = \bar{k} = 0$. Truncating the expansion at the leading order and neglecting a small constant term we have

$$g(m, \bar{k}, \bar{N}, t) \approx \frac{1}{6} m^2 \bar{k}\, \bar{N}\, t, \tag{6.32}$$

and then the following expression for the optimal probability of error

$$\bar{p}_e \geqslant \tfrac{1}{2} - \tfrac{1}{2}\sqrt{1 - \cos^2\left(\frac{1}{6} m^2 \bar{k}\, \bar{N}\, t\right)}. \tag{6.33}$$

In Section 4.3 we have shown how the Dirac automaton recovers the Dirac equation in the limit of small masses and momenta. This is now made rigorous by the optimal probability (6.33) of discriminating between the two evolutions. Indeed for $k, m \ll 1$ the probability of error approaches $\frac{1}{2}$ and then the optimal discrimination experiment is equivalent to a random guessing. This means that in the typical particle-physics regime the two evolutions are very similar.

By putting $\bar{p}_e = 0$, corresponding to $g(m, \bar{k}, \bar{N}, t) = \pi/2$, we obtain the minimum time

$$t_{min}(m, \bar{k}, \bar{N}) \approx 3\pi \frac{1}{m^2 \bar{k} \bar{N}}. \tag{6.34}$$

required for discriminating perfectly between the automaton and the Dirac evolution in a given experiment. Notice that this is an in-principle optimal result, without any specification of the actual apparatus needed to achieve it. For a proton UHECR (see [132]) with $\bar{k} = k_{CR} \approx 10^{-8}$ we have

$$t_{min}(m_p, k_{CR}, 1) \approx 3\pi 10^{46} \text{ Planck times } \approx 10^3 s. \tag{6.35}$$

6.2 Discrimination via particle fly-time

In this Section we consider an elementary experiment for discriminating between the Dirac automaton and the Dirac equation evolution based on particle fly-time. Then we will compare the result with the theoretical optimal in-principle testing of Section 6.1.

The dispersive differential equation (5.15) along with the leading order corrections to the drift and the diffusion coefficients in the relativistic and non relativistic regime (see Eqs. (5.22,5.23) and Eqs. (5.25,5.26)) provide a useful analytic tool for evaluating the macroscopic evolution of the automaton, which otherwise would not be computable in-practice. As a possible experiment for the falsification of the quantum automaton evolution we consider a measurement of fly-time of a particle. As for order of magnitude, we consider numerical values corresponding to UHECRs.

Consider then a proton with $m_p \approx 10^{-19}$ and momentum peaked around $k_{CR} \approx 10^{-8}$ in Planck units, with a spread σ. We ask what is the minimal time t_{CR} for observing a complete spatial separation Δ between the trajectory predicted by the cellular automaton model and the one described by the usual Dirac equation. Thus we require the separation between the two trajectories to be

$$\Delta x \geqslant \hat{\sigma} = \sigma^{-1}, \tag{6.36}$$

that is the initial width in the position space. Notice that UHECRs still belong to the relativistic regime $m_p, k_{CR} \ll 1$, where the automaton is well approximated by the usual Dirac evolution.

We describe the state evolution of the proton wave-packet using the dispersive differential equation (5.15) for an initial Gaussian state, as in the example of Eq. (5.18). The time required to have a separation $\hat{\sigma}$ between the automaton and the Dirac particle is

$$t \approx \hat{\sigma}/\Delta v = \hat{\sigma} \left| \frac{6\sqrt{(k^2+m^2)^3}}{m^2 k^2 (2m^2+k)} \right|, \tag{6.37}$$

where $\Delta v = v - v_D$ is the difference between the automaton and the Dirac drift coefficients in the relativistic regime (5.23). Since it is $m_p/k_{CR} \ll 1$ Eq. (6.37) further simplifies as follows

$$t_{CR} \approx 6\frac{\hat{\sigma}}{m_p^2}. \tag{6.38}$$

In order to be visible the separation $\Delta x = \hat{\sigma}$, the overall broadening $\hat{\sigma}_{br}(t)$ of the two packets must be much smaller than $\hat{\sigma}$. Using Eq. (5.19), which express the time broadening of a Gaussian in terms of the diffusion coefficient, and Eq. (5.23), which provides the automaton diffusion coefficient in the relativistic regime, one has

$$\begin{aligned} \hat{\sigma}_{br}(t) &= \hat{\sigma}\left(\sqrt{1+\left(\frac{D}{2\hat{\sigma}^2}t\right)^2} + \sqrt{1+\left(\frac{D_D}{2\hat{\sigma}^2}t\right)^2} - 2\right) \\ &\approx 2\hat{\sigma}\left(\sqrt{1+\frac{m_p^4}{4\hat{\sigma}^4 k_{CR}^6}t^2} - 1\right), \end{aligned} \tag{6.39}$$

where we used $m_p/k_{CR} \ll 1$. According to Eq. (6.38) we see that $\hat{\sigma} \gg \hat{\sigma}_{br}$ when

$$\hat{\sigma} \gg (k_{CR})^{-3} = 10^{22} \text{ Planck lengths} = 10^2 \text{fm}. \tag{6.40}$$

This result is physically consistent if we observe that the proton wave-packet should have a width bigger than the Fermi length. With $\hat{\sigma} = 10^2\,\text{fm}$ the fly-time for a complete separation between the two trajectories is

$$t_{CR} \approx 6 \times 10^{60} \text{ Planck times } \approx 10^{17} s, \tag{6.41}$$

comparable with the age of the universe. In Eq. (6.35) we evaluated the minimal theoretical time required for discriminating perfectly between the automaton and the Dirac evolution using UHECRs, that is 10^3 s. Unfortunately this means that the particle fly-time is not a good experimental setup for discriminating the automaton evolution and the usual Dirac one. On the other hand the theoretical optimal discrimination shows the in-principle possibility of testing the automaton and one should consider different kind of experiments.

Here we have considered an experiment on UHECRs which have the benefit of carrying a very high momentum. On the other hand they are very rare events and it is possible to consider only experiments involving one cosmic ray. Taking different, less energetic, particles one could engineer experiments involving many particles reducing the minimal time for the discrimination according to Eq. (6.34).

Alternative tests that can be considered are experiments based on quantum interferometry and/or ultra-cold atoms as in Refs. [28, 162–164].

CHAPTER 7

Deformed Relativity and emergent spacetime

The XX century has been the stage for the two theories that have completely modified our intuition in physics—say *Special Relativity* and QT. Special relativity was generalized in order to solve the tension with the Newtonian theory of gravity while the dichotomy between the resulting *General Theory of Relativity* and QT is still a severe issue in modern theoretical physics.

In the last decades many approaches to quantum gravity have been proposed, and some of them have become complete theories like *String Theory* [18, 19] and *Loop Quantum Gravity* [165–168]. Other relevant approaches are the *causal sets* of Bombelli et al. [20], the *non-commutative spacetime* of Connes [21] and the *quantized spacetime* of Synder [22].

As pointed out in [169] the existence of a fundamental scale of length or mass, which could be identified with the Planck scale, seems to be a model-independent feature of quantum gravity. The appearance of this minimum length is the result of combining quantum mechanics ($\hbar$), special relativity (c) and gravity (G),

$$\ell_P = \sqrt{\hbar G/c^3}, \tag{7.1}$$

with its inverse giving the Planck energy E_P. This scales are commonly regarded as the thresholds beyond which the old description of spacetime breaks down and a new phenomenology should appear. The existence of such a scale leads naturally to wonder what is the fate of Lorentz symmetry at the Planck

scale. Moreover in which reference frame l_P and E_P are the thresholds for the new phenomenology?

A possible way of tackling the above questions is a theory with two observer-independent scales, the speed of light and the Planck energy, as proposed by Amelino-Camelia in his seminal papers [23, 24, 170, 171] on *Deformed Special Relativity* (DSR) and later by Smolin and Magueijo who studied models of DSR is a quite general scenario [25, 62]. As we will discuss in the next Section all the DSR realizations have the common feature of predicting a Planck scale deformed kinematics which has the immediate benefit of providing a rich in-principle testable phenomenology. The deformed kinematics exhibits two major physical predictions, *(i)* a wavelength dependence of the speed of photons and *(ii)* modified threshold conditions for particles production in certain collision processes. It is worth mentioning that the possibility of having a deformed covariance in the presence of a varying speed of light was also considered independently by Magueijo in the pioneering works [172–174]. The relevance of the two classes of phenomena owns to the fact that they can be investigated with very high sensitivity in astrophysical context (see for example the Greisen-Zatsepin-Kuzmin theoretical threshold for UHECRs [132] and the equivalent for gamma cosmic-rays [175] which have been largely discussed in the last decade). However, in the class of DSRs not all the representatives have both the above features and the analysis of the emergent phenomenology is clearly model dependent. A particular realization comes from the simple automaton model introduced in this thesis.

In Chapters 4,5 and 6 we have proposed a QCA model for the Dirac field evolution and we have shown that assuming hypothetically the scale of the automaton to be the Planck scale, the usual QFT Dirac dynamics is recovered in the limit of small masses and momenta. The automaton theory does not enjoy an underling notion of Minkowski spacetime (which is for example the starting point of the algebraic approach to QFT), but only a discrete causal network of events with, at most, the metric given by the "event counting". Within this perspective most of the symmetries of QFT, along with the Lorentz covariance, do not correspond to symmetries of the QCA and thereby are broken at the Planck scale. However, in recovering the QFT dynamics as a large-scale approximation, one should also restore its symmetries—say the Lorentz covariance— as approximated symmetries of the emergent theory.

Here we move the first steps toward the emergence of Special Relativity from the automaton framework. In Section 7.1, after a short discussion about the literature on DSR, we show how the Dirac automaton provides a DSR model with an invariant energy scale. We also show that the emergent spacetime exhibits the feature of *relative locality* [63,64], namely different observers which are very far apart, construct different spacetimes starting from the invariant phase space. It is worth stressing that the deformed relativistic model based on the automaton dynamics introduces relative locality in a quantum scenario, which is a quite unusual feature according to [65]. Finally in the last Section

7.2, which is not strictly related to the automaton physics, we show how a digital version of the Lorentz transformations can be obtained by a kind of Einstein synchronization protocol on a classical *causal network*.

The content of this Chapter is based on Refs. [66], and on Refs. [67, 176].

7.1 Deformed Special Relativity from the Dirac automaton

The DSR proposal started off with the consideration that the Planck length (7.1) (or the Planck energy) represents a threshold below which the smooth spacetime geometry could be replaced by a discrete quantum geometry. On the other hand lengths and energies are not invariant under Lorentz transformations and different observers could disagree on the threshold for the appearance of the new quantum effects. This issue can be cured if the lorentz transformations may be modified as to preserve not only the speed of light but also a single energy (or momentum), above which all observes agree that the smoothness of spacetime breaks down. The first model of this kind was proposed by Amelino-Camelia in [23, 170, 171]. Later it has been shown [25, 62] that an invariant energy scale can be introduced keeping the principle of the relativity of inertial frames, and simply modifying the laws by which energy and momenta measured by different inertial observers are related to each other by introducing a non-linear action of the Lorentz transformations on momentum space. As a consequence the quadratic invariant of the Poincaré algebra, also called Casimir element of the algebra,

$$E^2 = p^2 + m^2 \tag{7.2}$$

is replaced by a non-linear invariant, which in turn corresponds to a modified dispersion relation.

One of the interesting aspects of DSR is that in some of its realizations it can be treated in a well established mathematical framework. Amelino-Camelia suggested to consider *quantum groups* (also called *Hopf algebras*[1]) for describing the symmetries of deformed relativity. Indeed it turned out [178,179] that a class of DSRs is equivalent to a "quantum group like" deformation of the usual Poincaré algebra, denoted *k-Poincarré algebra*, with the deformation parameter k representing an Hopf algebra invariant "energy scale" [180, 181].

From a very general point of view a DSR realization is a physical model whose momentum-energy space is described by a k-deformed Poincaré agebra, which is by definition a non-linear agebra. Alternatively, as pointed out in [25,62], a DSR model is described by a non-linear action of the Poincarré group on the momentum-energy space. Clearly in both cases the usual Poincaré linear action is restored at energies much below the Planck one. As observed by many

[1]For a complete reference on Hopf algebras see the book [177] by Majid.

authors working in the field of deformed relativity (see for example Kowalski-Glilman et al. [182]), the formalization of a full DSR theory has necessarily to deal with the following three aspects:

(a) Contrary to the usual Special Relativity, where the Poincaré algebra acts linearly on the momentum-energy space, in the DSR scenario there are essentially no restrictions on the non-linear momentum-energy transformations, allowing for a variety of mathematically equivalent models. On the other hand these models are likely to be inequivalent from the physical point of view and it is still an open problem how to single out one of them via physical principles.

(b) Usually it is provided an algebra which describes only the momentum-energy sector of the theory. On the other hand there is no unique prescription for deriving a picture of the full phase space, that is deforming the momentum-energy sector without providing a connection to the physics of the emergent spacetime could be unjustified. Instead it would be interesting to have a sort of "Einstein protocol" based on some physical model of spacetime at the Planck scale leading to a deformed momentum-energy sector of the kind proposed in DSR.

In the k-deformed Poincaré algebra scenario there is a way out of this issue. Indeed the dual nature of Hopf algebras [2] exhibits the deep connection between the k-Poincaré algebra and the non-commutative k-Minkowski spacetime: the k-Poincaré algebra describes the symmetries of a non-commutative spacetime. The full phase space of a k-deformed algebra was first derived in [183] for the Majid bicrossproduct realization [181] and then extended to other models in [184]. Finally in [182] the authors tried to figure out which properties of the emergent phase space are model dependent and which are not, in order to indicate the second ones as physically observable features of the theory.

Another relevant aspect on the phase space emergent from DSR is the relative locality phenomenon [64]. As a consequence of deforming the momentum space, *absolute locality*—say the assumption that all observers live in the same spacetime—can be violated. More precisely absolute locality is equivalent to the assumption that the momentum space is a linear manifold, which is not the case in DSR.

(c) While classical field theories with k-deformed symmetry have been recently

[2] An Hopf algebra $\mathcal{A}$ is actually a bialgebra, namely an object which is both an algebra (with the usual product $\mathcal{A} \otimes \mathcal{A} \rightarrow \mathcal{A}$ and unit), and a co-algebra (with the co-product $\mathcal{A} \rightarrow \mathcal{A} \otimes \mathcal{A}$ and co-unit) which satisfies logical compatibility rules. As the natural extension of the notion of group an Hopf algebra enjoyes the additional operation called antipode which gives the "inverse" of an element in the algebra upon the application of the co-product. Clearly also the antipode has to satisfy compatibility rules with the underlying bialgebra operations. The algebra and co-algebra structures are in a precise dual relation and the commutativity (non-commutativity) of the product is related to the commutativity (non-commutativity) of the co-product.

considered (see [185] and references therein), the ambiguities concerning how momenta are added and how composite quantities transform slows down the development of k-deformed quantum field theories (see [186] for some recent developments).

The one-dimensional Dirac automaton (4.31) derived in this thesis provides a potential microscopical description of the Dirac field free dynamics at the Planck scale, where it is a causal network of quantum systems in local unitary interaction. Here we show that the automaton provides a special modification of the dispersion relation (7.2) which turns out to be compatible with a deformed relativity with an invariant energy scale.

Concerning point (a), it is interesting to notice that the deformation considered here is motivated by the informational framework at the basis of the quantum field automaton theory, and it is verified a posteriori to satisfy some of the usual features of DSR, as the existence of an invariant energy scale. Regarding point (b) we will discuss the relative locality phenomenon assuming an ansatz for connecting the momentum space to the position one in the automaton DSR realization. Turning to issue (c) we proceed to the analysis of the one-particle sector as in the mainstream literature on DSR.

7.1.1 Invariance of the dispersion relation and deformed Lorentz covariance

The $(1+1)$-dimensional Dirac automaton (4.31) has been derived imposing its invariance with respect to the symmetries of the causal network and describes the one step evolution $\psi(x) \to U\psi(x)$ of a two components field $\psi(x) \coloneqq (\psi_r(x), \psi_l(x))^T$, with ψ_r and ψ_l denoting the left and right modes respectively. The states of the field are obtained by applying the creator operators $\psi^\dagger_{r(l)}$ to the vacuum state $|\Omega\rangle$. In the following we restrict to the one-particle sector for which a orthonormal basis is given by the states $|\eta\rangle\,|x\rangle \coloneqq \psi^\dagger_\eta\,|\Omega\rangle$, $\eta = r, l$. A generic one-particle state is $|\psi\rangle = \sum_{x,\eta} g_\eta(x)\,|\eta\rangle\,|x\rangle$ and Eq. (4.31) becomes a unitary matrix U on $\mathbb{C}^2 \otimes l_2(\mathbb{Z})$.

In the Fourier transformed basis $|\eta\rangle\,|k\rangle$, with $|\phi(k)\rangle := (2\pi)^{-1/2}\sum_x e^{-ikx}\,|x\rangle$, $k \in B \coloneqq [-\pi, \pi]$, the matrix U is written as in Eq. (4.38) and the momentum-energy sector of the automaton is characterized by the dispersion relation (4.41), here reported (in a slightly different form) for the convenience of the reader

$$\cos^2\omega = (1 - m^2)\cos^2 k. \tag{7.3}$$

Remember that the automaton model is naturally band-limited

$$k \in B_k \coloneqq [-\pi, \pi], \qquad \omega \in B_\omega \coloneqq [0, \pi], \tag{7.4}$$

and that $m \in [0, 1]$ with $m = 1$ corresponding to Planck mass. In Section 4.3 we discussed how in the limit $k, m \ll 1$ Eq. (7.3) reduces to

$$\omega^2 = k^2 + m^2, \tag{7.5}$$

(see also Fig. 4.7) and by interpreting the wave-vector k and the frequency ω as momentum and energy[3] we recover the Lorentz dispersion relation (7.2).

Disregarding the internal degrees of freedom, we consider the dispersion relation in Eq. (7.3) as the dynamical equation of the theory which should be independent on the reference frame. In one dimension the usual Lorentz group consists in the only boost transformations which in the momentum-energy sector are represented by the linear map

$$L_\beta : (k, \omega) \mapsto (k', \omega') = \gamma(k - \beta\omega, \omega - \beta k), \tag{7.6}$$

with $\gamma := (1 - \beta^2)^{-1/2}$. It is immediate to check that the Dirac automaton is not Lorentz covariant since its dispersion relation (7.3) changes under the standard boosts of Eq. (7.6). However, the fact that in the regime of small momenta and masses the usual Lorentz quadratic invariant (7.5) is restored suggests that in the automaton framework Lorentz invariance comes as an approximate symmetry, which is broken at high energy scales, exactly as in the DSR models.

We want now to investigate which are the invariance transformations for the automaton dynamics and verify their compatibility with a model of distorted relativity. Since the one-particle sector of the automaton is completely specified by the dispersion relation (7.3) we seek for the transformation which leave it invariant. Following the DSR proposal of preserving the Lorentz group structure, the linear Lorentz boosts (7.6) should be replaced by a non-linear representation of the kind

$$L^D_\beta : (k, \omega) \mapsto (k', \omega') = \mathcal{D}^{-1} \circ L_\beta \circ \mathcal{D}(k, \omega) =: L^D_\beta(k, \omega), \tag{7.7}$$

where $\mathcal{D} : \mathbb{R}^2 \to \mathbb{R}^2$ is a non-linear map. The specific form of $\mathcal{D}$ gives raise to a particular energy-momentum Lorentz deformation. As highlighted in [25] the non-linear map $\mathcal{D}$ has to satisfy some constraints:

(i) The Jacobian matrix $J_\mathcal{D}(k, \omega)$ of the map $\mathcal{D}$ evaluated in $k = \omega = 0$ must be

$$J_\mathcal{D}(0, 0) = I, \qquad (\text{equivalently } J_{L^D_\beta}(0, 0) = L_\beta) \tag{7.8}$$

with I the two by two identity matrix, ensuring the non-linear transformations to recover the standard boosts (7.6) in the regime of small momenta and energies.

(ii) According to Eq. (7.7) the map $\mathcal{D}$ must be invertible in a suitable (see next point (iii)) sector of the momentum-energy space. This sector is included in the interval $B = B_1 \times B_2$ of Eq. (7.4) and we will denote it B_{ph}: $\mathcal{D}$ restricted to B_{ph} is invertible.

[3] It is worth stressing that the correspondence between wave-vector and momentum can only be established within an interacting theory. For example a different possibility is to define the energy as the eigenvalue of the finite-difference Hamiltonian $2iH_f = U - U^\dagger$ (see Eq. (4.50)), which gives $\omega_f = \sin\omega$, and to identify the momentum with $p = \sin k$.

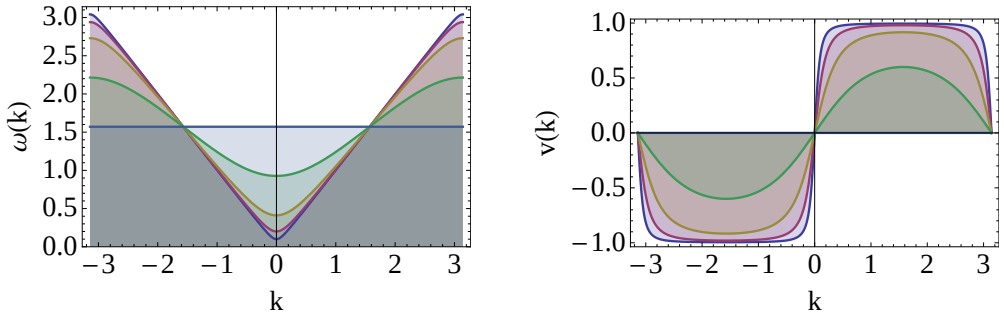


Figure 7.1: The automaton dispersion relation (left) and group velocity (right) for $m = 0.1, 0.2, 0.4, 0.8, 1$, from bottom to top on the left and from top to bottom on the right). Notice the fixed point $\omega = \pi/2$ in the dispersion relation at $k = \pm\pi/2$. The same points correspond to the maximum group velocity.

(iii) Since the ordinary Lorentz boosts span the whole interval $[0, \infty]$, both for momentum and energy, the range of the map $\mathcal{D}$ must include the whole interval $[0, \infty]$. Notice that the invertibility region of (ii) must include the interval $[0, \infty]$.

(iv) A DSR model exhibits an invariant energy scale only if the map $\mathcal{D}$ has a singular point, namely some energy ω_{inv} which is mapped to ∞ by $\mathcal{D}$

$$\mathcal{D}_2(k_{inv}, \omega_{inv}) = \infty, \qquad \omega_{inv} = \omega(k_{inv}). \tag{7.9}$$

Indeed Eq. (7.7) shows that the invariants of the new deformed theory are the inverse images via $\mathcal{D}$ of the Special Relativity invariants, namely infinite energy.

Let us now focus on the automaton dispersion relation (7.3) and observe that by simple manipulation it can be restated as

$$\frac{\sin^2 \omega}{\cos^2 k} - \tan^2 k = m^2. \tag{7.10}$$

Therefore we find that the non linear map $\mathcal{D}$ in Eq. (7.7) can be taken to be

$$\mathcal{D}(k, \omega) := \left(\tan k, \frac{\sin \omega}{\cos k} \right), \tag{7.11}$$

and corresponds to the following deformed Lorentz transformations

$$\begin{aligned} k' &= \arctan \left(\gamma \left(\tan k - \beta \frac{\sin \omega}{\cos k} \right) \right), \\ \omega' &= \arcsin \left(\gamma \left(\frac{\sin \omega}{\cos k} - \beta \tan k \right) \cos k' \right). \end{aligned} \tag{7.12}$$

The map (7.11) derived from the invariance of the automaton dispersion relation automatically satisfies the requirements (i-iv) and then provides a DSR

model with an invariant energy scale. One can easily verify that $J_{\mathcal{D}}(0,0) = I$ and in the relativistic regime, namely $k, \omega \ll 1$ and for β sufficiently small to keep k', ω' small, $\mathcal{D}$ approaches the identity function with the transformations (7.12) approaching the usual Lorentz boosts. According to (iv) since $\omega = \frac{\pi}{2}$ is mapped to ∞ by $\mathcal{D}$,

$$\mathcal{D}_2(\pm\tfrac{\pi}{2}, \tfrac{\pi}{2}) = \infty \Rightarrow \omega_{inv} = \tfrac{\pi}{2}, \tag{7.13}$$

it is an invariant energy. Looking at the modified dispersion relation (7.3) (see also left Fig. 7.1) we notice that it has two symmetrical fixed points at $k = \pm\frac{\pi}{2}$ which correspond to the invariant energy $\omega(k) = \frac{\pi}{2}$ independently on the mass m. Finally the invariant energy separates the two physical regions

$$B_{ph} = B^1_{ph} \cup B^2_{ph}, \qquad \begin{array}{l} B^1_{ph} = (\text{-}\frac{\pi}{2}, \frac{\pi}{2}) \times [0, \frac{\pi}{2}), \\ B^2_{ph} = (\text{-}\pi, \text{-}\frac{\pi}{2}) \cup (\frac{\pi}{2}, \pi) \times (\text{-}\frac{\pi}{2}, \pi] \end{array} \tag{7.14}$$

in which $\mathcal{D}$ is invertible with full range $[0, \infty]$.

In the right Fig. 7.1 we plot the automaton group velocity $v = \partial\omega(k)/\partial k$ (4.53) which, as discussed in Chapter 5, represents the velocity of propagation of "particles" (see for example the automaton velocity operator (5.50) expressed in the automaton particle-antiparticle basis). For massless particles the Dirac QCA dispersion relation (7.3) coincides with the undistorted one $\omega^2 = k^2$ and the group velocity is the maximal speed (the speed of light $c = 1$) on the causal network. In [187, 188] it is possible to find a discussion on the propagation velocity of particles in DSR. In [188] it is argued that any massless particle should have the same speed c, while in [187] the authors show that the appropriate description of the propagation speed is given by the group velocity.

One can easily check that the two "physical regions" $B^{1(2)}_{ph}$ remain separated under all possible boosts (7.12), namely

$$(k, \omega) \in B^{1(2)}_{ph} \Rightarrow (k', \omega') \in B^{1(2)}_{ph}. \tag{7.15}$$

The fixed ponts $k = \pm\pi/2$ at the separation between the two regions correspond to a maximum of the group velocity (see right Fig. 7.1). While in region B^1_{ph} an increase in k corresponds to an increased group velocity, in region B^2_{ph} we see the opposite trend with an increasing k corresponding to a decreasing group velocity. On the other hand, as one can verify using the transformations (7.12), a boosted observer who sees an increased group velocity in B^1_{ph}, also sees an increased group velocity in B^2_{ph} since in both cases the wave-vector k is mapped closer to the invariant point. Since the two regions exhibit the same kinematics they are undistinguishable in a non interacting framework. In Section 4.3.1 we have shown that the monotonicity of the automaton dispersion relation avoids the Fermion doubling issue arising in the Lattice Gauge Theory scenario, on the other side it seems that from the kinematical point of view the momentum-energy sector could present a novel "particle doubling". The answer to this question need the analysis of a more general interactive scenario

in which the two sectors could lead to a different phenomenology. In case of indistinguishability one should reduce the periodicity of the QCA momentum space (7.4) to the smaller region $k \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ [4].

7.1.2 Position representation and relative locality

The action of the boosts (7.12) on the states of the automaton (disregarding the internal degrees of freedom) reads

$$|\psi\rangle = \int \mathrm{d}\mu_k \, g(k) |\tilde{k}\rangle \xrightarrow{L^D_\beta} \int \mathrm{d}\mu_k \, g(k) |\tilde{k}'\rangle = \int \mathrm{d}\mu_{k'} \, g(k(k')) |\tilde{k}'\rangle \tag{7.16}$$

where k' is as in Eq. (7.12),

$$\mathrm{d}\mu_k \, = \mathrm{d}k \, (2(1-m^2)\tan\omega(k))^{-1} \tag{7.17}$$

is the invariant measure in the momentum space and we defined

$$|\tilde{k}\rangle = (2(1-m^2)\tan\omega(k))^{1/2} \, |k\rangle \, . \tag{7.18}$$

One can verify that the transformation (7.16) is unitary.

Let us now deepen our analysis considering how the features of the present framework affect the geometry of space and time. The aim is then to derive the position-time deformed Lorentz transformations corresponding to the momentum-energy ones (7.12). As remarked in the introduction to this Section (see point (b)) in the general DSR scenario there is not a unique way to connect the momentum space (k, ω) to the position space (x, t). This is mainly due to the non trivial geometry of the momentum space induced by the non-linear deformed Lorentz transformations.

Under the action of the deformed boost L^D_β a function $f(k, \omega)$ transforms as $f'(k, \omega) = f(k', \omega')$ and, following an ansatz due to Schützhold et al. [189], one can express the boosted function in the variables x, t by conjugating the boost L^D_β with the Fourier transform $\mathcal{F}$ [5] *i.e.* first we take the Fourier transform $\mathcal{F}$ to the momentum space, then the deformed boost (7.12) in the momentum space

[4] Following an analogy between the automaton lattice and a crystal in solid-state physics this corresponds to take a smaller Brillouin zone which, by definition, describes the full momentum space.

[5] Here $\mathcal{F}$ denotes the Fourier transform in both position and time variables: $\mathcal{F}(f)(\omega, k) = \sum_{t,x} e^{i(\omega t - kx)} f(\omega, k)$, $\mathcal{F}^{-1}(f)(t, x) = \int \mathrm{d}\mu \, e^{-i(\omega t - kx)} f(t, x)$.

is applied, and finally the function is Fourier transformed back [6]

$$\begin{aligned} f' &= \mathcal{F}^{-1} \circ L^D_\beta \circ \mathcal{F}\, f, \\ f'(x',t') &= \sum_{x,t\in\mathbb{Z}} \int \mathrm{d}\mu'\, e^{-i\chi(k',\omega',x,t,x',t')} f(x,t), \\ \chi(k',\omega',x,t,x',t') &:= k(\omega',k')x - k'x' - \omega(\omega',k')t + \omega' t'. \end{aligned} \tag{7.19}$$

Notice that, due to the non linearity of $\mathcal{D}$, the map (7.19) does not correspond to a change of coordinates from (x,t) to (x',t') and therefore we cannot interpret the variables x and t as the coordinates of points in a continuum spacetime interpolating the automaton cells: this may be regarded as manifestation of the quantum nature of spacetime. Following the idea that space-time should emerge from a collection of physical events labelled by means of a synchronization procedure (like the Einstein protocol), we now define a point of the "emergent spacetime" as the "crossing of the trajectories of two particles" (see Fig. 7.2).

In the automaton context a particle corresponds to a wave-packet narrowly peaked in the momentum space around some k_0. Thus in the following we consider smooth-states peaked around some momentum k_0 and having width σ according to Definition 5.1. We can imagine a collection of wave-packets $g_{k_0}(t,x)$ and the emergent spacetime as the collection of all the collision points of these particles. Within this perspective a collision point p is an extended object corresponding to a region in the (x,t) plane, whose coordinates (x_p,t_p) can be identified with the mean values of x and t in the collision area.

The deformed Lorentz transformations in position space can now be derived applying the transformation (7.19) to the points (x_p,t_p) given by the coincidences of localized wave-packets. In order to have a consistent definition of spacetime as emergent from coincidences we have to verify that a peaked state remains peaked after a non-linear Lorentz boost. If we look at the wave-packet $g_{k_0}(x,t)$ at a given time step and we move to the momentum representation, we can consider the dependence on k only, using the dispersion relation $\omega(k)$

[6] We are aware that the validity of this construction relies on the definition of the Fourier transformation and its inverse in the context of the distorted relativity theories where the integration process could turn out to be non-trivial both in momentum and in position space [190]. We have already discussed (see point (b) in the introduction to this Section) the example of k-Poincaré deformed relativity where the "curved momentum space" corresponds to the k-Minkowski spacetime whose non-commutative structure implies a non-trivial integral calculus.

In the continuos spacetime case, assuming the usual definition for the Fourier transformation corresponds to the prescription $ip \leftrightarrow \partial/\partial x$, $[X,P] = i$. However, in the Planckian regime the commutator $[X,P] = i$ could be modified as discussed in [182, 184] or in [191] where the author showed how quantum-gravitational effects and the presence of a minimal length could give a "generalized uncertainty principle" instead of the usual Heisenberg uncertainty relation.

The QCA commutator $[X,P]$ has been evaluated in Eq. (5.33) and differs from the usual Heisenberg uncertainty relation by an additional term. This term is relevan for highly localized states while it becomes negligible for wave-packets narrow-band in momentum.

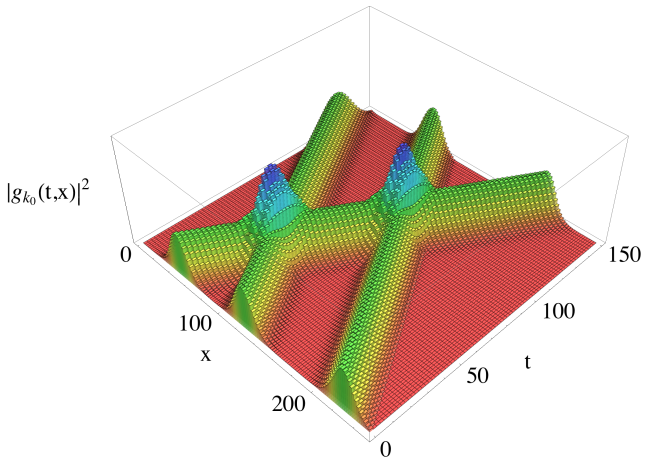


Figure 7.2: Two coincidences of travelling wave-packets in the automaton evolution.

(7.3) and writing $\mathcal{D}(k)$ for $\mathcal{D}(k,\omega(k))$ in Eq. (7.7). If the initial wave-packet is peaked in k_0 and has width σ in momentum space, the boosted state in the momentum representation

$$g'_{k'_0}(k') = g_{k_0}(k(k'))\left(\frac{\mathrm{d}k}{\mathrm{d}k'}\right)^{1/2}, \tag{7.20}$$

has respectively momentum and position width given by

$$\sigma' \approx \sigma\left(\frac{\mathrm{d}k}{\mathrm{d}k'}\bigg|_{k'=k'_0}\right)^{-1} = \sigma\,\frac{\mathrm{d}k'}{\mathrm{d}k}\bigg|_{k=k_0}, \qquad \hat{\sigma}' \approx \hat{\sigma}\,\frac{\mathrm{d}k}{\mathrm{d}k'}\bigg|_{k'=k'_0}, \tag{7.21}$$

where $\hat{\sigma} = \sigma^{-1}$ and k'_0 is the the boosted wave vector given by Eq. (7.12). The narrow-band wave-packet in the unprimed reference frame corresponds to a still narrow-band wave-packet peaked around k'_0 in the primed one, indeed the Jacobian

$$\frac{\mathrm{d}k'}{\mathrm{d}k} = \gamma\frac{1+\beta(v\cos k+\sin^2\omega(k)\sin k)}{\cos^2 k+\gamma^2(\sin k+\beta\sin\omega(k))^2} \tag{7.22}$$

assumes small values for any β, k, and m (except for $k=\pi/2, \beta=1$ and $k=0, \beta=1$) as one can check in Fig. 7.3. In Fig. 7.4 we also show some examples of deformed boosts on smooth-states.

In the following analytical derivation we work with continuous x and t interpolating exactly the automaton (4.31) discrete evolution as described in Remark 5.1 and replacing summations in Eq. (7.19) with integrals. Since for a wave-packet $g_{k_0}(t,x)$ peaked around k_0 the boosted wave-packet $g'_{k_0}(x',t')$ is still narrow-band and peaked around k'_0, the map (7.19) can be approximated by taking the first order Taylor expansion of $k(\omega',k')$ and $\omega(\omega',k')$ around k'_0 and $\omega'(k_0)$ in the function χ. Truncating the Taylor expansion at the first

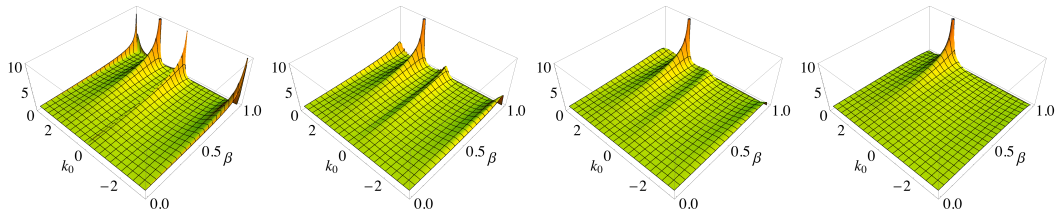


Figure 7.3: Plot of the Jacobian $(\mathrm{d}k'/\mathrm{d}k)_{k=k_0}$ providing the width (7.21) of the boosted wave-packet for an initial smooth-state peaked around k_0 and having width σ. The jacobian is plotted as a function of k_0 and β. From the left to right the Jacobian is depicted for different values of the mass: $m = 0.01, 0.2, 0.4, 0.8$.

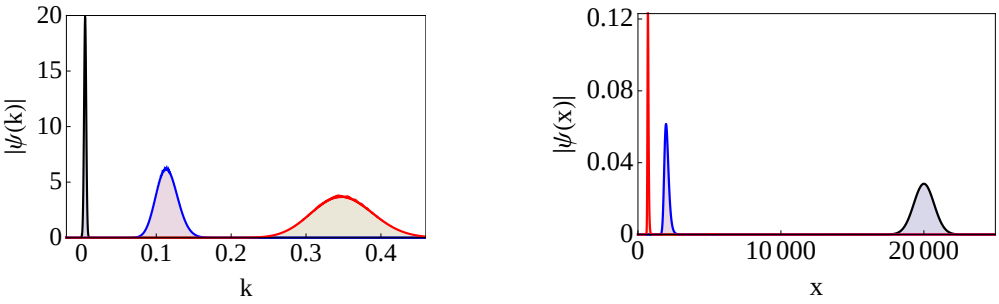


Figure 7.4: Simulation of the transformation of a smooth-state (see Definition 5.1) due to a deformed boost (7.12). Here the mass is $m = 0.01$ and the initial state is a Gaussian (in black) peaked around momentum $k_0 = 5 \times 10^{-3}$ and position $x_0 = 2 \times 10^4$. The initial momentum and position widths are respectively $\sigma = 1/500$ and $\hat{\sigma} = \sigma^{-1} = 500$. The left (right) figure shows the boosted state in the momentum (position) representation for two different values of β, say $\beta = -0.99$ (in blue) and $\beta = -0.999$ (in red). Using Eq. (7.12) we find that the boosted states are respectively peaked in $k'_0 = 0.11$ and $k'_0 = 0.35$, according to the figure. Notice the spreading (narrowing) of the state in the momentum (position) representation with the jacobian in Eq. (7.21) given by $(\mathrm{d}k'/\mathrm{d}k)_{k_0} = 10, 28$ for the two values of β.

order

$$\begin{aligned} k &\approx k(k'_0,\omega'_0) + \left.\frac{\partial k}{\partial k'}\right|_{k'_0}(k'-k'_0) + \left.\frac{\partial k}{\partial \omega'}\right|_{\omega'_0}(\omega'-\omega'_0) \\ \omega &\approx \omega(k'_0,\omega'_0) + \left.\frac{\partial \omega}{\partial k'}\right|_{k'_0}(k'-k'_0) + \left.\frac{\partial \omega}{\partial \omega'}\right|_{\omega'_0}(\omega'-\omega'_0) \end{aligned} \tag{7.23}$$

we find the following relation

$$g_{k_0}(x',t') \propto g_{k_0}\left(\left.\frac{\partial k}{\partial k'}\right|_{k'_0} x - \left.\frac{\partial k}{\partial \omega'}\right|_{\omega'_0} t\ ,\ \left.\frac{\partial \omega}{\partial k'}\right|_{k'_0} t - \left.\frac{\partial \omega}{\partial \omega'}\right|_{\omega'_0} x\right),$$

which corresponds to the momentum-energy dependent Lorentz transformations

$$\begin{pmatrix} x' \\ t' \end{pmatrix} \approx \begin{pmatrix} \partial_{k'}k & -\partial_{\omega'}k \\ -\partial_{k'}\omega & \partial_{\omega'}\omega \end{pmatrix}_{k'_0,\omega'_0} \begin{pmatrix} x \\ t \end{pmatrix}. \tag{7.24}$$

Since Eq. (7.24) defines a linear transformation of the variables x and t and the wave-packets move along straight lines, we could interpret (7.24) as the transformation of the coordinates x_p and t_p of a point p which is identified by the crossing of the trajectories of two particles having k's close to some common k_0. It is immediate to verify that the usual Lorentz boosts (7.6) are recovered for $k, \omega \ll 1$ and $k', \omega' \ll 1$.

However, the k-dependance of the transformations (7.24) prevents a well defined notion of a common spacetime shared by all the observers. Consider a point p which is given by the intersection of four wave-packets, the first pair peaked around k_1 and the second pair peaked around k_2 ($k_1 \neq k_2$). Because of the k dependence in (7.24), a boosted observer will actually see the first pair intersecting at a point which is different from the one where the second pair intersects (see Fig. 7.5). As an example suppose the two pairs crossing at $(x, 0)$ for some observer. According to Eq. (7.24) the spatial separation between the two crossings in the boosted frame is

$$|x'_1 - x'_2| \approx \left|(\partial_{k'} k)_{k'_1} - (\partial_{k'} k)_{k'_2}\right| |x| \tag{7.25}$$

and can be made arbitrarily big taking the spatial origin x big enough. In particular the spreading/narrowing of the wave-packets due to the boost (see Eq. (7.21)) can be made negligible with respect to the relative distance taking the origin in a position such that

$$|x| \gg \frac{\hat{\sigma}'_1(t'_1) + \hat{\sigma}'_2(t'_2)}{\left|(\partial_{k'} k)_{k'_1} - (\partial_{k'} k)_{k'_2}\right|}. \tag{7.26}$$

The loss of coincidence effect, first noticed in Ref. [189], is the characteristic trait of the so-called *relative locality* [63, 64]. According to relative locality spacetime is no longer an objective entity but different observers infer different spacetimes and interactions that are local in the spacetime constructed by one observer may be non-local to another one. In this scenario the only invariant object under change of reference frame is the full phase space of the theory. Therefore the sole coincidence in spacetime is not an objective event (and it is not paradoxical that different observers could disagree on it) while the points of the full phase space must be regarded as objective—say different observers agree on the coincidence of two "trajectories in the phase space".

Relative locality appears as an unavoidable features of all the models in which the momentum space has a non-flat geometry. This has been proved in [63] where the authors show that absolute locality is equivalent to the assumption that momentum space is a linear manifold. This can be easily seen as follows. In order to recover absolute locality, and then the invariance of coincidences in spacetime, we require that the transformation (7.24) does not depend on k_0 and ω_0, that is

$$\begin{pmatrix} \partial_{k'} k & -\partial_{k'} \omega \\ -\partial_{\omega'} k & \partial_{\omega'} \omega \end{pmatrix}_{k'_0, \omega'_0} = \begin{pmatrix} \alpha_1(\beta) & \alpha_3(\beta) \\ \alpha_3(\beta) & \alpha_4(\beta) \end{pmatrix} \qquad \forall\ k'_0, \omega'_0, \tag{7.27}$$

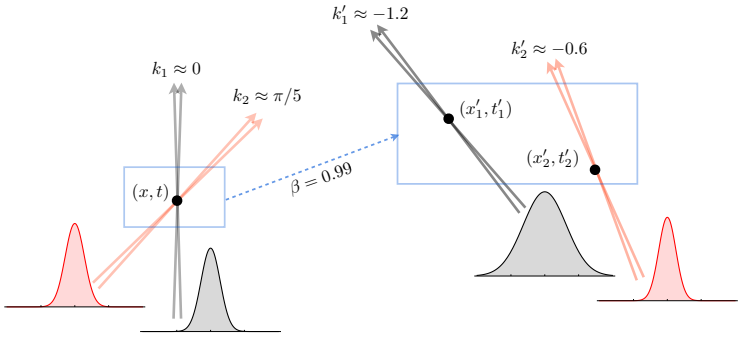


Figure 7.5: Example of the relative locality phenomenon in automaton DSR model. Here we consider an automaton having $m = 0.4$ and two pairs of Gaussian wave-packets respectively peaked around $k_1 \approx 0$ and $k_2 \approx \pi/5$ (group velocities $v_1 = 0$, $v_2 = 0.8$). In the rest frame the packets have spatial width $\hat{\sigma}_1 = \hat{\sigma}_2 = 10^3$ and cross at $(x, t) = (10^5, 0)$. After a boost $\beta = 0.99$, the transformed wave-vectors are $k_1' = -1.2$ and $k_2' = -0.6$ (group velocities $v_1' = -0.9$, $v_2' = -0.8$) and by Eq. (7.24) it is possible to find the coordinates in the boosted frame: $(x_1', t_1') \approx (7 \times 10^5, -6 \times 10^5)$ and $(x_2', t_2') \approx (4 \times 10^5, -4 \times 10^5)$.

for some functions α_i. Noticing that the map (7.27) depends only on β and that in the regime of small momenta and energies it has to correspond to the usual Lorentz map L_β (see constraint (7.8)), we find that (7.27) must coincide with the Lorentz transformation. Any non linear mapping $\mathcal{D}$ of the momentum space (7.7) is then incompatible with the assumption that all observers live in the same spacetime.

7.1.3 Boost of perfectly-localized states and distorted Lorentz contraction

We have already noticed that a novel feature of the automaton description of field dynamics is to accommodate perfectly-localized states, see for example Fig. 4.6 and Eq. (5.27). It is then interesting to evaluate the deformed Lorentz transformations for these states. For a state $g(x) = \delta_{x0}$ perfectly localized in $x = 0$ it is $g(k) = 1/\sqrt{2\pi}$ and according to Eq. (7.20) one has

$$g'(x') = \mathcal{F}\left(g'(k')\right), \qquad g'(k') = \frac{1}{\sqrt{2\pi}} \left(\frac{\mathrm{d}k}{\mathrm{d}k'}\right)^{1/2}, \tag{7.28}$$

namely the boosted state in position representation is simply the Fourier transform of the Jacobian square root. In the left Fig. 7.6 it is shown how a perfectly-localized state delocalizes when submitted to a boost. It is also interesting look at the transformed state in the momentum representation (see right Fig. 7.6) where the it coincides with the Jacobian square root. The flat distribution becomes peaked in the extremal points $k = 0, \pm\pi$ and in the invariant point $k = \frac{\pi}{2}$ ($k = -\frac{\pi}{2}$) for negative (positive) β. Since in $k = 0, \pm\pi$ the group ve-

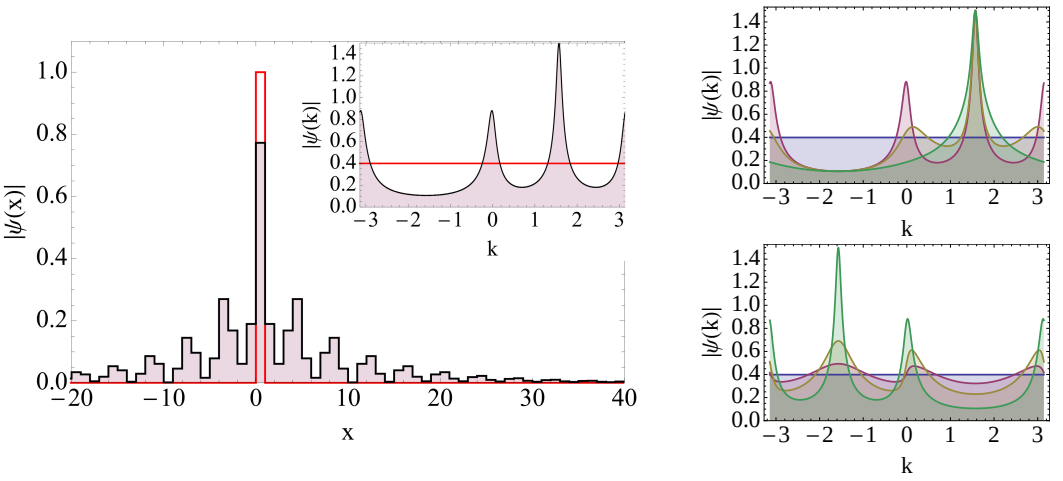


Figure 7.6: **Left figure:** delocalization (in black) of a state (in red) perfectly localized at $x = 0$ in the rest frame after a boost of $\beta = -0.99$ for mass $m = 0.1$. The top figure depicts the momentum representation, the bottom one the position representation. **Right figure:** momentum representation of a boosted localized state for different values of the mass (top) and of the boost velocity β (bottom). In the top figure it is $\beta = -0.99$ and $m = 0.1, 0.3, 0.8$, we notice that the higher is the mass the larger is the peak around the invariant momentum $k = \pi/2$ with the other peaks disappearing. In the bottom it is $m = 0.1$ and $\beta = 0.4, 0.8, 0.99$, we notice that the higher is β the more peaked is the transformed function.

locity is zero, the wave-packet acquires a drift contribute corresponding to the maximal group velocity $v(\frac{\pi}{2})$ $(v(\text{-}\frac{\pi}{2}))$ (see Fig. 7.1).

Besides relative locality, another example of how the non-linear momentum-energy transformations affect the bahaviour in the spacetime constructed by different observers, is the distortion to the length contraction factor. Consider two reference systems sharing the origin $(x, t) = (x', t') = 0$ with the primed reference boosted by β. In the unprimed frame consider a smooth-state peaked

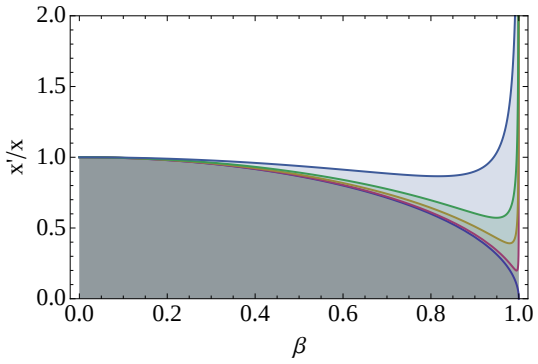


Figure 7.7: Deformed Lorentz contraction factor (7.30) as a function of the boost velocity β and for four different values of the mass $m = 0.1, 0.2, 0.3, 0.5$, from bottom to top. At the very bottom is the usual Lorentz contraction factor γ^{-1}.

around $k_0 = 0$ and with average position x. Exploiting the transformations (7.24) we find that the average position x' in the boosted reference frame is equal to

$$x' = \left.\frac{\partial k}{\partial k'}\right|_{k'_0} x \; + \; \left.\frac{\partial k}{\partial \omega'}\,\frac{\partial \omega'}{\partial k'}\right|_{k'_0} x, \tag{7.29}$$

where the second term on the RHS of the equation is due to the fact that the wave-packet in the boosted frame is peaked around $k'_0 \neq 0$ and is then moving with a group velocity $v' = (\partial\omega'/\partial k')_{k'_0}$ for a time interval $t' = (\partial k/\omega')_{k'_0}$. Therefore the deformed length contraction coefficient is

$$\frac{x'}{x} = \left.\frac{\mathrm{d}k}{\mathrm{d}k'}\right|_{k'_0} = \left(\left.\frac{\mathrm{d}k'}{\mathrm{d}k}\right|_{k_0=0}\right)^{-1} = \gamma^{-1}[1 - \beta^2(1 - m^2), \tag{7.30}$$

where we have used Eq. (7.22) in $k_0 = 0$. In Fig. 7.7 the length contraction factor is plotted as a function of the boost velocity β and for different values of the mass. Notice that for boosts which bring k'_0 close the invariant momentum, the contraction instead becomes a dilation.

7.2 Digital Lorentz transformation on a classical causal network

In last Section we have studied the phase space of the Dirac QCA which corresponds to a model of DSR. In the context of DSR the emergent spacetime has been derived as a "linear projection" of the non trivial phase space. On the contrary it would be interesting to have the spacetime emergent from the synchronization protocol of a "clock" defined on the QCA providing an operational procedure for the emergence of spacetime from the automaton elementary description. The notion of clock in a quantum scenario has been considered since the seminal paper by Salecker and Wigner [192] where they argued that a quantum clock has a "limited precision" due to his size and mass. In order to be a clock, a physical system has to be localized in a finite region of space for an interval of time long enough to allow the time measure in that interval. The trade-off between the spatial resolution and the temporal resolution of the clock appears also in the automaton context. Indeed one can imagine a clock as a "localized smooth state" (see Definition 5.1) evolving with a certain group velocity and diffusion coefficient as described in Chapter 5: if the state is perfectly localized (maximal resolution) then it diffuses immediately and is not suitable for measuring long time intervals, on the other hand if it is very delocalized its spatial resolution decreases. This is also related with the so called Margolous Levitin bound [193] later studied as the *quantum speed limit* by Lloyd et al. [194]. The deepening in this direction is beyond the purpose of the present thesis and is still under investigation.

In this Section we consider a slightly different and simpler scenario which is not directly related with the QCA framework. We we analyze the mechanism of emergence of spacetime from the homogeneous topology of a classical causal network in $1+1$ dimensions. Upon considering the causal connections as exchanges of classical information, we can establish coordinate systems via a kind of Einstein protocol, leading to a digital version of the Lorentz transformations. In a computational analogy first noticed by Leslie Lamport [195], the foliation construction can be regarded as the synchronization protocol with a global clock of the calls to independent subroutines (the causally independent events) in a parallel distributed computation. In this scenario a clock is simply a periodic sequence of causally connected events on the network and each clock corresponds to a foliation of the network. Hence, the Lorentz time-dilation corresponds to an increased number of leaves within a clock *tic-tac*, whereas space-contraction results from the corresponding decreased density of events per leaf of the foliation. We will see that the operational procedure of building up the coordinate system introduces an in-principle indistinguishability between neighboring events, resulting in a network that is coarse-grained, the thickness of the event being a function of the observer's clock. The digital version of the Lorentz transformation is an integer relation which differs from the usual analog transformation by a multiplicative real constant corresponding to the event thickness. The composition rule for velocities is independent on such constant, and it is the same in both the analog and the digital versions.

7.2.1 Causal networks

We now introduce the main notion of causal network (CN) as a partially ordered set of events with the partial order representing the causal relation between two events.

Definition 7.1 (Causal Network) *A* CN *is a set* N *of elements called events* $a, b, c \ldots, \in \mathsf{N}$ *equipped with a partial order relation* $\preceq$ *which is:*

(i) Reflexive: $\forall a \in \mathsf{N}$ *we have* $a \preceq a$.

(ii) Antisymmetric: $\forall a, b \in \mathsf{N}$, *we have* $a \preceq b \preceq a \Rightarrow a = b$.

(iii) Transitive: $\forall a, b, c \in \mathsf{N}$, $a \preceq b \preceq c \Rightarrow a \preceq c$.

(iv) Locally finite: $\forall a, c \in \mathsf{N}$, $|\{b \in \mathsf{N} : a \preceq b \preceq c\}| < \infty$, *where* $|\mathsf{S}|$ *denotes the cardinality of the set* S.

In the following we will also write $a \prec b$ *to state that* $a \preceq b$ *with* $a \neq b$. *A causal network is represented by a graph with points being the events and the edges drawn between any two points* a *and* b *for which* $a \preceq b$—i.e. *that are causally connected, as in Fig. 7.8. Usually a* CN *is unbounded in all directions, while a finite region on it is called causal set. In order to satisfy transitivity, the* CN *is a* directed acyclic graph *(*DAG*),* i.e. *loops are forbidden (arrows on edges are usually not drawn by orienting the graph* e.g. *from the bottom to the top).*

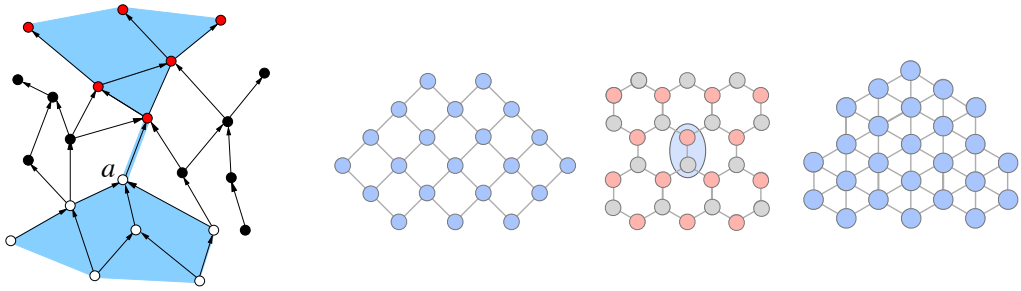


Figure 7.8: **Left figure:** Illustration of the past/future light-cone of the event a in a CN. **Right figure:** The three possible homogeneous topologies for causal networks in $1+1$ dimensions.

Due to the causal structure the CN is endowed with the notion of *light-cone* J_a of an event $a \in \mathsf{N}$, along with those of *past/future light-cone* $\mathsf{J}_a^-/\mathsf{J}_a^+$, respectively (see Fig. 7.8)

$$\mathsf{J}_a^- \coloneqq \{b \in \mathsf{N} : b \preceq a\}, \quad \mathsf{J}_a^+ \coloneqq \{b \in \mathsf{N} : a \preceq b\}, \qquad \mathsf{J}_a \coloneqq \mathsf{J}_a^- \cup \mathsf{J}_a^+. \tag{7.31}$$

Accordingly, one has that $a \preceq b$ is equivalent to $a \in \mathsf{J}_b^-$ and to $b \in \mathsf{J}_a^+$. We will call *independent* or *space-like* two events $a, b \in \mathsf{N}$ that are not causally related, say $a \notin \mathsf{J}_b$ (or $b \notin \mathsf{J}_a$), and *causally dependent* or *time-like* otherwise, namely $a \in \mathsf{J}_b$ (or $b \in \mathsf{J}_a$). We call a CN *connected* if for every $a, b \in \mathsf{N}$ there exists $c \in \mathsf{J}_a \cap \mathsf{J}_b$. Two events that are not space-like are connected by at least a causal chain, *e.g.* $a \preceq b$ are connected by the causal chain

$$\mathsf{C}(a,b) \coloneqq \{c_i\}_{i=1}^N, \qquad a \equiv c_1 \prec c_2 \prec \dots c_N \equiv b. \tag{7.32}$$

It is convenient to orient the chain, generalizing its definition to include the case $b \preceq a$. The verse of the chain is taken into account by a signed cardinality $|\mathsf{C}(a,b)|_\pm \coloneqq \sigma|\mathsf{C}(a,b)|$ with $\sigma = +$ for $a \prec b$, and $\sigma = -$ for $b \prec a$.

Definition 7.2 (Homogeneous causal network) *For any event a denote by $\mathsf{I}_{in}(a) = \{i_1(a), i_2(a), \dots i_n(a)\}$ and $\mathsf{I}_{out}(a) = \{o_1(a), o_2(a), \dots o_m(a)\}$ the set of input and output links of a. We now say that a* CN *is* topologically homogeneous *if for each couple of events $a, b \in \mathbb{N}$ one has the isomorphism*

$$i_j(a) \simeq i_j(b), \quad j = 1, \dots n \qquad o_j(a) \simeq o_j(b), \quad j = 1, \dots m. \tag{7.33}$$

It is easy to see that in $1+1$ dimensions there are only three possible lattices corresponding to a homogeneous CN: the square, the triangular, and the honeycomb ones as shown in the right Fig. 7.8. However, the honeycomb-lattice has two inequivalent types of events (having one input and two output links and viceversa), and the corresponding "undressed" topology—where each couple of connected inequivalent events are merged into a single event—reduces to the square lattice. The triangular-lattice, on the other hand, has redundant causal links (the middle vertical ones). We are thus left with the square-lattice

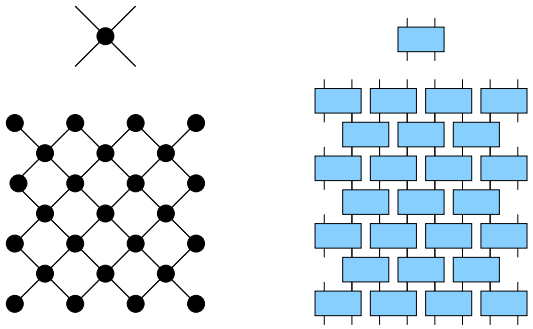


Figure 7.9: Homogeneous CN and equivalent representation as a circuit.

as the most general homogeneous CN in $1+1$ dimensions. In this case there are only two types of link: toward the right and toward the left—shortly r-link and l-link (see Fig.7.9 where the network is also depicted as a circuit whose elementary gate has the same input-output structure all over the network). Notice that the CN of the 1d Dirac automaton derived in Section 4.2 is a square-lattice (see Figs. 4.3 and 4.4). Indeed the automaton description of QFT was motivated by a homogeneous computation which corresponds to a homogeneous CN.

In the following we will call *light signals* those sequences of events that are connected only by r-links or only by l-links. Their "speed" is equal to "one event-per-step", and it is the maximum speed allowed. Regarding the homogeneous CN as the description of physical systems in interaction, the causality is sufficient to guarantee a bound for the speed of the "information flow".

7.2.2 Foliations and digital coordinates

Now we can introduce a notion of simultaneity for the events on the CN. Being the equivalent of a *world line* an unbounded causal chain on the CN can be identified with an observer. Hence, an observer will be denoted as

$$\mathsf{C} = \{c_i\}_{i\in\mathbb{Z}}, \qquad c_i \preceq c_{i+1}\ \forall i \in \mathbb{Z}, \tag{7.34}$$

where c_0 represents the origin while the index $i \in \mathbb{Z}$ plays the role of the observer's proper time. Thanks to the topological homogeneity, we can translate the origin of the observer to any event $a \in \mathsf{N}$ (more in general we could denote by C the equivalence class of all observers translated over all the events of the CN). Among all the possible observers we consider a special subclass denoted *clocks*

Definition 7.3 (Clock) *A clock is a causal chain of events periodically oscillating between two different "positions" of the* CN*, namely the oscillation (tic-tac) is a periodic sequence of α r-links followed by β l-links identifying*

events in the same "position" (see Fig. 7.10). We denote the chain of events in the same position as $\mathsf{C}^{\alpha\beta}$.

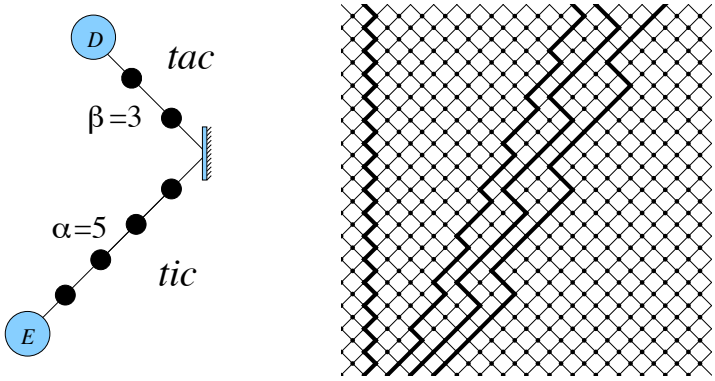


Figure 7.10: The *tic-tac* of the clock is represented by the number α of r-links between the emitter E and the mirror, and the number β of l-links between the mirror and the detector D. From the left to the right we have the rest-frame clock, corresponding to $\alpha = \beta = 1$, and boosted-frames for $\alpha = 3, \beta = 1$, $\alpha = 6, \beta = 2$, and $\alpha = 7, \beta = 2$, respectively, corresponding to digital speed $v = \frac{1}{3}$, $v = \frac{1}{3}$ and $v = \frac{2}{7}$, respectively. The case $\alpha = 6, \beta = 2$ has doubled imprecision, compared to the case $\alpha = 3, \beta = 1$.

A clock introduces a notion of simultaneity for the events on the CN. From the clock $\mathsf{C}^{\alpha\beta}$ a light signal is sent (sending event $a_s \in \mathsf{C}^{\alpha\beta}$) to an event $a \in \mathsf{N}$ and then received back at the clock (receiving event $a_s \in \mathsf{C}^{\alpha\beta}$) . The event $c_i \in \mathsf{C}^{\alpha\beta}$ at the intermediate step between the sending a_s and the receiving a_r event is taken as synchronous with the event a at the turning point—denoted as $a \overset{\mathsf{C}^{\alpha\beta}}{\sim} b$—while the number of *tic-tacs* divided by two is taken as the distance from the turning point a and the clock.

The given notion of simultaneity allows us to associate each observer with a *foliation* of the CN. For each event $c_i \in \mathsf{C}^{\alpha\beta}$ there is a *leaf* $\mathsf{L}_i(\mathsf{C}^{\alpha\beta})$, which is the set of events simultaneous to c_i with respect to the observer $\mathsf{C}^{\alpha\beta}$, namely

$$\mathsf{L}_i(\mathsf{C}^{\alpha\beta}) \coloneqq \{b \in \mathsf{N};\, b \overset{\mathsf{C}^{\alpha\beta}}{\sim} c_i\}. \tag{7.35}$$

The collection of all leaves for all the events in $\mathsf{C}^{\alpha\beta}$ is the *foliation* $\mathsf{L}(\mathsf{C}^{\alpha\beta})$ of N associated to the observer $\mathsf{C}^{\alpha\beta}$

$$\mathsf{L}(\mathsf{C}^{\alpha\beta}) \coloneqq \{\mathsf{L}_i(\mathsf{C}^{\alpha\beta}), \forall i \in \mathbb{Z}\}. \tag{7.36}$$

Notice that to any clock $\mathsf{C}^{\alpha\beta}$, with origin in c_0, is associated a reference frame $\mathfrak{R}$ which is just the foliation of the network built up using the Einstein protocol leading to the following digital coordinates for an event $a \in \mathsf{N}$

$$\begin{aligned} t(a) &\coloneqq |\mathsf{C}^{\alpha\beta}(c_0, a_r)|_{\pm} + \frac{1}{2}|\mathsf{C}^{\alpha\beta}(a_s, a_r)|_{\pm}, \\ s(a) &\coloneqq \frac{1}{2}|\mathsf{C}^{\alpha\beta}(a_s, a_r)|_{\pm}, \end{aligned} \tag{7.37}$$

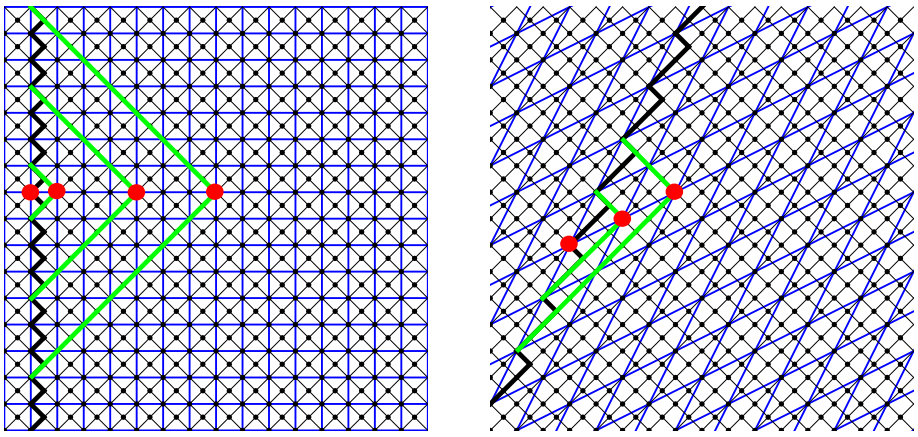


Figure 7.11: Illustration of the Einstein protocol for establishing a reference frame. **Left figure:** Rest-frame. The blue lines represent the reference frame established using the clock with $\alpha = \beta = 1$ (see Fig. 7.10). The green lines represent light signals bouncing between the clock and four particular events in the network. These events are synchronous, since the intermediate time between the sending and the receiving event on the clock is the same for all of them. They lie on the same leaf of the foliation, but at different distance from the clock, $0, 1, 4, 7$, respectively: the spatial coordinate is obtained by counting the *tic-tacs* between the the sending and the receiving event $(0, 2, 8, 14$,respectively) divided by two. **Right figure:** Reference frame (blue lines) for $\alpha = 3, \beta = 1$, built up using the same protocol as in the felt figure.

where the foil $\mathsf{L}_{t(a)}$ contains the event a and $s(a)$ is the number of events on the same foil in between the clock and the event a.

In Fig. 7.11 the Einstein protocol is illustrated in two particular reference frames. The figure on the left corresponds to the rest-frame, with the blue lines depicting the coordinate system established using the clock with $\alpha = \beta = 1$. The green lines represent light signals bouncing between the clock and four particular events in the network. These events are synchronous, since the intermediate time between the sending and the receiving event on the clock is the same for all of them. They lie on the same leaf of the foliation, but at different position, $0, 1, 4, 7$, respectively: the spatial coordinate is obtained by counting the *tic-tacs* between the the sending and the receiving event divided by two. The right figure represents a frame for $\alpha = 3, \beta = 1$, built up using the same protocol as in the left figure.

7.2.3 Boosted frames and digital Lorentz transformation

One can associate to any frame $\mathfrak{R}$ a parameter v, in terms of the α and β of its clock, by simply evaluating the "distance in space" and the "distance in time"

between the starting and the ending point of the *tic-tac*, namely

$$v = \frac{\alpha - \beta}{\alpha + \beta}. \tag{7.38}$$

Then any couple of reference frames $\mathfrak{R}_1$ and $\mathfrak{R}_2$ can be related by the same parameter

$$v_{12} = \frac{\alpha_{12} - \beta_{12}}{\alpha_{12} + \beta_{12}}, \tag{7.39}$$

where now it is

$$\alpha_{12} = \alpha_2/\alpha_1 \in \mathbb{Q}, \qquad \beta_{12} = \beta_2/\beta_1 \in \mathbb{Q} \tag{7.40}$$

($\mathbb{Q}$ denoting rational numbers). The parameter $v_{12} \in \mathbb{Q}$ is interpreted as the relative velocity of the frame $\mathfrak{R}_2$ with respect to the frame $\mathfrak{R}_1$ and $\mathfrak{R}_2$ is said to be boosted with respect to $\mathfrak{R}_1$. Accordingly one has

$$v_{21} = \frac{\alpha_{21} - \beta_{21}}{\alpha_{21} + \beta_{21}} = -v_{21}. \tag{7.41}$$

Now, considering three reference frames $\mathfrak{R}_1$, $\mathfrak{R}_2$ and $\mathfrak{R}_3$ and using the trivial identities

$$\alpha_{13} = \alpha_{12}\alpha_{23}, \qquad \beta_{13} = \beta_{12}\beta_{23}, \tag{7.42}$$

one has

$$v_{13} = \frac{\alpha_{12}\alpha_{23} - \beta_{12}\beta_{23}}{\alpha_{12}\alpha_{23} + \beta_{12}\beta_{23}}, \tag{7.43}$$

which by simple algebraic manipulations immediatley gives

$$v_{13} = \frac{\left(\frac{\alpha_{12}-\beta_{12}}{\alpha_{12}+\beta_{12}}\right) + \left(\frac{\alpha_{23}-\beta_{23}}{\alpha_{23}+\beta_{23}}\right)}{1 + \left(\frac{\alpha_{12}-\beta_{12}}{\alpha_{12}+\beta_{12}}\right)\left(\frac{\alpha_{23}-\beta_{23}}{\alpha_{23}+\beta_{23}}\right)} = \frac{v_{12} + v_{23}}{1 + v_{12}v_{23}}, \tag{7.44}$$

which corresponds to the composition rule of parallel velocities in Special Relativity.

We have shown how the Einstein protocol on the CN allows to associate coordinates (7.37) to a given event in a certain reference frame $\mathfrak{R}$. Now we derive the coordinates in a boosted frame in terms of the rest frame ones, along with the relative coordinate systems between any couple of boosted frames.

Consider a the rest reference frame $\mathfrak{R}_1$ with $\alpha_1 = \beta_1 = 1$ and a boosted frame $\mathfrak{R}_2$ with certain α_2, β_2. In order to connect the coordinate systems in the two frames we have chosen the same origin $(0, 0)$ on both $\mathfrak{R}_2$ and $\mathfrak{R}_1$. The generic event on $\mathfrak{R}_2 \cap \mathfrak{R}_1$ has coordinates (s_2, t_2) and (s_1, t_1) in the two frames, respectively. One can see that a spatial step in $\mathfrak{R}_2$ corresponds to $(\alpha_{12}+\beta_{12})/2$ space and $(\alpha_{12} - \beta_{12})/2$ time steps in $\mathfrak{R}_1$ while a time step in $\mathfrak{R}_2$ corresponds

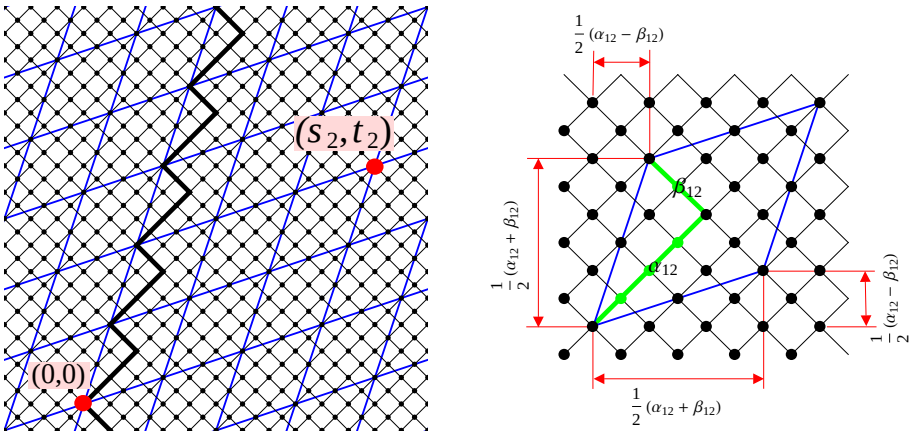


Figure 7.12: Illustration of the derivation of Eqs. (7.45) and (7.46), leading to the digital version of the Lorentz transformations in Ref. (7.51). **Left figure:** the reference frame $\mathfrak{R}_1$ with $\alpha_1 = \beta_1 = 1$ is represented by the tiny network in black, whereas the coarser network in blue represents the boosted reference frame $\mathfrak{R}_2$ with $\alpha_2 = 4$, $\beta_2 = 2$. According to Eq. (7.39) the relative velocity of $\mathfrak{R}_2$ with respect to $\mathfrak{R}_1$ is $v_{12} = 1/3$. In order to connect the coordinate systems in the two frames we have chosen the same origin $(0, 0)$ on both $\mathfrak{R}_2$ and $\mathfrak{R}_1$. The generic event on $\mathfrak{R}_2$ has coordinates $(s_2, t_2) = (3, 2)$ and $(s_1, t_1) = (11, 9)$ in the two frames, respectively. **Right figure:** a spatial step in $\mathfrak{R}_2$ corresponds to $(\alpha_{12} + \beta_{12})/2$ space and $(\alpha_{12} - \beta_{12})/2$ time steps in $\mathfrak{R}_1$. In the same way a time step in $\mathfrak{R}_2$ corresponds to $(\alpha_{12} - \beta_{12})/2$ space and $(\alpha_{12} + \beta_{12})/2$ time steps in $\mathfrak{R}_1$. This correspondence allows to determine the coordinates (s_1, t_1) of a given event in the frame $\mathfrak{R}_1$ in terms of its coordinates (s_2, t_2) in the frame $\mathfrak{R}_2$. The resulting transformations are in Eqs. (7.45) and (7.46).

to $(\alpha_{12} - \beta_{12})/2$ space and $(\alpha_{12} + \beta_{12})/2$ time steps in $\mathfrak{R}_1$ (one can also check this in right Fig. 7.12). Accordingly we have

$$s_1 = \tfrac{1}{2}(\alpha_{12} + \beta_{12})s_2 + \tfrac{1}{2}(\alpha_{12} - \beta_{12})t_2, \tag{7.45}$$
$$t_1 = \tfrac{1}{2}(\alpha_{12} + \beta_{12})t_2 + \tfrac{1}{2}(\alpha_{12} - \beta_{12})s_2, \tag{7.46}$$

and the invariance of the topology with boosts guarantees that the same equations hold between any couple of boosted frames. By elementary manipulation Eqs. (7.45) and (7.46) can be written in the more familiar way

$$s_1 = \tfrac{1}{2}(\alpha_{12} + \beta_{12})\,(s_2 + v_{12}t_2)\,, \tag{7.47}$$
$$t_1 = \tfrac{1}{2}(\alpha_{12} + \beta_{12})\,(t_2 + v_{12}s_2)\,. \tag{7.48}$$

Upon defining the following constant depending on the clocks $\mathsf{C}^{\alpha_1\beta_1}$ and $\mathsf{C}^{\alpha_2\beta_2}$ of the two frames

$$\chi_{12} := \sqrt{\alpha_{12}\beta_{12}}, \tag{7.49}$$

and using the identity

$$\tfrac{1}{2}(\alpha_{12} + \beta_{12}) = \frac{\chi_{12}}{\sqrt{1 - v_{12}^2}}, \tag{7.50}$$

we obtain the digital Lorentz transformations

$$s_1 = \chi_{12}\frac{s_2 + v_{12}t_2}{\sqrt{1-v_{12}^2}}, \quad t_1 = \chi_{12}\frac{t_2 + v_{12}s_2}{\sqrt{1-v_{12}^2}}. \tag{7.51}$$

The equations in (7.51) differ from the usual analog Lorentz transformations by the multiplicative factor χ_{12}, which makes the transformations rational, compensating the irrationality of the boost factor $\sqrt{1-v_{12}^2}$. The digital-analog conversion is thus just a rescaling of both space and time coordinates by the factor $(\alpha\beta)^{\frac{1}{2}}$ depending on the boost. In the next Section we observe that the events of the boosted frame can be seen as coarse-graining of the events in the rest frame. The rescaling factor corresponds to the square-root of the volume of the coarse-grained event measured as the number of rest-frame events that it contains. Such event volume also affects the Lorentz space-contraction and time-dilation factor, which in the digital case is given by $\frac{1}{2}(\alpha_{12}+\beta_{12})$, whereas in the analog case is rescaled by the ratio of event volumes, leading to $\frac{1}{2\sqrt{\alpha_{12}\beta_{12}}}(\alpha_{12}+\beta_{12})$. Thus, for example, for $\alpha_{12}=1$ and $\beta_{12}=3$, corresponding to $v_{12}=1/2$, the digital factor is 2 whereas the analog one is $2/\sqrt{3}$. As we will notice in next Section the rescaling factor is due to the different precision of the clocks used for building up the reference frames coordinate systems.

7.2.4 Clock precision, events coarse-graining and the Lorentz limit

Due to the indivisibility of the clock *tic-tac*, we see that there are indiscernible events, for which the synchronisation occurs in the middle of the *tic-tac*. We are thus led to identify events, and merge them into thicker coarse-grained events. This is done as follows. We identify the events along the *tic* and those along the *tac* so that the *tic-tac* is always regarded as the bouncing between two nearest neighboring events. Then we merge events into minimal sets so that the topology is left the invariant, namely the usual square-lattice one (see Figs. 7.13 and 7.14).

We can distinguish between two different kinds of coarse-graining: one due to the boosting (in yellow in the figures), and one due to the intrinsic imprecision of the clock (in gray). The difference between the two is clarified in Fig. 7.14: in the left figure events along the *tic* and events along the *tac* are identified in the boosted frame. Then events are merged into minimal sets (in yellow) so that the topology is left invariant (the merged events are again events of a square-lattice network). In the middle figure the coarse graining associated to the intrinsic imprecision is added in gray, and finally, in the bottom figure the circuit is stretched so to have all synchronous events on horizontal lines, and events located in the same position on vertical lines. Notice that in the special case of the rest-frame, see Fig. 7.13, the coarse-graining is just due to the intrinsic imprecision of the clock.

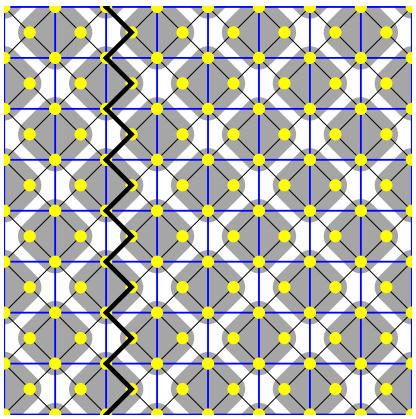


Figure 7.13: Illustration of the coarse-graining procedure due to the intrinsic imprecision of the clock (in gray).

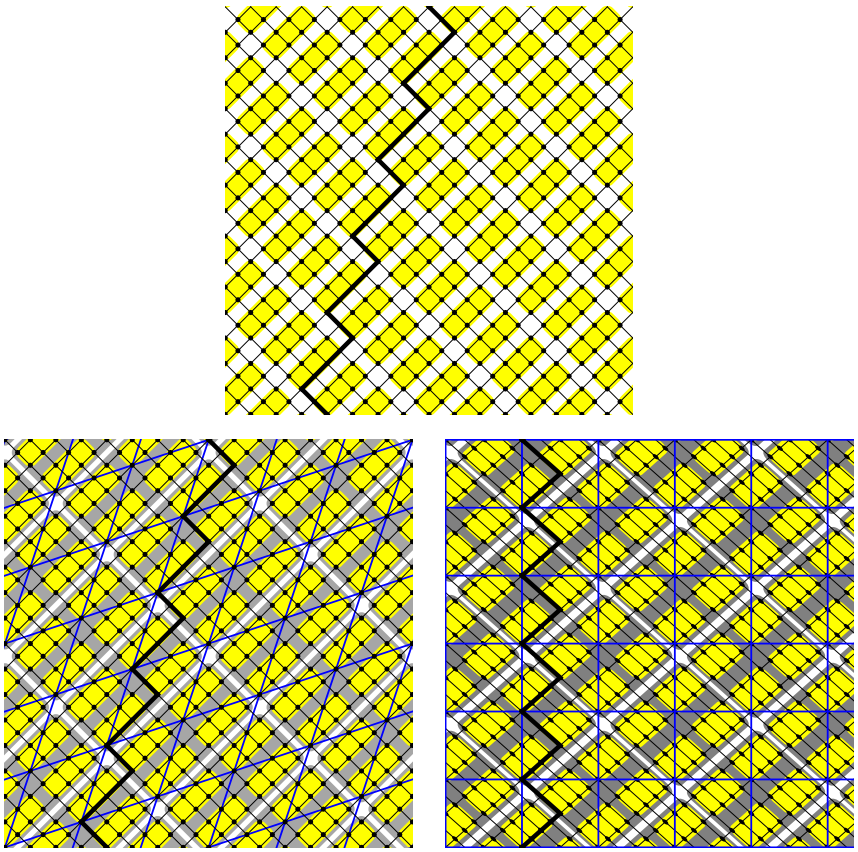


Figure 7.14: **First figure:** events along the *tic* and events along the *tac* are identified in the boosted frame. Then events are merged into minimal sets (in yellow) so that the topology is left invariant (the merged events are again events of a square-lattice network). **Second figure:** the coarse graining associated to the intrinsic imprecision is added in gray. **Third figure:** the circuit is stretched so to have all synchronous events on horizontal lines, and events located in the same position on vertical lines.

It is reasonable to wonder if the digital Lorentz transformations derived in Section 7.2.3 admit a limit in which the usual Lorentz transformations are recovered and this will be considered in the forthcoming publications. In a reasonable physical scenario one should compare reference frames built up using clocks having similar precision. Given a clock $\mathsf{C}^{\alpha\beta}$ we can define its precision as the number of events in the corresponding coarse-graining, namely

$$p(\mathsf{C}^{\alpha\beta}) = \alpha\beta. \tag{7.52}$$

Moreover, as in the Dirac automaton case, one can imagine the CN scale to be very small with any macroscopical clock including numerous events of the network

$$\alpha, \beta \gg 1. \tag{7.53}$$

In this scenario we should compare reference frames $\mathfrak{R}_1$ and $\mathfrak{R}_2$ such that

$$\alpha_1\beta_1 \approx \alpha_2\beta_2 \gg 1. \tag{7.54}$$

In this case we have $\chi_{23} \approx 1$ and the digital Lorentz transformations in (7.51) are approximated by the usual ones. Notice that since $\alpha_i\beta_i \gg 1$ we can obtain almost any relative velocity according to Eq. (7.39).

CHAPTER 8

Quantum Theory with superselection: the Fermionic case

In Chapter 4 we have defined a Quantum Field Cellular Automaton as an automaton evolving Bosonic or Fermionic quantum fields. Within the field QCA perspective the difference between the two scenarios is in the nature of the local systems on the lattice: in the Bosonic case they are commuting quantum systems—say qubits—while in the Fermionic case they are anticommuting systems—say Fermionic modes. One could wonder if a field automaton for Fermionic systems can be simulated by an automaton evolving commuting quantum systems. This issue was first raised by Richard Feynman in the 1982 when in his seminal work on physical computation [42] he wondered about the possibility of simulating Fermions by local quantum systems in interaction—what we would call nowadays a *quantum computer*.

Since the pioneering paper by Feynman the relation between Fermions and quantum systems has been largely investigated both from the computational and from the physical viewpoint. On one hand one can simply consider the Fermionic modes as the elementary systems of a computational model different from QT and study its features as a theory of information processing. On the other hand one can focus on the compatibility between local operations in the Fermionic and in the quantum case, namely on the possibility of mimic local interactions of Fermionic modes by local interactions of qubits and viceversa.

The Jordan-Wigner map [128] transforms isomorphically the Fermionic an-

ticommuting algebra into a commuting qubit algebra. Such a correspondence has been a valuable instrument for solving the 1d XY spin-chains [196, 197] and has been used to extend to the Fermionic case computational notions as the entanglement [198], the entropic area law [199], and the universal computation [200]. However, the Jordan-Wigner isomorphism is not a "physical isomorphism" since it does not map local Fermionic operations into local quantum ones, and nonlocal to nonlocal ones. One could say the Jordan-Wigner map is not *isolocal*. This leads to some ambiguities in defining the partial trace [201–204], and in assessing the local nature of operations [205].

Here the Wigner superselection rule comes to help. The Wigner superselection rule for Fermionic systems forbids superpositions between states with odd and even particle number, based on the simple argument of the impossibility of discriminating a 2π rotation from the identity [32, 206]. The Wigner superselection rule avoids the problems connected to the isolocality [198], but it is still not clear if it can restore the isolocality of the Jordan-Wigner map showing the reciprocal simulability of Fermionic and quantum interactions.

In this Chapter we consider the Fermionic model from the aforementioned first perspective, namely as a probabilistic theory of the kind considered in Chapter 2, and derive the consequences of the Wigner superselection rule on the informational features of the theory—say on its degree of holism and on the monogamy of entanglement[1]. After introducing for the first time a notion of superselection rule for a general probabilistic theory, corresponding to a linear constraint over the convex set of states, we provide a link between the cardinality of the rule (the number of linearly-independent constraints) and the degree of holism of the resulting theory. Then we define the *Fermionic Quantum Theory* (FQT) and show that it is rigorously a superselected version of QT which lacks local tomography and monogamy of entanglement. The monogamy violation goes hand in hand with the existence of mixed maximally entangled states and of non trivial MESs [90] (see Section 2.4.3) for bipartite states (unlike QT where non-trivial MESs occur starting from the tripartite case).

The content of this Chapter is based on Ref. [68].

8.1 Superselection rules for a general probabilistic theory

In Chapter 2 we have introduced the framework of operational probabilistic theories. Here we restrict to general probabilistic theories that are convex (all

[1]Local tomography has been extensively discussed in Section 2.3, and corresponds to the possibility of discriminating between two multipartite states using only local measurements [10, 207]. As stated in Section 2.3 a theory that lacks local tomography is called holistic [71]. The monogamy of entanglement, defined in Section 2.4.4, is loosely speaking the limitation on the sharing of entanglement. For example, if two qubits are completely entangled with each other, neither of them can be entangled with any other object.

sets of transformation are convex, see Remark 2.2) and causal [10] (namely, the probability of the preparation is independent of the choice of the observation test, see Axiom 2.1). As usual we denote by $\mathrm{St}(\mathrm{A})$ and $\mathrm{Eff}(\mathrm{A})$ the convex set of (generally subnormalized) states and the convex set of effects of a system A, while $\mathsf{Transf}(\mathrm{A},\mathrm{B})$ denotes the convex set of transformations from the input system A to the output system B. We remind that the dimension of a system A is given by $D_{\mathrm{A}} \coloneqq \dim(\mathsf{St}_{\mathbb{R}}(\mathrm{A})) = \dim(\mathsf{Eff}_{\mathbb{R}}(\mathrm{A}))$, with $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ and $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$ the the linear spans of the set of states and effects. Remind also that these two linear spaces are in a duality relation $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A}) = \mathsf{St}_{\mathbb{R}}(\mathrm{A})^*$, with effects corresponding to non negative linear functional on the set of states and viceversa.

For simplicity we assume the no restriction hypotesis of Remark 2.1, namely all the complete state preserving transformations are transformations of the probabilistic theory. Since also effects are special kind of transformations—say transformations with no output system—then in the presence of the no restriction hypotesis a system A is fully specified by its set of states. In this case we can identify the theory itself as a collection of systems and their set of states

$$\mathcal{T} := \{\mathrm{A}, \mathsf{St}(\mathrm{A})\}. \tag{8.1}$$

Imposing a superselection rule σ on a theory $\mathcal{T}$ corresponds to sectioning linearly all sets of transformations for each multipartite system, which under the no restriction hypothesis is equivalent to linear sectioning the sole sets of states. We can give the following formal definition of superselection rule:

Definition 8.1 (Superselection rule) *Given a probabilistic theory $\mathcal{T}$ a superselection rule σ*

$$\sigma : \mathcal{T} \to \bar{\mathcal{T}}, \qquad \mathrm{A} \mapsto \sigma(\mathrm{A}) =: \bar{\mathrm{A}}, \tag{8.2}$$

maps any system $\mathrm{A} \in \mathcal{T}$ to a new system $\sigma(\mathrm{A}) \in \bar{\mathcal{T}}$ whose set of states $\mathrm{St}(\sigma(\mathrm{A}))$ is a linear section of $\mathrm{St}(\mathrm{A})$, namely for any system A one has

$$\mathrm{St}(\sigma(\mathrm{A})) \coloneqq \{\rho \in \mathrm{St}(\mathrm{A});\, (s_i^\sigma \,|\, \rho) = 0,\, i = 1, \dots, V_{\mathrm{A}}^\sigma\},$$

where

$$s_i^\sigma \in \mathrm{Eff}_{\mathbb{R}}(\mathrm{A}), \qquad i = 1, \dots V_{\mathrm{A}}^\sigma \tag{8.3}$$

are V_{A}^σ linear independent constraints.

Notice the set of states and effects of the constrained theory $\bar{\mathcal{T}}$ are included in the sets of the original theory $\mathcal{T}$

$$\mathrm{St}(\bar{\mathrm{A}}) \subseteq \mathrm{St}(\mathrm{A}), \qquad \mathrm{Eff}(\bar{\mathrm{A}}) \subseteq \mathrm{Eff}(\mathrm{A}), \tag{8.4}$$

and one has

$$D_{\bar{\mathrm{A}}} = D_{\mathrm{A}} - V_{\mathrm{A}}^{\sigma}. \tag{8.5}$$

Moreover the linearity of σ preserves the convexity of the theory, that is the sets $\mathrm{St}(\bar{\mathrm{A}})$, $\mathrm{Eff}(\bar{\mathrm{A}})$ and $\mathsf{Transf}(\mathrm{A},\mathrm{B})$ are still convex sets. We will see that for certain superselection rules the set of states of the constrained theory splits into two or more sectors each one corresponding to a different "quantum number", with superselection preventing from superimposing states in different sectors. However, since the superselected theory is still convex, one can always mix them.

For consistency of the theory, the superselection map σ must commute with the systems parallel composition (see Section 2.1.1) and this implies the following definition of parallel composition for the constrained theory

$$\sigma(\mathrm{A})\sigma(\mathrm{B}) \coloneqq \sigma(\mathrm{AB}). \tag{8.6}$$

The number of linear constraints that one can impose on a certain theory is not completely arbitrary and in the following proposition we derive both a lover an upper bound for the number of constraints on a composite system AB given the constraints on the component ones.

Proposition 8.1 (Bounds on the linear constraints) *Consider a probabilistic theory $\mathcal{T}$ and a superselection rule σ as in Definition 8.1. Then, for the consistency of the superselected theory $\bar{\mathcal{T}}$, we have the following lower and upper bound for the number of linear constraints on composite systems*

$$V_{\mathrm{AB}}^{\sigma} \geqslant D_{\mathrm{A}} V_{\mathrm{B}}^{\sigma} + D_{\mathrm{B}} V_{\mathrm{A}}^{\sigma} - 2 V_{\mathrm{A}}^{\sigma} V_{\mathrm{B}}^{\sigma}, \tag{8.7}$$

$$V_{\mathrm{AB}}^{\sigma} \leqslant D_{\mathrm{A}} V_{\mathrm{B}}^{\sigma} + D_{\mathrm{B}} V_{\mathrm{A}}^{\sigma} - V_{\mathrm{A}}^{\sigma} V_{\mathrm{B}}^{\sigma} + D_{\mathrm{AB}} - D_{\mathrm{A}} D_{\mathrm{B}}. \tag{8.8}$$

Proof. We first prove the upper bound (8.8). In a general probabilistic theory it is $\mathsf{St}_{\mathbb{R}}(\mathrm{A}) \otimes \mathsf{St}_{\mathbb{R}}(\mathrm{B}) \subseteq \mathsf{St}_{\mathbb{R}}(\mathrm{AB})$ from which it follows the relation

$$D_{\mathrm{AB}} \geqslant D_{\mathrm{A}} D_{\mathrm{B}} \tag{8.9}$$

for the dimension of the composite systems. Applying (8.9) to the systems of the constrained theory $\bar{\mathcal{T}}$ and using Eq. (8.5) we get the bound (8.8).

The lower bound (8.7) is proved showing that all the local constraints on the component systems A and B are also constraints of the composite system AB, namely for any $b \in \mathrm{Eff}(\bar{\mathrm{B}})$ and any $i = 1, \ldots, V_{\mathrm{A}}^{\sigma}$, one has that $s_i^{\sigma} \otimes b \in \mathrm{Eff}_{\mathbb{R}}(\mathrm{AB})$ is a constraint for $\overline{\mathrm{AB}}$. Indeed, suppose by contradiction that for $\rho \in \mathrm{St}(\overline{\mathrm{AB}})$ it is

$$(s_i^{\sigma} \otimes b | \, \rho) \neq 0, \tag{8.10}$$

then since $|\rho_b) = (b|\,\rho) \in \mathsf{St}(\bar{\mathrm{A}})$ is a valid state for the system $\bar{\mathrm{A}}$, we have $(s_i^{\sigma}|\,\rho_b) \neq 0$ against the hypothesis. The same argument holds reversing the

roles of the subsystems $\bar{\mathrm{A}}$ and $\bar{\mathrm{B}}$, and we conclude that the composite system $\overline{\mathrm{AB}}$ must satisfy at least a number of constraints

$$D_{\bar{\mathrm{A}}} V_{\mathrm{B}}^{\sigma} + D_{\bar{\mathrm{B}}} V_{\mathrm{A}}^{\sigma}, \tag{8.11}$$

which, using Eq. (8.5), gives the bound (8.7). ■

In principle, given a theory $\mathfrak{T}$, one could build "bottom-up" a superselected theory defining the constraints only on the elementary systems (the ones that cannot be obtained by composition of other systems) and taking the minimal number of linear constraints (8.7) on the composite ones. These theories are denoted *minimally superselected* according to the following definition:

Definition 8.2 (Minimally superselected theory) *We call minimal a superselected theory that saturates the lower bound in Eq.* (8.7)*. For such a theory the constraints on bipartite systems are of the form* $s_i^{\sigma} \otimes b$ *and* $a \otimes r_j^{\sigma}$*, with* $a \in \mathrm{Eff}(\bar{\mathrm{A}})$ *and* $b \in \mathrm{Eff}(\bar{\mathrm{B}})$*.*

If instead we saturate the upper bound Eq. (8.8) we have a *maximal superselected* theory:

Definition 8.3 (Maximally superselected theory) *We call maximal a superselected theory that saturates the upper bound in Eq.* (8.8)*.*

8.1.1 Superselection-holism trade-off.

In Section 2.3 we introduced rigorously the notion "tomography" for a probabilistic theory. In simple terms it regards the possibility of reconstructing any multipartite state from the statistics of measurements on the component systems, or equivalently the possibility of discriminating between multipartite states using measurements on the component systems. Accordingly a theory has been defined strictly n-local-tomographic if the set of n-local effects (but not the $n-1$ one) is separating for multipartite states[2]. In Definition 2.12 we denoted such a theory n-holistic—say it has degree of holism n.

For convenience of the reader we remind that a local-tomographic theory (*i.e.* $n = 1$) is characterized by the following relation on the dimension of the composite systems

$$D_{\mathrm{AB}} = D_{\mathrm{A}} D_{\mathrm{B}}, \tag{8.12}$$

while a strictly bilocal-tomographic theory, as shown by Hardy and Wootters in [71], must satisfy

$$D_{\mathrm{AB}} > D_{\mathrm{A}} D_{\mathrm{B}}, \tag{8.13}$$

$$D_{\mathrm{ABC}} \leqslant D_{\mathrm{A}} D_{\mathrm{B}} D_{\mathrm{C}} + \widetilde{D}_{\mathrm{AB}} D_{\mathrm{C}} + \widetilde{D}_{\mathrm{BC}} D_{\mathrm{A}} + \widetilde{D}_{\mathrm{CA}} D_{\mathrm{B}}, \tag{8.14}$$

[2]Remember that a set of effects E is called separating for a set of states S if any two states of S are discriminated by an effect of E (see Definition 2.7).

with

$$\widetilde{D}_{\mathrm{AB}} := D_{\mathrm{AB}} - D_{\mathrm{A}} D_{\mathrm{B}}. \tag{8.15}$$

In particular a bilocal-tomographic theory can saturate the bound (8.14) and in that case we call it *maximally bilocal-tomographic*

Definition 8.4 (Maximal bilocal tomography) *When the upper bound in Eq.* (8.14) *is saturated we say that the theory is maximally bilocal-tomographic, and it requires all* 2*-local effects to separate multipartite states.*

In principle the constraints on the systems of a probabilistic theory could affect its degree of holism. Taking a superselection of the theory $\mathcal{T}$, we explore the relation between the number of linear constraints on its systems and the degree of holism of the constrained theory $\bar{\mathcal{T}}$.

Physically it is interesting to consider superselected versions of QT and we consider the case in which the original theory $\mathcal{T}$ is local-tomographic. In this scenario we have the following striking relation:

Proposition 8.2 (Superselection-holism trade-off) *Let $\bar{\mathcal{T}}$ be a superselection of a local-tomographic theory $\mathcal{T}$. Then we have*

(i) Minimal superselection $\Rightarrow$ $\bar{\mathcal{T}}$ maximally bilocal-tomographic.

(ii) Maximal superselection $\Rightarrow$ $\bar{\mathcal{T}}$ local-tomographic.

Proof. First we prove the implication (i). The superselected theory $\bar{\mathcal{T}}$ is maximally bilocal-tomographic if it saturates the bound (8.14), namely

$$D_{\overline{\mathrm{ABC}}} = D_{\bar{\mathrm{A}}} D_{\bar{\mathrm{B}}} D_{\bar{\mathrm{C}}} + \widetilde{D}_{\overline{\mathrm{AB}}} D_{\bar{\mathrm{C}}} + \widetilde{D}_{\overline{\mathrm{BC}}} D_{\bar{\mathrm{A}}} + \widetilde{D}_{\overline{\mathrm{CA}}} D_{\bar{\mathrm{B}}}. \tag{8.16}$$

We prove this equality evaluating the LHS and the RHS of the equation and imposing the minimal superselection given by the lower bound (8.7)

$$V^{\sigma}_{\mathrm{AB}} = D_{\mathrm{A}} V^{\sigma}_{\mathrm{B}} + D_{\mathrm{B}} V^{\sigma}_{\mathrm{A}} - 2 V^{\sigma}_{\mathrm{A}} V^{\sigma}_{\mathrm{B}}. \tag{8.17}$$

LHS: using Eq. (8.5) we have

$$D_{\overline{\mathrm{ABC}}} = D_{\mathrm{ABC}} - V^{\sigma}_{\mathrm{ABC}}, \tag{8.18}$$

and taking the partition $\overline{\mathrm{ABC}} = \bar{\mathrm{A}}(\overline{\mathrm{BC}})$, the minimal superselection constraint (8.17) gives

$$D_{\overline{\mathrm{ABC}}} = D_{\mathrm{ABC}} - (D_{\mathrm{A}} V^{\sigma}_{\mathrm{BC}} + D_{\mathrm{BC}} V^{\sigma}_{\mathrm{A}} - 2 V^{\sigma}_{\mathrm{A}} V^{\sigma}_{\mathrm{BC}}). \tag{8.19}$$

Now we use again (8.17) for expanding V^{σ}_{BC} and V^{σ}_{BC} in the last equation and after a simple manipulation we get

$$\begin{aligned} D_{\overline{\mathrm{ABC}}} = D_{\mathrm{A}} D_{\mathrm{B}} D_{\mathrm{C}} &- (D_{\mathrm{A}} D_{\mathrm{B}} V^{\sigma}_{\mathrm{C}} + D_{\mathrm{A}} D_{\mathrm{C}} V^{\sigma}_{\mathrm{B}} + D_{\mathrm{B}} D_{\mathrm{C}} V^{\sigma}_{\mathrm{A}}) \\ &+ 2(D_{\mathrm{A}} V^{\sigma}_{\mathrm{B}} V^{\sigma}_{\mathrm{C}} + D_{\mathrm{A}} V^{\sigma}_{\mathrm{C}} V^{\sigma}_{\mathrm{B}} + D_{\mathrm{B}} V^{\sigma}_{\mathrm{C}} V^{\sigma}_{\mathrm{A}}) \\ &- 4 V^{\sigma}_{\mathrm{A}} V^{\sigma}_{\mathrm{B}} V^{\sigma}_{\mathrm{C}}, \end{aligned} \tag{8.20}$$

where we used the identity $D_{\mathrm{XY}} = D_{\mathrm{X}} D_{\mathrm{Y}}$ (remember that the non-superselected theory is local-tomographic).

RHS: After using the identities (8.15) and (8.5) we exploit the minimal superselection constraint (8.17). Finally we use the local tomography condition $D_{\mathrm{XY}} = D_{\mathrm{X}} D_{\mathrm{Y}}$ for the non-superselected theory and by direct computation we get exactly the expression in (8.20) showing that the superselected theory is maximally bilocal-tomographic.

Let us now consider the implication (ii). The bound (8.8), corresponding to the maximal superselection, has been derived in Proposition 8.1 using the general constraint (8.9) for the superselected theory $\bar{\mathcal{T}}$, that is $D_{\overline{\mathrm{AB}}} \geqslant D_{\bar{\mathrm{A}}} D_{\bar{\mathrm{B}}}$. On the other hand when the last bound is saturated

$$D_{\overline{\mathrm{AB}}} = D_{\bar{\mathrm{A}}} D_{\bar{\mathrm{B}}}, \tag{8.21}$$

the resulting theory $\bar{\mathcal{T}}$ is local-tomographic. ■

While the relation (i) of Proposition 8.2 is quite surprising the relation (ii) is very intuitive. Suppose to have a local-tomographic theory $\mathcal{T}$ and to impose a superselection rule on its elementary systems. This in general breaks down the local tomography of the theory with the number of local effects of $\bar{\mathcal{T}}$ reduced by the linear constraints. However, if we remove from the bipartite systems all the states that cannot be tomographied using the superselected local effects, we are left with a theory $\bar{\mathcal{T}}$ which is again local-tomographic.

Remark 8.1 (n-local-tomographic superselected theories) In Proposition 8.2 we have considered the extremal cases of minimal and maximal superselection which respectively lead to theories having degree of holism $n = 2$ and $n = 1$. On the other hand between these two cases there is a full range of possible constraints where one can find superselected theories with any degree of holism.

In the following we focus on the Fermionic theory which can be regarded as the QT with a minimal superselection on the elementary qubit systems.

8.2 The Fermionic Quantum Theory

We consider a network where each site, called local Fermionic mode (LFM), can be either empty or occupied by a spinless Fermion. In the following we will denote the multipartite Fermionic systems as

$$1_{\mathrm{F}}, 2_{\mathrm{F}}, \ldots, \mathrm{N}_{\mathrm{F}}, \ldots. \tag{8.22}$$

We define creation and annihilation operators $\psi_i^\dagger$ and ψ_i—where i labels the i-th LFM on the network—as operators satisfying the canonical anticommutation relations

$$[\psi_i^\dagger, \psi_j]_+ = \delta_{ij} I, \qquad [\psi_i, \psi_j]_+ = 0, \tag{8.23}$$

and we will denote by FQT the probabilistic theory having LFMs as elementary systems.

The vacuum state of the theory $|\Omega\rangle$ corresponds to the unique joint eigenvector $|\Omega\rangle$ of the operators $\psi_j^\dagger \psi_j$ with zero eigenvalue zero. It corresponds to all modes unoccupied and is then annihilated by the field operators

$$\psi_i |\Omega\rangle = 0 \qquad \forall \psi_i. \tag{8.24}$$

The vacuum state of the single LFM will be denoted by $|0\rangle$, with $\psi^\dagger |0\rangle = |1\rangle$, while the vacuum state $|\Omega\rangle$ of $N < \infty$ LFMs is unique and it is given by

$$|\Omega\rangle = \otimes_{i=1}^N |0\rangle_i \,. \tag{8.25}$$

A representation of the Fermionic algebra can be given fixing an order $1, 2, \dots N$ of the LFMs and considering the orthonormal basis for $\mathbb{C}^{2^N}$ given by

$$|q_1, q_2, \dots, q_N\rangle := \psi_1^{\dagger q_1} \psi_2^{\dagger q_2} \dots \psi_N^{\dagger q_N} |\Omega\rangle \tag{8.26}$$

with $q_i = 0, 1$ for $i = 1, \dots, N$ corresponding to the occupation number at the i-th site. In this representation, corresponding to the Jordan-Wigner construction [54], the Hilbert space $\mathcal{H}$ of N LFMs is then the Fock space spanned by the basis vectors $|q_1, \dots, q_N\rangle$,

$$\mathcal{H} = \mathrm{span}(|q_1, \dots, q_N\rangle \,;\, q_i \in \{0, 1\}). \tag{8.27}$$

We now can define locality of transformations. We say that the transformation of the $\mathrm{N_F}$ system is local on the subsystem $\mathrm{M_F}$ with $M < N$ if the Kraus operators belong the representation of the field algebra of $\mathrm{M_F}$.

One can decompose the Fock Hilbert space (8.27) in the direct sum

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1, \tag{8.28}$$

where $\mathcal{H}_0$ and $\mathcal{H}_1$ are respectively the eigenspaces of the parity operator

$$P = \tfrac{1}{2}(I + \prod_{i=1}^n (\psi_i \psi_i^\dagger - \psi_i^\dagger \psi_i)) \tag{8.29}$$

whose eigenvalues $p = 0, 1$, correspond respectively to an even/odd total occupation number.

The main difference between FQT and QT, due to the anticommuting nature of Fermionic field, is that in the Fermionic case the parity is conserved [32,206], namely any physical state or observable commutes with the operator P. This implies that the linear space of states and effects

$$\mathrm{St}_{\mathbb{R}}(\mathrm{N_F}) = \mathrm{Eff}_{\mathbb{R}}(\mathrm{N_F}) \tag{8.30}$$

corresponds to the operator space spanned by products of an even number of ψ_i and $\psi_i^\dagger$. In the following we will seek out the consequences of this parity prescription on the structure of FQT.

8.2.1 Maximal bilocal tomography

We denote by $\mathrm{N_Q}$ the multipartite system of N qubits, having Hilbert dimension $d_{\mathrm{N_Q}} = 2^N$. The linear space of states is the set of $2^N \times 2^N$ Hermitian matrixes

$$\mathrm{St}_{\mathbb{R}}(\mathrm{N_Q}) = \mathrm{Eff}_{\mathbb{R}}(\mathrm{N_Q}) = \mathrm{Herm}(\mathbb{C}^{2^N}) \tag{8.31}$$

and according to Eq. (2.17) we have

$$D_{\mathrm{N_Q}} = d_{\mathrm{N_Q}}^2 = 2^{2N}. \tag{8.32}$$

For example the system 1_{Q} corresponds to the qubit whose pure states are

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,, \qquad |\alpha|^2 + |\beta|^2 = 1, \tag{8.33}$$

and whose dimension is

$$D_{1_{\mathrm{Q}}} = 2^2 = 4. \tag{8.34}$$

Notice that QT is local-tomographic

$$D_{\mathrm{N_Q}} = D_{1_{\mathrm{Q}}}^N = 2^{2N}. \tag{8.35}$$

Consider now the Fermionic case. Here the single-LFM system 1_{F} has only two possible pure states $|0\rangle, |1\rangle$ and corresponds to the classical bit. Indeed $|0\rangle$ and $|1\rangle$ have different parity and one cannot consider their superpositions (8.33). Due to the parity prescription the set of states is a direct sum of two disjoint components (see Eq. (8.28)) with the linear space of states for the system of N LFMs given by

$$\mathrm{St}_{\mathbb{R}}(\mathrm{N_F}) = \mathrm{Eff}_{\mathbb{R}}(\mathrm{N_F}) = \mathrm{Herm}(\mathbb{C}^{2^{N-1}}) \oplus \mathrm{Herm}(\mathbb{C}^{2^{N-1}}), \tag{8.36}$$

namely the direct sum of two copies of the state-space of $N-1$ qubits. It is easy to see that the dimension of the system N_{F} is

$$D_{\mathrm{N_F}} = 2^{2N-1} = \frac{1}{2} D_{\mathrm{N_Q}}. \tag{8.37}$$

Now we can prove that FQT is a minimal superselection of QT:

Proposition 8.3 (Minimal superselection) *The* FQT*, namely the probabilistic theory having* LFMs 1_{F} *as elementary systems, is a minimal superselection of* QT *with the following constraints on the qubit system* 1_{Q}

$$\mathrm{St}(\sigma(1_{\mathrm{Q}})) \coloneqq \{\rho \in \mathrm{St}(1_{\mathrm{Q}});\, \mathrm{Tr}[\sigma_x \rho] = \mathrm{Tr}[\sigma_y \rho] = 0\}. \tag{8.38}$$

Proof. We have already noticed that, due to the parity prescription, the single Fermionic mode 1_{F} has only two pure states $|0\rangle, |1\rangle$. Then the density matrices $\rho \in \mathrm{St}(1_{\mathrm{F}})$ must be diagonal

$$\mathrm{Tr}[\sigma_x \rho] = \mathrm{Tr}[\sigma_y \rho] = 0, \tag{8.39}$$

showing that the superselection on the elementary systems is as in Eq. (8.38) with

$$1_{\mathrm{F}} = \sigma(1_{\mathrm{Q}}), \qquad D_{1_{\mathrm{F}}} = D_{1_{\mathrm{Q}}} - V^{\sigma}_{1_{\mathrm{Q}}}, \qquad V^{\sigma}_{1_{\mathrm{Q}}} = 2. \tag{8.40}$$

Now we have to show that the whole FQT is built bottom-up extending in the minimal way the constraint (8.38) on the composite systems. According to Definition 8.2 we can simply check that FQT achieves the lower bound (8.7),

$$V^{\sigma}_{\mathrm{N_Q M}_Q} = D_{\mathrm{N_Q}} V^{\sigma}_{\mathrm{M}_Q} + D_{\mathrm{N_Q}} V^{\sigma}_{\mathrm{M}_Q} - 2 V^{\sigma}_{\mathrm{N_Q}} V^{\sigma}_{\mathrm{M}_Q}. \tag{8.41}$$

Observing that by definition it is $D_{\mathrm{N_F}} = D_{\mathrm{N_Q}} - V_{\mathrm{N_Q}}$, using Eq. (8.37) we have

$$D_{\mathrm{N_F}} = V_{\mathrm{N_Q}} = \frac{1}{2} D_{\mathrm{N_Q}}, \tag{8.42}$$

and Eq. (8.41) is trivially satisfied. ■

Therefore, as a corollary of Proposition 8.2, FQT is maximally bilocal-tomographic. This can also be verified observing that for elementary Fermionic systems the bounds (8.13) and (8.14) are satisfied

$$8 = D_{2_{\mathrm{F}}} > D^2_{1_{\mathrm{F}}} = 4, \tag{8.43}$$

$$D_{3_F} = D^3_{1_{\mathrm{F}}} + 3 \widetilde{D}_{2_{\mathrm{F}}} D_{1_{\mathrm{F}}} = 16, \tag{8.44}$$

where $\widetilde{D}_{2_{\mathrm{F}}} = D_{2_{\mathrm{F}}} - D^2_{1_{\mathrm{F}}} = 4$ is the dimension of the non-local component of 2_{F}. The theory is maximally bilocal-tomographic as one can check counting the number of independent local and 2-local effects for the system $\mathrm{N_F}$ of N LFMs and observing that it is exactly the dimension of $\mathrm{N_F}$

$$\sum_{k=0}^{\lfloor N/2 \rfloor} \binom{N}{2k} D^{N-2k}_{1_{\mathrm{F}}} \widetilde{D}^{k}_{2_{\mathrm{F}}} = 2^{2N-1} = D_{\mathrm{N_F}}. \tag{8.45}$$

We emphasize that FQT provides an example of a bilocal-tomographic theory whose systems do not satisfy the dimensional prescription in Ref. [71]: after showing that the dimension of the non-local component of a bipartite system $\widetilde{D}_{\mathrm{AB}} = D_{\mathrm{AB}} - D_{\mathrm{A}} D_{\mathrm{B}}$ can be factorized as $\widetilde{D}_{\mathrm{AB}} = L_{\mathrm{A}} L_{\mathrm{B}}$, and assuming that the two functions

$$D_{\mathrm{A}} + L_{\mathrm{A}}, \qquad D_{\mathrm{A}} - L_{\mathrm{A}}, \tag{8.46}$$

are strictly increasing functions of the number of perfectly discriminable states d_{A}, the authors prove that in a bilocal-tomographic theory, the dimension of the system A must be

$$D_{\mathrm{A}} = \frac{1}{2}(d_{\mathrm{A}}^{r} + d_{\mathrm{A}}^{s}), \tag{8.47}$$

for some integers r, s satisfying $r \geqslant s > 0$. This is not true for the Fermionic computation where for example $D_{2_{\mathrm{F}}} = 8$ cannot be achieved in this way. The strict monotonicity of the function $D_{\mathrm{A}} - L_{\mathrm{A}}$ is too restrictive and excludes the Fermionic case, which has $D_{\mathrm{N_F}} - L_{\mathrm{N_F}} = 0$ for any $\mathrm{N_F}$, from the set of admissible bilocal-tomographic theories.

Remark 8.2 (The Real Quantum Theory case) It is worth mentioning that FQT is not the unique minimal superselection of QT. Another example is given by RQT (see Section 2.2.3). We denote by $\mathrm{N_R}$ the multipartite system of N *rebits*, with Hilbert space having dimension $d_{\mathrm{N_R}} = 2^N$. In the literature RQT is defined [71] as the QT of real matrices and as such, the set of states $\mathrm{St}_{\mathbb{R}}(\mathrm{N_R})$ corresponds to the symmetric component of the quantum counterpart $\mathrm{St}_{\mathbb{R}}(\mathrm{N_Q})$ (8.31) and has dimension

$$D_{\mathrm{N_R}} = d_{\mathrm{N_R}}(d_{\mathrm{N_R}} + 1)/2. \tag{8.48}$$

The RQT is actually a superselection of QT (see Definition 8.1) since the reality of quantum states is given by the linear constraint $\rho - \rho^T = 0$, with T denoting transposition with respect to a fixed basis taken as real. Thus one has $\mathrm{N_R} = \sigma(\mathrm{N_Q})$ where the number of constraints for the system $\mathrm{N_R}$ is given by

$$V^{\sigma}_{\mathrm{N_R}} = D_{\mathrm{N_Q}} - D_{\mathrm{N_R}} = \tfrac{1}{2} d_{\mathrm{N_R}} d_{\mathrm{M}_R}(d_{\mathrm{N_R}} d_{\mathrm{M}_R} - 1). \tag{8.49}$$

One can easily check that RQT is minimally superselected, since the number of constraints for the composite systems

$$V^{\sigma}_{\mathrm{N_R M}_R} = \tfrac{1}{2} d_{\mathrm{N_R}} d_{\mathrm{M}_R}(d_{\mathrm{N_R}} d_{\mathrm{M}_R} - 1) \tag{8.50}$$

saturates the lower bound (8.8). Then according to Proposition 8.2 RQT is maximally bilocal-tomographic, as pointed out in [71].

Notice that for the rebit system 1_{R} the linear constraint is

$$\mathrm{Tr}[\sigma_y \rho] = 0, \tag{8.51}$$

with

$$1_{\mathrm{R}} = \sigma(1_{\mathrm{Q}}), \qquad D_{1_{\mathrm{R}}} = D_{1_{\mathrm{Q}}} - V^{\sigma}_{1_{\mathrm{Q}}} \qquad V^{\sigma}_{1_{\mathrm{Q}}} = 1, \tag{8.52}$$

and the full RQT is achieved taking the minimal extension of this constraint to the composite systems.

Remark 8.3 (n-local-tomographic superselected theory) As noticed in Remark 8.1 a theory with V_{AB}^{σ} strictly included in the bounds of Eqs. (8.7) and (8.8) can have an arbitrary degree of holism. Due to the parity constraint, the FQT retains only superpositions of pure states with total occupation number equal modulo 2. If instead we allow only superpositions with total occupation number equal modulo 3 we get a theory that is k-local-tomographic with $k \geqslant 3$.

8.2.2 Fermionic entanglement

Now we focus on another characteristic trait of QT—say the entanglement—and explore how its properties change in FQT as a consequence of the parity superselection.

In Section 2.4 we have discussed the problem of defining the entanglement in a general probabilistic theory and we stressed that entanglement must be quantified in operational terms. For bipartite states in QT all measures of entanglement refer to a standard unit of entanglement, the *ebit* (which is the amount of entanglement of a singlet state (2.37)), and the so called *entanglement of formation* is the number of ebits that are needed to achieve the state by LOCC (see Section 2.4.1). A full theory of entanglement would require a complete analysis of the transformations of states under LOCC, which is a very difficult task form many reasons, *e.g.* a complet characterization of LOCC is still lacking also in QT. However, independently of such analysis, we can show that in FQT one can assess features that are very different from those of entanglement in QT. These are:

(i) Existence of mixed states with maximal entanglement of formation.
(ii) Need of MES (see Section 2.4.3) for bipartite states.
(iii) Violation of monogamy of entanglement.

Similar features have been shown for RQT in Refs. [208] (for (i)) and [209] (for (iii)). These phenomena are due to the fact that both theories are superselected versions of QT. One would conjecture that these features may be related to the non-local-tomographic nature of the theories, however, this remains an open issue.

The Fermionic entanglement of formation and concurrece

In Section 2.4.1 we have defined the entanglement of formation and the concurrence in the usual quantum scenario and we refer to that Section for the details regarding the original definitions and the relative literature. For the convenience of the reader we remind that in QT, the entanglement of formation for a generally mixed state ρ is defined as

$$E(\rho) \coloneqq \min_{\mathcal{D}_\rho} \sum_i p_i E(|\Psi_i\rangle), \tag{8.53}$$

where $\mathcal{D}_\rho \coloneqq \{\{p_i, |\Psi_i\rangle\} \mid \rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|\}$ is the set of all the pure decompositions of the mixed state ρ and $E(|\Psi\rangle)$ is the entanglement entropy of the pure state $|\Psi\rangle$, while for a mixed state ρ of two qubits the quantum concurrence $C(\rho)$ is defined as

$$C(\rho) \coloneqq \min_{\mathcal{D}_\rho} \sum_i p_i C(|\Psi_i\rangle), \tag{8.54}$$

with the formula for the concurrence of a pure state given in Eq. (2.52).

We know from Section 2.4.2 that the entanglement of formation extends to mixed states the operational meaning of the entanglement entropy: if two parties wants to prepare n copies of a pure state $|\Psi_{\mathrm{AB}}\rangle$ by means of LOCC, they need $nE(|\Psi_{\mathrm{AB}}\rangle)$ pairs of singlet states (2.37) in order to achieve their task. We also know that both the entanglement of formation and the concurrence are zero if and only if the state ρ is separable, and for two qubits they reach the maximum value 1 if and only if ρ is a maximally entangled state in the sense of MES (see Section 2.4.3), namely that it can be used to reach any other state by LOCC operations.

The entanglement of formation and the concurrence can be specialized to FQT by the convex-roof extensions[3] of the relative quantum quantities in Eqs. (8.53) and (8.54) as follows

$$E_{\mathrm{F}}(\rho) \coloneqq \min_{\mathcal{D}^{\mathrm{F}}_\rho} \sum_i p_i E(|\Psi_i\rangle), \tag{8.55}$$

$$C_{\mathrm{F}}(\rho) \coloneqq \min_{\mathcal{D}^{\mathrm{F}}_\rho} \sum_i p_i C(|\Psi_i\rangle), \tag{8.56}$$

where $\mathcal{D}^{\mathrm{F}}_\rho$ is the set of all the pure decompositions of ρ that satisfy the parity superselection rule. Notice that on pure states $|\Psi\rangle$ we have the same entanglement entropy $E(|\Psi\rangle)$ and concurrence $C(|\Psi\rangle)$ of QT. The same extension has been considered for RQT in Ref. [208] considering the decompositions $\mathcal{D}^{\mathrm{R}}_\rho$ on real states.

In the following proposition we show that, due to the parity superselection, the Fermionic entanglement of formation and concurrence can be expressed in terms of the usual quantum quantities. Accordingly we can recover a lower bound on the operational asymptotic measure of entanglement in the Fermionic case.

Proposition 8.4 (Fermionic measure of entanglement) *In* FQT *the entanglement of formation* (8.55) *and the concurrence* (8.56) *of a mixed state* ρ *are*

$$E_{\mathrm{F}}(\rho) = p_0 E(\rho_0) + p_1 E(\rho_1), \tag{8.57}$$

$$C_{\mathrm{F}}(\rho) = p_0 C(\rho_0) + p_1 C(\rho_1), \tag{8.58}$$

[3]Both the entanglement of formation (8.53) and the concurrence (8.53) are defined as the minimum over all the convex decompositions of ρ. In this way the entanglement of formation and the concurrence of pure states are extended to the mixed ones. This is a special case of the so called convex-roof extension.

where

$$\rho = p_0\rho_0 + p_1\rho_1, \qquad p_0 + p_1 = 1, \quad \mathsf{Supp}(\rho_i) \subseteq \mathcal{H}_i, \tag{8.59}$$

is the unique parity-decomposition of ρ and $E(\rho_i)$ is the usual quantum entanglement of formation. Moreover (8.57) *provides a lower bound for the for the operational asymptotic measure of entanglement in the bipartite case.*

Proof. Consider first the entanglement of formation. A Fermionic density matrix ρ can always be decomposed in its even and odd parity components as in Eq. (8.59) and this decomposition must be unique since $\mathsf{Supp}(\rho_0)$ and $\mathsf{Supp}(\rho_1)$ are orthogonal. Thus we have

$$E_{\mathrm{F}}(\rho) = E_{\mathrm{F}}(p_0\rho_0 + p_1\rho_1), \tag{8.60}$$

and by convexity of E_{F} (see Eq.(8.55)), one has

$$E_{\mathrm{F}}(\rho) \leqslant p_0 E_{\mathrm{F}}(\rho_0) + p_1 E_{\mathrm{F}}(\rho_1). \tag{8.61}$$

Now suppose by absurd that $E_{\mathrm{F}}(\rho) < p_0 E_{\mathrm{F}}(\rho_0) + p_1 E_{\mathrm{F}}(\rho_1)$ and let $\{\bar{p}_i, |\bar{\Psi}_i\rangle\}$ be the pure decomposition achieving the minimum of $E_{\mathrm{F}}(\rho)$. Exploiting the uniqueness of the parity-decomposition (8.59) we have

$$\begin{aligned} E_{\mathrm{F}}(\rho) &= \sum_i \bar{p}_i E_{\mathrm{F}}(|\bar{\Psi}_i\rangle) = \\ &= p_0 \sum_{\text{even}} \frac{\bar{p}_i^0}{p_0} E(|\bar{\Psi}_i^0\rangle) + p_1 \sum_{\text{odd}} \frac{\bar{p}_i^1}{p_1} E(|\bar{\Psi}_i^1\rangle) < p_0 E_{\mathrm{F}}(\rho_0) + p_1 E_{\mathrm{F}}(\rho_1) \end{aligned} \tag{8.62}$$

which is clearly an absurd since $E_{\mathrm{F}}(\rho_i)$ are defined as minimum among all the (admissible) pure decompositions. Moreover, since the convex decomposition of a state ρ can only contain states in the support of ρ we have

$$\mathcal{D}^{\mathrm{F}}_{\rho_i} \equiv \mathcal{D}_{\rho_i}, \tag{8.63}$$

which gives Eq. (8.57). The same proof applies to Eq. (8.58) for the Fermionic concurrence.

In order to prove that (8.57) provides a lower bound for the operational asymptotic measure of entanglement we observe that the Fermionic entanglement of formation of ρ is expressed as average of the respective quantum quantities, which have a clear operational meaning. We know that in the asymptotic limit, allowing all quantum bipartite LOCCs operations, the two parties can generate n copies of ρ using $n(p_0 E_{\mathrm{F}}(\rho_0) + p_1 E_{\mathrm{F}}(\rho_1))$ singlets (see Section 2.4.2). Since bipartite fermionic LOCCs are all admissible quantum LOCCs (one cannot do better in the Fermionic case) and any resource-state in the fermionic MES has a quantum entanglement entropy smaller than (or equal to) one we can conclude that (8.57) is a lower bound for the fermionic entanglement.

We noticed that in QT the entanglement of formation and the concurrence are connected as in Eq. (2.47) by the relation $E(\rho) = \mathcal{E}(C(\rho))$, with $\mathcal{E}(x) \coloneqq h(\frac{1+\sqrt{1-x^2}}{2})$ and h the binary Shannon entropy. This is true also in RQT [208] but not in the FQT case. Nevertheless we have that

$$E_{\mathrm{F}}(\rho) \geqslant \mathcal{E}(C_{\mathrm{F}}(\rho)), \tag{8.64}$$

and for a maximally entangled state Φ it is $E_{\mathrm{F}}(\Phi) = \mathcal{E}(C_{\mathrm{F}}(\Phi)) = 1$. According to Proposition 8.4 when $E_{\mathrm{F}}(\Phi) = \mathcal{E}(C_{\mathrm{F}}(\Phi)) = 1$ the entanglement of formation corresponds to the operational measure of entanglement.

Mixed states with maximal entanglement of formation

Using the quantities E_{F} and C_{F} we can show that in FQT there exist mixed states with maximal entanglement of formation. Consider the following state

$$\Phi \coloneqq \tfrac{1}{4}\left(I \otimes I + \sigma_x \otimes \sigma_x\right), \tag{8.65}$$

which corresponds to the equal mixture of the Fermionic pure states

$$|\Psi_0\rangle = \tfrac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right), \qquad |\Psi_1\rangle = \tfrac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right). \tag{8.66}$$

Despite being mixed we have that

$$E_{\mathrm{F}}(\Phi) = \frac{1}{2}E(|\Psi_0\rangle) + \frac{1}{2}E(|\Psi_1\rangle) = 1, \tag{8.67}$$

$$C_{\mathrm{F}}(\Phi) = \frac{1}{2}C(|\Psi_0\rangle) + \frac{1}{2}C(|\Psi_1\rangle) = 1, \tag{8.68}$$

i.e. Φ has maximal entanglement of formation in the operational sense. It is easy to see that Φ cannot be written in a separable form, namely no mixture of parity-superselected product states can produce the above state Φ. Indeed according to Proposition 8.3 any mixture ρ_{sep} of parity-superselected product states must satisfy the constraints

$$\mathrm{Tr}[\rho_{sep}(\sigma_x \otimes \sigma_x)] = \mathrm{Tr}[\rho_{sep}(\sigma_y \otimes \sigma_y)] = 0, \tag{8.69}$$

but the state ρ does not have zero trace with $(\sigma_x \otimes \sigma_x)$ and is then non separable in FQT. We can obtain other states with the same property of Φ (8.65) replacing σ_x in with any linear combination of σ_x and σ_y.

Also RQT has mixed maximally entangled states. Remark 8.2 shows that RQT have the linear constraint $\mathrm{Tr}[\sigma_y \rho] = 0$ for one rebit, and the same argument presented in FQT holds in RQT with σ_x replaced by σ_y [208] in the state (8.65).

In QT the state Φ (8.65) is separable since it can be regarded as the equal mixture of the pure product states

$$|+\rangle|+\rangle, \quad |-\rangle|-\rangle, \qquad |\pm\rangle = \tfrac{1}{\sqrt{2}}\left(|0\rangle \pm |1\rangle\right) \tag{8.70}$$

which gives

$$E(\Phi) = C(\Phi) = 0. \tag{8.71}$$

Such a decomposition, however, is not allowed in the Fermionic case, because the states $|\pm\rangle$ violate the parity superselection rule (and in the RQT case, because $|\pm\rangle$ violate the real superselection rule).

Maximally entangled sets for bipartite states

The state Φ (8.65), despite having maximum entanglement of formation, is not maximally entangled in the sense that it can be transformed by LOCC into any other state. It actually happens that for two LFMs the concept of maximally entangled state under LOCC has to be superseded by the concept of MES introduced in Section 2.4.3. In QT this only happens for n-partite quantum entanglement with $n \geqslant 3$ [90] while in the bipartite case a single maximally entangled state can be found (as proved in Section 2.4.3).

In the Fermionic theory the single LFM is a bit system and the only LOCCs operations that one can perform are of the kind

$$I \cdot I, \quad \sigma_1 \cdot \sigma_1, \quad \sigma_2 \cdot \sigma_2, \quad \sigma_3 \cdot \sigma_3, \tag{8.72}$$

which do not allow to transform the pure states $|0\rangle$, $|1\rangle$ in any of their superpositions. This implies that starting from a pure bipartite state $|\Phi\rangle$ having Schmidt decomposition

$$|\Phi\rangle = \alpha\,|00\rangle + \beta\,|11\rangle \tag{8.73}$$

one can never change by LOCC the magnitude of the coefficients α and β but can at most change the parity sector via $\sigma_x \cdot \sigma_x$ or $\sigma_y \cdot \sigma_y$ on a single LFM, or apply an arbitrary relative phase. The only case in which one can vary the magnitude of the Schmidt coefficients is the LOCC generation of an arbitrary factorized state: Alice measures its LFM in the computational base $\{|0\rangle, |1\rangle\}$ and according to the result classically communicates to Bob which local operation he has to apply on his LFM, the identity $I \cdot I$ or the bit flip $\sigma_x \cdot \sigma_x$. Clearly, one cannot do the opposite, namely transforming a factorized state into a superimposed one.

Examples of MES for two LFMs are the set of all even non-factorized pure states with positive coefficients, and the set of all odd non-factorized pure states with positive coefficients.

Violation of entanglement monogamy

Now it is very easy to see that the Fermionic entanglement is not monogamous in the sense of the entanglement of formation.

Consider the 3_{F} pure state

$$|\Phi_p\rangle := \tfrac{1}{2}(|000\rangle + |110\rangle + |011\rangle + |101\rangle). \tag{8.74}$$

If we trace the state $|\Phi_p\rangle$ over any one of the three LFMs we find that the reduced bipartite state is the mixed state Φ of Eq. (8.65) which has maximal entanglement of formation. Therefore, in the FQT, as well as in RQT [209], the amount of entanglement can be totally shared by each pair systems, a feature forbidden in QT.

8.2.3 Possible physical effects of monogamy violation

Regarded as a probabilistic theory the FQT results to be very different from QT and this could have interesting physical consequences, *i.e.* in the physics of black holes information.

In 2007 Preskill and Hayden [210] showed that old enough black holes—say half of the black hole's qubits have already been radiated away—behave as mirrors for the information thrown into them, reemitting it as Hawking radiation. In a quantum scenario, they evaluated the "number of emitted qubits" necessary for recovering the information previously entered the black hole: if k qubits entangled with a reference are thrown into the old black hole, for any positive c, once $k+c$ qubits are reemitted the transevent horizon entanglement between the reference and the interior subsystems will have virtually vanished, with it appearing instead (with a fidelity of at least $1-2^{-c}$) as entanglement between the reference and the outgoing radiation. Another interesting paper about the same effect is [211]. Later in [212] the mirroring behavior has been generalized in a scenario beyond QT but with local discriminability showing how in that case the black hole mirrors the information even faster than in the quantum case.

All the previous results are grounded on the so called *decoupling theorem* (see Ref. [213] for an exhaustive review on the subject) which exploit the monogamy of the entanglement. One could consider the same process in the Fermionic case where the entanglement is not monogamous.

Monogamy plaies a crucial role also in the so called "firewall paradox" for black holes and some authors tried to solve it by informational argument, see for example the recent work [214] of Erik and Herman Verlinde. Again the FQT is a concrete physical model having a non monogamous entanglement and should be investigated as a possible way of solving the paradox.

CHAPTER 9

Conclusions and future perspectives

The aim of this thesis was to present in a coherent way the motivations and the consequences of describing physical dynamics as the processing of quantum information. The problem connects a broad spectrum of notions in the fields of Quantum Information Theory, Foundations of Quantum Theory (QT), Quantum Field Theory (QFT) and physics beyond the Standard Model. This motivated us to consider the most elementary quantum framework satisfying few informational-theoretical principles that an informational theory of dynamics should obey and to procede step by step in the exploration of their consequences. Since any physical evolution is described by a quantum algorithm the two principles at the basis of our approach have been the finite complexity of the algorithm and the universality of the "physical law" expressed by the algorithm. The simplest quantum dynamical model carrying these two features is the Quantum Cellular Automaton (QCA), namely a set of quantum systems with minimal Hilbert-space dimension interacting with a finite number of neighbors on a homogeneous causal network. These considerations led to the idea of considering QCAs as a microscopical description of quantum fields and to consider spacetime as emergent from the automata discrete causal network.

The research program of this presentation started from the derivation in Chapter 4 of a QCA describing the Dirac evolution in one dimension. It has been shown that assuming the automaton lattice at a scale much smaller than the usual scale of particle physics—say the Planck scale—the Dirac equation

is recovered in the large-scale limit of small wave-vectors and masses. We have seen how the QCA represents a novel discretization method for QFT. The automaton provides a discrete unitary evolution and does not correspond to a finite-difference version of the usual QFT Lagrangians (or Hamiltonians). A first implication is that the QCA avoids the fermion doubling phenomenon affecting the Lattice Field Theories based on finite-difference differential equations.

Chapter 5 was devoted to the Dirac QCA dynamics in the one-particle sector, recovering many characteristic traits of the Dirac equation. An approximate dispersive differential equation describes analytically the automaton free evolution with drift and diffusion coefficients depending on the particle's momentum and mass. While the coefficients coincide with the Dirac (Schrödinger) ones in the relativistic (non-relativistic) limit of the automaton, they significantly differ for high wave-vectors or masses, providing the first phenomenological corrections to QFT. In the free case the Dirac QCA exhibits the Zitterbewegung of the average position due to the interference between the positive- and the negative-frequency eigenstates of the unitary operator. The frequency and the amplitude of the trembling have been evaluated and are compatible with the Dirac equation predictions. Adding a square potential barrier it was possible to study the scattering of the Dirac QCA. This resultes in another Dirac typical feature, that is the revival of the transmitted wave for an high enough potential barrier also known as "Klein paradox". The interpretation of this phenomenon in the automaton scenario is very clear and the revival of the transmitted wave occurs for any particle wave-vector and mass. Indeed the values of the potential corresponding to a total reflection are in a bounded region whose width is equal to the gap between the positive and the negative frequencies (energies) of the QCA.

The experimental verification of the Dirac QCA phenomenology has been rigorously treated in Chapter 6 in a quantum informational scenario. We have considered the QCA and the usual Dirac evolution as two black boxes evaluating the optimal probability of discrimination between them. The optimality is defined as the minimization of the error probability with respect to all possible preparations and measurements performed by the experimenter. The result is an upper bound on the probability of success as a function of the number, momenta and masses of the particles involved in the experiment. We concluded that for high energy particles—say cosmic rayes—or for a considerable number of particles—say in a particle beam—the discrimination can in-principle be achieved. We have also presented a discrimination experiment based on particles fly-time whose results are far from the optimal bound. A richer phenomenology will be available introducing the interaction in the QCA dynamics. A QCA description of Electrodynamics and the connection to gauge theories on the automaton is the next step of this research program.

The analysis of Chapter 6 closed the part of the thesis related to the Dirac QCA dynamics in a fixed reference frame. The Dirac QCA was shown to be not

invariant under Lorentz transformations and the usual covariance is manifestly broken in the automaton framework. In Chapter 7 we have characterized the invariance transformations of the automaton in the momentum-energy sector starting from its distorted dispersion relation. We got a model of Deformed Special Relativity (DSR) with an invariant energy scale whose novelty is to correspond to a precise microscopical model of field dynamics. The spacetime emerging from the QCA deformed momentum space is not absolutely local, namely any local observer constructs a different spacetime and the only invariant object is the full phase space of the theory. We have proved the loss of spacetime coincidences changing reference frame in the QCA framework. This actually happens for any non-linear deformation of the momentum space, with absolute locality recovered only in the linear Lorentz case. At the end of Chapter 7 we presented a result which is not directly connected with the automata theory. We showed how starting from an operational Einstein synchronization protocol on a classical causal network it is possible to derive a digital version of the Lorentz transformations. This result is included here to provide an additional motivation in the study of the automaton spacetime. In the DSR scenario the position-time coordinates are obtained as the linear projection of the coordinates in the phase space; it would be interesting to have an operational "Einstein like" synchronization protocol, based on a suitable notion of clock, for the direct construction of the coordinate system. This is a very hard topic and deserve further investigation.

The thesis concludes in Chapter 8 where are derived the informational features of the Fermionic systems, a fundamental building block of QFT. It turned out that the Fermionic Quantum Theory (FQT), regarded as a theory of information processing, is substantially different from QT. On one side it exhibits a limitation, since it does not satisfy local tomography and local measurements are not sufficient to discriminate multipartite states. On the other hand, while in QT the entanglement is monogamous and can only be shared by two parties, in FQT there are no limitations on the sharing of multipartite entanglement and this provides the theory with a powerful additional resource. The new features arise as direct consequences of the parity superselection rule on the Fermionic systems which avoids superpositions between states having an even and an odd number of particles. The notion of superselection has also been generalized in the scenario of operational probabilistic theories proving a trade-off between the number of constraints imposed on a local tomographic theory and the degree of holism (how much non-local-tomographic is the theory) of the superselected theory. We observe that, while the Fermionic computation and the standard quantum computation have been shown to be equivalent from the complexity point of view, our findings about Fermionic entanglement suggest that the same may not hold for distributed Fermionic computation. Moreover the non monogamy of the entanglement could have striking physical consequences, *i.e.* in the physics of black holes information. Indeed it has been shown that black holes behave as mirrors for the quantum information thrown

into them, reemitting it as Hawking radiation. However this behavior seems to depend on the entanglement monogamy in QT and one could consider the same process in the Fermionic case. Monogamy plaies a crucial role in the so called "firewall paradox" for black holes and FQT should be investigated as a possible way out of the paradox.

This thesis has to be considered a preliminary analysis of the field automaton as a theoretical model, based on operational and informational principles, aiming at the description of physical dynamics. The presented results confirm that the automaton discrete evolution is compatible with the QFT free dynamics which is restored as a large-scale effective theory. On the other hand the QCA predictions depart from the usual QFT ones at small length distances (or at high energy scales) where the notion of relativistic spacetime breaks down and the discreetness of the automaton becomes relevant. The QCA proposal could be an alternative way to explore the physics beyond the Standard Model. Besides the aforementioned future perspectives, due to the key role of information in the present approach, it could be significant to understand if the information associated with a part of the emergent space obeys the holographic principle [215, 216] whose validity is supported by black hole physics [30, 217] and by theoretical predictions as the AdS/CFT correspondence [218]. The holographic principle, which is usually studied in relativistic contexts, has been recently considered [219, 220] in the completely different context of emergent spacetime as a possible origin of gravity and inertia, thus reversing the common logic that lead people from the laws of gravity to holography. However in this scenario the holographic principle is assumed and a satisfactory informational justification for it is still missing.

Acknowledgements

The completion of this thesis benefited from the expertise and the favorable atmosphere of the Quantum Information Theory Group of Pavia University, led by Prof. Mauro D'Ariano. Undertaking my PhD within the group has been an excellent opportunity for enriching my scientific and professional education starting developing the skill-set needed in research. First and foremost I thank my advisor Mauro for his great contributions of time and ideas during my PhD studies, the enthusiasm and joy he has for research was motivational for me. I am also grateful to Mauro for funding me during the last year. I want to express my deeply-felt thanks to all members of the group with whom I share both research and friendship: Alessandro Bisio and Paolo Perinotti, who contributed to my training, Chiara Macchiavello, Franco Manessi, Lorenzo Maccone, Matteo Rossi and Massimiliano Sacchi. I am also sincerely thankful to some past member of the group: Giulio Chiribella, Timoteo Colnaghi, Michele Dall'Arno, Stefano Facchini, Michal Sedlak and Marco Zaopo.

I would like to thank Giovanni Amelino-Camelia and João Magueijo the official reading committee of this thesis for their time, interest, and careful reading. The final version of the presentation benefited from their comments.

I acknowledge all Professors and Researchers who contributed to the Doctorate degree program in Physics at Pavia University for coordinating PhD courses and series of Physics Colloquia of distinguished international speakers.

Last but not the least, I gratefully acknowledge my family and Valentina for coming along me throughout my life.

APPENDIX A

The Sampling Theorem

The Sampling Theorem [221] shows that a continuous band-limited function may be represented exactly by samples of its points taken at uniform intervals of ℓ if ℓ is sufficiently small. In other words, since the continuous function may be reconstructed perfectly from its samples, sampling at a high enough rate is information-lossless. Before stating the main theorem we give the definition of band-limited function:

Definition A.1 (Band-limited function) *Consider a function $f \in L_2(\mathbb{R})$. We say f to be band-limited to Δ if*

$$\hat{f}(k) = 0, \quad \textit{for} \quad |k| \geqslant \Delta. \tag{A.1}$$

The theorem is then as follows:

Theorem A.1 (Nyquist-Shannon Sampling Theorem) *Let $f \in L_2(\mathbb{R})$ be band-limited to Δ. Then f at all points x can be recovered from its samples $\{f(n\,\ell), n \in \mathbb{Z}\}$ at integer multiples of $\ell = 1/(2\Delta)$ as follows:*

$$f(x) = \sum_{n \in \mathbb{Z}} f(n\,\ell) \mathrm{sinc}\left(\frac{x - n\ell}{\ell}\right). \tag{A.2}$$

The theorem just states that a band-limited function can always be recovered as a weighted sum of its values on a discrete set of points where the weights are given by the sinc function. Before proving the theorem remember that the

sinc function and its Fourier transform are

$$\operatorname{sinc}(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x}, & x \neq 0, \\ 1 & x = 0 \end{cases}, \qquad \widehat{\operatorname{sinc}}(k) = \begin{cases} 1, & |k| \leqslant 1, \\ 0 & |k| > 1 \end{cases}, \tag{A.3}$$

and in particular an easy change of variable shows that the Fourier transform of $\operatorname{sinc}\left(\frac{x-n\ell}{\ell}\right)$ is $\ell \exp(-2\pi i k n \ell)$ if $|k| \leqslant \Delta$ while it is 0 for $|k| \geqslant \Delta$, according to the band-limitness of f in (A.2).

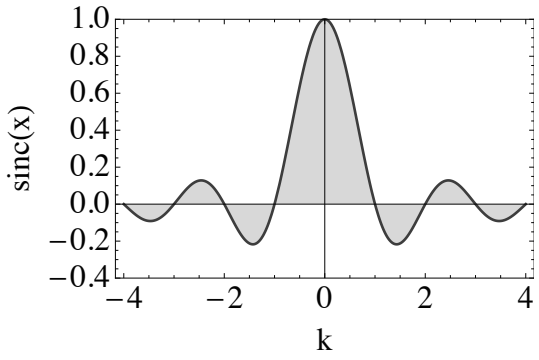


Figure A.1: Plot of the sinc function

The usual proof of the Sampling Theorem presented here is just an easy application of the Fourier series expansion of a limited function.

Proof. By hypothesis f is band-limited to Δ. Therefore, by definition of band-limited function, its Fourier transform $\hat{f}$ is limited to the interval $k \in [-\Delta, \Delta]$ and then can be expanded in Fourier series as follows

$$\hat{f}(k) = \sum_{n=-\infty}^{\infty} \hat{f}_n e^{-2\pi i n k \ell}, \quad k \in [-\Delta, \Delta], \tag{A.4}$$

$$\hat{f}_n = \frac{1}{2\Delta} \int_{-\Delta}^{\Delta} \mathrm{d}k \; \hat{f}(k) e^{2\pi i n k \ell}. \tag{A.5}$$

Now we can Fourier transform both sides of (A.4) and using the Fourier transform of the sinc function given in (A.3) we get

$$f(x) = \sum_{n \in \mathbb{Z}} \frac{\widehat{f_n}}{\ell} \operatorname{sinc}\left(\frac{x - n\,\ell}{\ell}\right). \tag{A.6}$$

Finally evaluating both sides at $x = n\ell$ we have $f(n\ell) = \widehat{f_n}/\ell$, which completes the proof. ■

We can now observe that in general it is possible to approximate the function f taking a sampling at a rate $\ell' > 1/2\Delta$.

APPENDIX B

The stationary phase approximation

The stationary phase approximation allows the evaluation of the integral of a function against a rapidly oscillating exponential

$$\mathcal{I}(t) = \int_a^b \mathrm{d}k \; f(k) e^{i\omega(k)t}. \tag{B.1}$$

In this scenario the integral becomes small as the frequency increases.

Under very weak assumptions the Riemann-Lebesgue lemma allows to state that asymptotically the integral goes to zero:

Lemma B.1 (Riemann-Lebesgue) *Consider a function $f \in L_1[a,b]$. Then we have*

$$\lim_{t\to\pm\infty} \int_a^b \mathrm{d}k \; f(k) e^{ikt} = 0. \tag{B.2}$$

A similar result holds for $f \in L_1(\mathbb{R})$.

Imposing some regularity conditions on the functions f and ω, it is possible to derive the asymptotical behaviour for the oscillating integral (B.1) via the so called *stationary phase approximation*. Thus let us assume f and ω smooth enough to admit Taylor expansion, *i.e.* $f, \omega \in C_0^\infty[a,b]$. The idea at the basis of the stationary phase approximation is that the contribution to the integral $\mathcal{I}$ in the regions where there is a rapid oscillation of the exponential term is

approximately zero, with the only significant non-zero contributions occurring in the regions of the integration range where

$$\omega^{(1)}(k) = 0, \tag{B.3}$$

i.e. at *points of stationary phase* ($\omega^{(i)}$ denotes the nth derivative).

Suppose that at some point $k_0 \in [a, b]$ it is $\omega^{(1)}(k_0) = 0$, with $\omega^{(1)}(k) \neq 0$ everywhere else. Assume moreover that $\omega^{(2)}(k_0) \neq 0$ and $f(k_0) \neq 0$. Since for $t \gg 0$ the only contribution to the integral is in a small neighbor of the stationary point k_0 we have

$$\mathcal{I}(t) \approx e^{i\omega(k_0)t} \int_{k_0-\epsilon}^{k_0+\epsilon} \mathrm{d}k \; f(k) e^{i(\omega(k)-\omega(k_0))t} \tag{B.4}$$

$$\approx f(k_0) e^{i\omega(k_0)t} \int_{k_0-\epsilon}^{k_0+\epsilon} \mathrm{d}k \; e^{i(\omega(k)-\omega(k_0))t}, \tag{B.5}$$

where we have pulled f outside the integral since $f(k) \approx f(k_0)$ in the vicinity of the stationary point and f is assumed slowly varying. Now expanding $\omega(k)$ near k_0 one has

$$\mathcal{I}(t) \approx f(k_0) e^{i\omega(k_0)t} \int_{k_0-\epsilon}^{k_0+\epsilon} \mathrm{d}k \; e^{\frac{i}{2}\omega^{(2)}(k_0)(k-k_0)^2 t} \tag{B.6}$$

$$\approx f(k_0) e^{i\omega(k_0)t} \int_{-\infty}^{\infty} \mathrm{d}k \; e^{\frac{i}{2}\omega^{(2)}(k_0)(k-k_0)^2 t}, \tag{B.7}$$

and computing exactly the integral in (B.7) we finally get the leading order asymptotic approximation

$$\mathcal{I}(t) \xrightarrow{t \gg 0} f(k_0) e^{i\omega(k_0)t} \sqrt{\frac{2\pi i}{\omega^{(2)}(k_0) t}}. \tag{B.8}$$

Bibliography

[1] Bridgman, P. W., Bridgman, P. W., Bridgman, P. W., and Bridgman, P. W., *The logic of modern physics*, Macmillan New York, 1927.

[2] Ludwig, G. and Hein, C. A., *Foundations of quantum mechanics II*, New York, NY (United States); Springer-Verlag, 1985.

[3] Hardy, L., Arxiv preprint quant-ph/0101012 (2001).

[4] Fuchs, C. A., arXiv preprint quant-ph/0205039 (2002).

[5] Ariano, G. M. D., AIP Conference Proceedings **810** (2006) 114.

[6] D'Ariano, G. M., Philosophy of Quantum Information and Entanglement **85** (2010).

[7] Chiribella, G., D'Ariano, G. M., and Perinotti, P., Phys. Rev. A **81** (2010) 062348.

[8] Masanes, L. and Müller, M. P., New Journal of Physics **13** (2011) 063001.

[9] Dakic, B. and Brukner, C., Quantum theory and beyond: is entanglement special?, in *Deep Beauty: Understanding the Quantum World through Mathematical Innovation*, edited by Halvorson, H., pages 365–392, Cambridge University Press, 2011.

[10] Chiribella, G., D'Ariano, G., and Perinotti, P., Phys. Rev. A **84** (2011) 012311.

[11] Kuhlmann, M., Quantum field theory, in *The Stanford Encyclopedia of Philosophy*, edited by Zalta, E. N., http://plato.stanford.edu/archives/win2012/entries/quantum-field-theory, winter 2012 edition, 2012.

[12] Cao, T. and Schweber, S., Synthese **97** (1993) 33.

[13] Teller, P., *An interpretive introduction to quantum field theory*, Princeton Univ Pr, 1997.

[14] Auyang, S., *How is quantum field theory possible?*, Oxford University Press, USA, 1995.

[15] Cao, T., *Conceptual foundations of quantum field theory*, Cambridge Univ Pr, 2004.

[16] Thaller, B., *The Dirac Equation*, New York, NY (United States); Springer-Verlag, 1992.

[17] Reeh, H. and Schlieder, S., Il Nuovo Cimento (1955-1965) **22** (1961) 1051.

[18] Polchinski, J., arXiv preprint hep-th/9611050 (1996).

[19] Polchinski, J. G., *String theory*, Cambridge university press, 2003.

[20] Bombelli, L., Lee, J., Meyer, D., and Sorkin, R., Physical review letters **59** (1987) 521.

[21] Connes, A. and Lott, J., Nuclear Physics B-Proceedings Supplements **18** (1991) 29.

[22] Snyder, H., Physical Review **71** (1947) 38.

[23] Amelino-Camelia, G., International Journal of Modern Physics D **11** (2002) 35.

[24] Amelino-Camelia, G. and Piran, T., Physical Review D **64** (2001) 036005.

[25] Magueijo, J. and Smolin, L., Physical Review D **67** (2003) 044017.

[26] Hogan, C., Arxiv preprint arXiv:1002.4880 (2010).

[27] Hogan, C., Arxiv preprint arXiv:1204.5948 (2012).

[28] Pikovski, I., Vanner, M., Aspelmeyer, M., Kim, M., and Brukner, C., Nature Physics **331** (2012) 393.

[29] Bekenstein, J. D., Lettere Al Nuovo Cimento (1971–1985) **4** (1972) 737.

[30] Hawking, S. W., Communications in mathematical physics **43** (1975) 199.

[31] Haag, R. and Kastler, D., Journal of Mathematical Physics **5** (1964) 848.

[32] Streater, R. F. and Wightman, A. S., *PCT, spin and statistics, and all that*, Princeton University Press, 1964.

[33] Haag, R., *Local quantum physics*, volume 2, Springer Berlin, 1996.

[34] D'Ariano, G. M., AIP Conference Proceedings **1232** (2010) 3.

[35] D'Ariano, G. M., AIP Conference Proceedings **1327** (2011) 7.

[36] D'Ariano, G. M., Phys. Lett. A **376** (2011).

[37] D'Ariano, G. M., AIP Conference Proceedings **1508** (2012) 146.

[38] D'Ariano, G. M., Il Nuovo Saggiatore **28** (2012) 13.

[39] von Neumann, J., *Theory of self-reproducing automata*, University of Illinois Press, Urbana and London, 1966.

[40] Gardner, M., Scientific American **223** (1970) 120.

[41] Wolfram, S., *A new kind of science*, volume 5, Wolfram media Champaign, 2002.

[42] Feynman, R., International journal of theoretical physics **21** (1982) 467.

[43] Grossing, G. and Zeilinger, A., Complex Systems **2** (1988) 197.

[44] Schumacher, B. and Werner, R., Arxiv preprint quant-ph/0405174 (2004).

[45] Arrighi, P., Nesme, V., and Werner, R., Journal of Computer and System Sciences **77** (2011) 372.

[46] Gross, D., Nesme, V., Vogts, H., and Werner, R., Communications in Mathematical Physics (2012) 1.

[47] 't Hooft, G., Nuclear Physics B **342** (1990) 471.

[48] Nakamura, T., Journal of mathematical physics **32** (1991) 457.

[49] Bialynicki-Birula, I., Physical Review D **49** (1994) 6920.

[50] Meyer, D., Journal of Statistical Physics **85** (1996) 551.

[51] Yepez, J., Quantum Information Processing **4** (2006) 471.

[52] Kogut, J. B., Rev. Mod. Phys. **51** (1979) 659.

[53] Kogut, J. B., Rev. Mod. Phys. **55** (1983) 775.

[54] Nielsen, H. B. and Ninomiya, M., Physics Letters B **105** (1981) 219.

[55] D'Ariano, G. M. and Tosini, A., Quantum Information Processing **9** (2010) 95.

[56] Kari, J., Theoretical Computer Science **334** (2005) 3.

[57] Ambainis, A., Bach, E., Nayak, A., Vishwanath, A., and Watrous, J., One-dimensional quantum walks, in *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pages 37–49, ACM, 2001.

[58] Vogts, H., *Discrete Time Quantum Lattice Systems*, Braunschweig, Techn. Univ., 2009.

[59] Bisio, A., D'Ariano, G., and Tosini, A., arXiv preprint arXiv:1212.2839 (2012).

[60] D'Ariano, G. M. and Perinotti, P., arXiv preprint arXiv:1306.1934 (2013).

[61] Bisio, A., D'Ariano, G. M., and Tosini, A., Phys. Rev. A **88** (2013) 032301.

[62] Magueijo, J. and Smolin, L., Phys. Rev. Lett. **88** (2002) 190403.

[63] Amelino-Camelia, G., Freidel, L., Kowalski-Glikman, J., and Smolin, L., Phys. Rev. D **84** (2011) 084010.

[64] Amelino-Camelia, G., Freidel, L., Kowalski-Glikman, J., and Smolin, L., International Journal of Modern Physics D **20** (2011) 2867.

[65] Amelino-Camelia, G., Astuti, V., and Rosati, G., The European Physical Journal C **73** (2013) 1.

[66] Bibeau-Delisle, A., Bisio, A., D'Ariano, G. M., Perinotti, P., and Tosini, A., unpublished (2013).

[67] D'Ariano, G. M. and Tosini, A., Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics **44** (2013) 294 .

[68] D'Ariano, G. M., Manessi, F., Perinotti, P., and Tosini, A., arXiv preprint arXiv:1307.7902 (2013).

[69] Spekkens, R. W., Phys. Rev. A **75** (2007) 032110.

[70] Mac Lane, S., *Categories for the working mathematician*, volume 5, Springer verlag, 1998.

[71] Hardy, L. and Wootters, W. K., Foundations of Physics **42** (2012) 454.

[72] Barrett, J., Phys. Rev. A **75** (2007) 032304.

[73] Chiribella, G., D'Ariano, G. M., and Perinotti, P., EPL (Europhysics Letters) **83** (2008) 30004.

[74] Chiribella, G., D'Ariano, G. M., and Perinotti, P., Phys. Rev. A **80** (2009) 022339.

[75] D'Ariano, G. M., Manessi, F., and Perinotti, P., arXiv preprint arXiv:1301.7578 (2013).

[76] Peres, A., Phys. Rev. Lett. **77** (1996) 1413.

[77] Horodecki, M., Horodecki, P., and Horodecki, R., Physics Letters A **223** (1996) 1.

[78] Wootters, W. K., Phys. Rev. Lett. **80** (1998) 2245.

[79] Wootters, W. K., Quantum Information & Computation **1** (2001) 27.

[80] DiVincenzo, D. P., Phys. Rev. A **51** (1995) 1015.

[81] Vedral, V., Plenio, M. B., Rippin, M., and Knight, P. L., arXiv preprint quant-ph/9702027 (1997).

[82] Plenio, M. B. and Virmani, S., arXiv preprint quant-ph/0504163 (2005).

[83] Bennett, C. H., DiVincenzo, D. P., Smolin, J. A., and Wootters, W. K., Physical Review A **54** (1996) 3824.

[84] Hill, S. and Wootters, W. K., Physical review letters **78** (1997) 5022.

[85] Bennett, C. H. et al., Phys. Rev. A **59** (1999) 1070.

[86] Bennett, C. H., Bernstein, H. J., Popescu, S., and Schumacher, B., Phys. Rev. A **53** (1996) 2046.

[87] Sakurai, J. J., Reading, MA: Addison Wesley,| c1985, edited by Tuan, San Fu **1** (1985).

[88] Rungta, P., Bužek, V., Caves, C. M., Hillery, M., and Milburn, G. J., Phys. Rev. A **64** (2001) 042315.

[89] Popescu, S. and Rohrlich, D., Phys. Rev. A **56** (1997) R3319.

[90] de Vicente, J. I., Spee, C., and Kraus, B., arXiv preprint arXiv:1305.7398 (2013).

[91] Bruß, D., Phys. Rev. A **60** (1999) 4344.

[92] Dür, W., Vidal, G., and Cirac, J. I., Phys. Rev. A **62** (2000) 062314.

[93] Koashi, M., Bužek, V., and Imoto, N., Phys. Rev. A **62** (2000) 050302.

[94] Dennison, K. A. and Wootters, W. K., Phys. Rev. A **65** (2001) 010301.

[95] Coffman, V., Kundu, J., and Wootters, W. K., Phys. Rev. A **61** (2000) 052306.

[96] Raggio, G. and Werner, R., Helvetica Physica Acta **62** (1989) 980.

[97] Fannes, M., Lewis, J., and Verbeure, A., Letters in mathematical physics **15** (1988) 255.

[98] Terhal, B. M., IBM Journal of Research and Development **48** (2004) 71.

[99] Koashi, M. and Winter, A., Phys. Rev. A **69** (2004) 022309.

[100] Horodecki, R., Horodecki, P., Horodecki, M., and Horodecki, K., Rev. Mod. Phys. **81** (2009) 865.

[101] Wolfram, S., Reviews of modern physics **55** (1983) 601.

[102] Aharonov, Y., Davidovich, L., and Zagury, N., Physical Review A **48** (1993) 1687.

[103] Knight, P., Roldán, E., and Sipe, J., journal of modern optics **51** (2004) 1761.

[104] Ahlbrecht, A., Vogts, H., Werner, A., and Werner, R., Journal of Mathematical Physics **52** (2011) 042201.

[105] Toffoli, T. and Margolus, N. H., Physica D: Nonlinear Phenomena **45** (1990) 229.

[106] Kari, J., Physica D: Nonlinear Phenomena **45** (1990) 379.

[107] Hedlund, G. A., Theory of Computing Systems **3** (1969) 320.

[108] Richardson, D., Journal of Computer and System Sciences **6** (1972) 373.

[109] Arrighi, P., Nesme, V., and Werner, R., One-dimensional quantum cellular automata over finite, unbounded configurations, in *Language and Automata Theory and Applications*, pages 64–75, Springer, 2008.

[110] Toffoli, T., Journal of Computer and System Sciences **15** (1977) 213.

[111] Morita, K. and Harao, M., IEICE TRANSACTIONS (1976-1990) **72** (1989) 758.

[112] Amoroso, S. and Patt, Y. N., Journal of Computer and System Sciences **6** (1972) 448.

[113] Kari, J., Journal of Computer and System Sciences **48** (1994) 149.

[114] Margolus, N., Physica D: Nonlinear Phenomena **10** (1984) 81.

[115] Toffoli, T. and Margolus, N., *Cellular automata machines*, MIT press, 1987.

[116] Kari, J., Mathematical Systems Theory **29** (1996) 47.

[117] Kari, J., Fundamenta Informaticae **38** (1999) 93.

[118] Nayak, A. and Vishwanath, A., arXiv preprint quant-ph/0010117 (2000).

[119] Aharonov, D., Ambainis, A., Kempe, J., and Vazirani, U., Quantum walks on graphs, in *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pages 50–59, ACM, 2001.

[120] Kempe, J., Contemporary Physics **44** (2003) 307.

[121] Carteret, H. A., Ismail, M. E., and Richmond, B., Journal of Physics A: Mathematical and General **36** (2003) 8775.

[122] Grimmett, G., Janson, S., and Scudo, P. F., Physical Review E **69** (2004) 026119.

[123] Konno, N., Namiki, T., and Soshi, T., Interdisciplinary Information Sciences **10** (2004) 11.

[124] Konno, N., Journal of the Mathematical Society of Japan **57** (2005) 1179.

[125] Meyer, D. A., International Journal of Modern Physics C **8** (1997) 717.

[126] Kitaev, A., Annals of Physics **321** (2006) 2.

[127] Susskind, L., Phys. Rev. D **16** (1977) 3031.

[128] Jordan, P. and Wigner, E., Zeitschrift für Physik **47** (1928) 631.

[129] Farrelly, T. C. and Short, A. J., arXiv preprint arXiv:1303.4652 (2013).

[130] Valcárcel, G., Roldán, E., and Romanelli, A., New Journal of Physics **12** (2010) 123022.

[131] Reitzner, D., Nagaj, D., and Bužek, V., Acta Physica Slovaca. Reviews and Tutorials **61** (2011) 603.

[132] Takeda, M. et al., Physical Review Letters **81** (1998) 1163.

[133] Bridson, M., Gowers, T., Barrow-Green, J., and Leader, I., The Princeton companion to mathematics, T. Gowers, J. Barrow-Green, and I. Leader, Eds 10.

[134] Huang, K., American Journal of Physics **20** (1952) 479.

[135] Cannata, F. and Ferrari, L., Physical Review B **44** (1991) 8599.

[136] Ferrari, L. and Russo, G., Physical Review B **42** (1990) 7454.

[137] Cannata, F., Ferrari, L., and Russo, G., Solid State Communications **74** (1990) 309.

[138] Lurié, D. and Cremer, S., Physica **50** (1970) 224.

[139] Shen, S.-Q., Phys. Rev. Lett. **95** (2005) 187203.

[140] Bernardes, E., Schliemann, J., Lee, M., Egues, J. C., and Loss, D., Phys. Rev. Lett. **99** (2007) 076603.

[141] Lamata, L., León, J., Schätz, T., and Solano, E., Physical review letters **98** (2007) 253005.

[142] Gerritsma, R. et al., Nature **463** (2010) 68.

[143] Cserti, J. and Dávid, G., Physical Review B **74** (2006) 172305.

[144] Rusin, T. M. and Zawadzki, W., Physical Review B **76** (2007) 195439.

[145] Schliemann, J., Loss, D., and Westervelt, R. M., Physical review letters **94** (2005) 206801.

[146] Zawadzki, W., Physical Review B **72** (2005) 085217.

[147] Zawadzki, W. and Rusin, T., Physics Letters A **374** (2010) 3533.

[148] Geim, A. and Novoselov, K., Nature materials **6** (2007) 183.

[149] Zawadzki, W. and Rusin, T., Journal of Physics: Condensed Matter **23** (2011) 143201.

[150] Vaishnav, J. Y. and Clark, C. W., Physical review letters **100** (2008) 153002.

[151] Zhang, X., Phys. Rev. Lett. **100** (2008) 113903.

[152] Katsnelson, M., Novoselov, K., and Geim, A., Nature Physics **2** (2006) 620.

[153] Katsnelson, M. I. and Novoselov, K., Solid State Communications **143** (2007) 3.

[154] Sansoni, L. et al., Physical review letters **108** (2012) 010502.

[155] Schrödinger, E., *Über die kräftefreie Bewegung in der relativistischen Quantenmechanik*, Akademie der wissenschaften in kommission bei W. de Gruyter u. Company, 1930.

[156] Lock, J. A., Am. J. Phys **47** (1979) I979.

[157] Romera, E. and de los Santos, F., Phys. Rev. B **80** (2009) 165416.

[158] Romera, E., Phys. Rev. A **84** (2011) 052102.

[159] Kurzyński, P., Physics Letters A **372** (2008) 6125.

[160] Helstrom, C., *Quantum detection and estimation theory*, volume 84, Academic press New York, 1976.

[161] D'Ariano, G., Lo Presti, P., and Paris, M., Physical review letters **87** (2001) 270404.

[162] Moyer, M., Scientific American (2012).

[163] Hogan, C., Physical Review D **85** (2012).

[164] Amelino-Camelia, G., Laemmerzahl, C., Mercati, F., and Tino, G., Physical review letters **103** (2009) 171302.

[165] Rovelli, C. and Smolin, L., Nuclear Physics B **8** (1990) 80.

[166] Ashtekar, A., Rovelli, C., and Smolin, L., Physical Review Letters **69** (1992) 237.

[167] Rovelli, C. and Smolin, L., Nuclear Physics B **442** (1995) 593.

[168] Carlip, S., Reports on progress in physics **64** (2001) 885.

[169] Garay, L. J., International Journal of Modern Physics A **10** (1995) 145.

[170] Amelino-Camelia, G., Modern Physics Letters A **17** (2002) 899.

[171] Amelino-Camelia, G., Physics Letters B **510** (2001) 255.

[172] Albrecht, A. and Magueijo, J., Phys. Rev. D **59** (1999) 043516.

[173] Magueijo, J., Phys. Rev. D **62** (2000) 103521.

[174] Magueijo, J., Reports on Progress in Physics **66** (2003) 2025.

[175] Finkbeiner, D. P., Davis, M., and Schlegel, D. J., The Astrophysical Journal **544** (2000) 81.

[176] D'Ariano, G. M. and Tosini, A., arXiv preprint arXiv:1008.4805 (2010).

[177] Majid, S., *Foundations of quantum group theory*, Cambridge university press, 2000.

[178] Bruno, N., Amelino-Camelia, G., and Kowalski-Glikman, J., Physics Letters B **522** (2001) 133.

[179] Kowalski-Glikman, J., Physics Letters A **286** (2001) 391.

[180] Lukierski, J., Ruegg, H., Nowicki, A., and Tolstoy, V. N., Physics Letters B **264** (1991) 331.

[181] Majid, S. and Ruegg, H., Physics Letters B **334** (1994) 348.

[182] Kowalski-Glikman, J. and Nowak, S., International Journal of Modern Physics D **12** (2003) 299.

[183] Nowicki, A., arXiv preprint math/9803064 (1998).

[184] Kowalski-Glikman, J. and Nowak, S., Physics Letters B **539** (2002) 126.

[185] Arzano, M. and Marciano, A., Physical Review D **75** (2007) 081701.

[186] Arzano, M. and Marciano, A., Physical Review D **76** (2007) 125005.

[187] Amelino-Camelia, G., D'Andrea, F., and Mandanici, G., Journal of Cosmology and Astroparticle Physics **2003** (2003) 006.

[188] Daszkiewicz, M., Imilkowska, K., and Kowalski-Glikman, J., Physics Letters A **323** (2004) 345.

[189] Schützhold, R. and Unruh, W., Journal of Experimental and Theoretical Physics Letters **78** (2003) 431.

[190] Amelino-Camelia, G. and Majid, S., International Journal of Modern Physics A **15** (2000) 4301.

[191] Maggiore, M., Physics Letters B **304** (1993) 65.

[192] Salecker, H. and Wigner, E. P., Physical Review **109** (1958) 571.

[193] Margolus, N. and Levitin, L. B., Physica D: Nonlinear Phenomena **120** (1998) 188.

[194] Giovannetti, V., Lloyd, S., and Maccone, L., Science **306** (2004) 1330.

[195] Lamport, L., Communications of the ACM **21** (1978) 558.

[196] Onsager, L., Physical Review **65** (1944) 117.

[197] Lieb, E., Schultz, T., and Mattis, D., Rev. Mod. Phys **36** (1964) 856.

[198] Bañuls, M.-C., Cirac, J. I., and Wolf, M. M., Phys. Rev. A **76** (2007) 022311.

[199] Wolf, M. M., Phys. Rev. Lett. **96** (2006) 010404.

[200] Bravyi, S. B. and Kitaev, A. Y., Annals of Physics **298** (2002) 210 .

[201] Montero, M. and Martín-Martínez, E., Phys. Rev. A **83** (2011) 062323.

[202] Brádler, K. and Jáuregui, R., Phys. Rev. A **85** (2012) 016301.

[203] Montero, M. and Martín-Martínez, E., Phys. Rev. A **85** (2012) 016302.

[204] Friis, N., Lee, A. R., and Bruschi, D. E., Phys. Rev. A **87** (2013) 022338.

[205] Verstraete, F. and Cirac, J., Journal of Statistical Mechanics: Theory and Experiment **2005** (2005) P09012.

[206] Weinberg, S., *The quantum theory of fields: Foundations.*, volume 1, Cambridge university press, 2005.

[207] D'Ariano, G. M., Probabilistic theories: What is special about quantum mechanics?, in *Philosophy of Quantum Information and Entanglement*, edited by Bokulich, A. and Jaeger, G., chapter 5, Cambridge University Press., 2010.

[208] Caves, C., Fuchs, C., and Rungta, P., Foundations of Physics Letters **14** (2001) 199.

[209] Wootters, W., Foundations of Physics **42** (2012) 19.

[210] Hayden, P. and Preskill, J., Journal of high energy physics **2007** (2007) 120.

[211] Braunstein, S. L., Pirandola, S., and Życzkowski, K., Phys. Rev. Lett. **110** (2013) 101301.

[212] Müller, M. P., Oppenheim, J., and Dahlsten, O. C., Journal of High Energy Physics **2012** (2012) 1.

[213] Abeyesinghe, A., Devetak, I., Hayden, P., and Winter, A., Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science **465** (2009) 2537.

[214] Verlinde, E. and Verlinde, H., arXiv preprint arXiv:1306.0515 (2013).

[215] 't Hooft, G., arXiv preprint gr-qc/9310026 (1993).

[216] Susskind, L., arXiv preprint hep-th/9409089 (1994).

[217] Bekenstein, J. D., Physical Review D **7** (1973) 2333.

[218] Maldacena, J., International journal of theoretical physics **38** (1999) 1113.

[219] Jacobson, T., Physical Review Letters **75** (1995) 1260.

[220] Verlinde, E., Journal of High Energy Physics **2011** (2011) 1.

[221] Shannon, C. E., Proceedings of the IRE **37** (1949) 10.

List of publications

- D'Ariano, G. M. and Tosini, A., *Testing axioms for quantum theory on probabilistic toy-theories*, Quantum Information Processing **9** (2010) 95.
- D'Ariano, G. M., and Tosini, A., *Emergence of Space-Time from topologically homogeneous causal networks*, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics **44** (2013) 294.
- Bisio, A., D'Ariano, G., and Tosini, A., *Quantum Field as a Quantum Cellular Automaton: the Dirac free evolution in one dimension* , arXiv preprint arXiv:1212.2839 (2012).
- Bisio, A., D'Ariano, G. M., and Tosini, A., *The Dirac Quantum Cellular Automaton in one dimension: Zitterbewegung and scattering from potential*, Phys. Rev. A 88 (2013).
- D'Ariano, G. M., Manessi, F., Perinotti, P., and Tosini, A., *Fermionic computation is non-local tomographic and violates monogamy of entanglement*, arXiv preprint arXiv:1307.7902 (2013).
- Bibeau-Delisle, A., Bisio, A., D'Ariano, G. M., Perinotti, P., and Tosini, *Doubly-Special Relativity from Quantum Cellular Automata*, A., arXiv preprint arXiv:1310.6760 (2013).
- D'Ariano, G. M. and Tosini, A., *Space-time and special relativity from causal networks*, arXiv preprint arXiv:1008.4805 (2010).