IOP Publishing | Royal Swedish Academy of Sciences Physica Scripta

Phys. Scr. T163 (2014) 014013 (9pp) doi:10.1088/0031-8949/2014/T163/014013

Determinism without causality

G M D'Ariano[1,2], F Manessi[2] and P Perinotti[1,2]

[1] *QUIT group*, Dipartimento di Fisica, via Bassi 6, I-27100 Pavia, Italy
[2] *INFN Gruppo IV*, Sezione di Pavia, via Bassi, 6, I-27100 Pavia, Italy

E-mail: dariano@unipv.it, franco.manessi01@ateneopv.it and perinotti@unipv.it

Received 14 April 2014, revised 20 June 2014
Accepted for publication 21 July 2014
Published 19 December 2014

Abstract

Causality has often been confused with the notion of determinism. It is mandatory to separate the two notions in view of the debate about quantum foundations. Quantum theory provides an example of causal non-deterministic theory. Here we introduce a toy operational theory that is deterministic and non-causal, thus proving that the two notions of causality and determinism are totally independent.

Keywords: determinism, causality, operational theory, non-causal deterministic theory

(Some figures may appear in colour only in the online journal)

1. Introduction

Causality is the subject of a very extensive literature, encompassing hundreds of contemporary books and technical articles. It hits a wide spectrum of disciplines, ranging from pure philosophy to law, economics, natural sciences, and, in particular, physics. Perhaps the most natural connection with physics is in philosophy, from the early work of Aristotle, to the cornerstone of René Descartes, who broke the ground for the modern view of David Hume and Immanuel Kant, up to the contemporary works on physical causation of Wesley Salmon [1] and Phil Dowe [2].

The recent reconsideration of the foundations of physics, with particular focus on quantum theory, has brought research in theoretical physics to explore issues in the territory shared with philosophy. A paradigmatic case is the issue of realism raised by the founding fathers von Neumann [3] and Einstein [4] in regard of the completeness of quantum theory.

The problem of causality has remained in the realm of philosophy, and stayed only in the background of physics, without the status of a physical law or the rank of a principle. Most of the time causality creeps in the form of ad hoc assumptions based on empirical evidence—like the discard of advanced potentials in electrodynamics or the Kramers–Kronig relations—or it is part of the interpretation of the theory—e.g. in special relativity—or else it is hidden in the theoretical framework, as in Hardy axiomatization of quantum theory [5].

A notion that is traditionally connected with causality in physics and philosophy is *determinism*, which is deeply entangled with causality, to the extent that the two are often merged into *causal determinism*, or even confused, as in the exemplar quotation from Max Planck: *'An event is causally determined if it can be predicted with certainty'* [6]. This confusion between the two notions is the source of the typical misleading way of regarding quantum correlations as 'spooky action at a distance'—the commonplace of perfect Einstein-Podolsky-Rosen (EPR) correlations interpreted as causation.

The advent of quantum mechanics has led us to consider the alternative point of view of contemplating causal relations as intrinsically probabilistic [7]. Although a probabilistic context for causal relations had already been considered by several authors [1, 2, 8], a precise mathematical formulation has been given only recently in the context of the operational axiomatization of quantum theory [9]. The main ingredient is the notion of *test*, which includes not only a complete set of alternative events, but also the specification of input and output *systems*, which encompass the possibility of concomitant causes and multiple effects, allowing for a network description as in [8]. The famous criticism of causality by Bertrand Russell [10] was indeed addressed against its deterministic notion, along with the overlooking of concomitant causations. All such criticisms are completely overcome by the operational-probabilistic formulation.

The property of causality within classical theory is trivialized by the irrelevance of the notion of measurement, which is identified with that of state itself. Complementarity is the feature that breaks the classical identification between observation and preparation (measurement and state). In the operational-probabilistic framework measurements and states

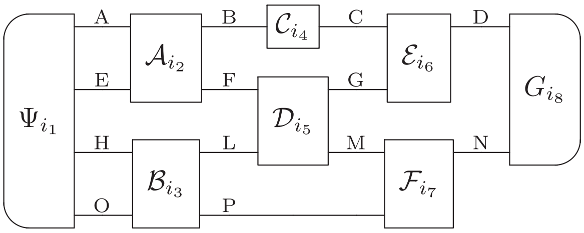


Figure 1. The closed circuit in the figure represents the joint probability $\Pr[i_1, i_2, \dots i_8 \mid \Psi, \mathcal{A}, \dots, \mathbf{G}]$ of outcomes $i_1, i_2, \dots i_8$ conditioned by the choice of tests $\Psi, \mathcal{A}, \dots, \mathbf{G}$. Since the output of the event $\mathcal{A}_{i_2}$ is connected to the input of the event $\mathcal{D}_{i_5}$ through the system F, the event $\mathcal{A}_{i_2}$ immediately precedes the event $\mathcal{D}_{i_5}$ ($\mathcal{A}_{i_2} \prec_1 \mathcal{D}_{i_5}$). Similarly, since between the event $\mathcal{B}_{i_3}$ and the event $\mathcal{E}_{i_6}$ there is $\mathcal{D}_{i_5}$ such that $\mathcal{B}_{i_3} \prec_1 \mathcal{D}_{i_5} \prec_1 \mathcal{E}_{i_6}$, the event $\mathcal{B}_{i_3}$ precedes the event $\mathcal{E}_{i_6}$ ($\mathcal{B}_{i_3} \prec \mathcal{E}_{i_6}$). If the closed circuit of the figure belongs to a causal theory, we have e.g. that the marginal probability of the event $\mathcal{D}_{i_5} \in \mathcal{D}$ cannot depend on the choice of any test $\mathcal{X}$ such that $\mathcal{X} \not\prec \mathcal{D}$, i.e. $\Pr[i_5 \mid \Psi, \mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathbf{G}] = \Pr[i_5 \mid \Psi, \mathcal{A}, \mathcal{B}]$.

correspond to special tests: the observation and the preparation tests. Causality is defined as the independence of the probability of the preparation test from the choice of the observation test: this definition of causality distills all the intuitive guises in which it appears in physics, with an intimate relation with the Einsteinian notion. In this formulation it is the first axiom of quantum theory in the derivation of [11]. A preliminary version of the present causality postulate has been given in [12] rephrased as *no-signaling from the future*, of which the first embryo can be found in a footnote in [13]. One can easily realize that the present formulation of causality is the only possible one, and covers even the original intuitive notion by Hume. The new mathematical definition of causality exhibits its full conceptual power in providing a non-trivial technical characterization of causal theories in terms of the uniqueness of the deterministic effect [9], as we will also review in this paper.

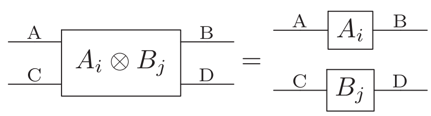
In the same framework we can also naturally introduce a precise notion of determinism, which must be taken as separate from that of reversibility of the evolution. The notion of determinism arose within the clockwork-universe vision of classical mechanics, assessing that the state of a system at an initial time completely determines the state at any later time. Classical mechanics, however, identifies the *state* (the point in the phase-space) with the *measurement-outcome*, while the two notions are radically different in quantum theory, and more generally in *operational probabilistic theories* [9, 11]. These allow us to define determinism outside the framework of classical mechanics which is already deterministic, avoiding the confusion between state and measurement-outcome. In a probabilistic context [9] determinism is identified with the property of a theory of having all probabilities of physical events equal to either zero or one—a definition which has no causal connotation.

Quantum theory provides a relevant example of operational probabilistic theory that is causal and not deterministic. In this paper we introduce a toy theory that is deterministic and non-causal. The purpose is to prove in this way that neither does causality imply determinism, nor determinism imply causality, namely that the the two notions are logically independent. In the concluding section we will further discuss the relation between the definition of causality in section 2 and the customary problem of physical causation along with the cause–effect connection.

2. Review on operational probabilistic theories

Before starting we need to review the basic definitions and notations for operational probabilistic theories (OPTs). For a detailed discussion see [9]. The basic notion in the operational framework is that of *test*. A test $\mathcal{A} = \{\mathcal{A}_i\}$ describes an elementary operation which generally produces the readout of an outcome i, heralding the occurrence of an *event* $\mathcal{A}_i$. Tests are also specified by an input and an output label, e.g. A, B, which identify the *system types* (*systems*, for short). The test $\mathcal{A}$ and its building events $\mathcal{A}_i \in \mathcal{A}$ can be represented by means of boxes as $\overset{\mathrm{A}}{\text{—}}\boxed{\mathcal{A}}\overset{\mathrm{B}}{\text{—}}$ and $\overset{\mathrm{A}}{\text{—}}\boxed{\mathcal{A}_i}\overset{\mathrm{B}}{\text{—}}$ respectively. The role of labelling input and output systems is to provide rules for connecting tests in sequences: an output wire labeled A can be connected only to an input wire with the same label A. Notice that the input/output relation has no causal connotation, and does not entail an underlying 'time arrow'. Here 'input/output' has to be understood as a functional dependence, namely the relation that links the variable x to the function evaluation $f(x)$. As will be clear shortly, only in a causal theory is it possible to understand the input/output relation as a time-arrow.

The event $\mathcal{B}_j \bigcirc \mathcal{A}_i$ belonging to the sequential composition $\mathcal{B} \bigcirc \mathcal{A}$ of the tests $\mathcal{A}$ and $\mathcal{B}$ is represented as $\overset{\mathrm{A}}{\text{—}}\boxed{\mathcal{A}_i}\overset{\mathrm{B}}{\text{—}}\boxed{\mathcal{B}_j}\overset{\mathrm{C}}{\text{—}}$ (a similar graphical representation holds also for the test $\mathcal{B} \bigcirc \mathcal{A}$ itself). For every system A there exists a unique singleton test $\{\mathcal{I}_\mathrm{A}\}$ such that $\mathcal{I}_\mathrm{B} \bigcirc \mathcal{A} = \mathcal{A} \bigcirc \mathcal{I}_\mathrm{A}$ for every event $\mathcal{A}$ with input A and output B. For every couple of systems (A, B) we can form the composite system C := AB, on which we can perform tests $\mathcal{A} \otimes \mathcal{B}$ with events $\mathcal{A}_i \otimes \mathcal{B}_j$ in *parallel composition* represented as follows



and satisfying the following condition:

$$(\mathcal{C}_h \otimes \mathcal{D}_k) \bigcirc (\mathcal{A}_i \otimes \mathcal{B}_j) = (\mathcal{C}_h \bigcirc \mathcal{A}_i) \otimes (\mathcal{D}_k \bigcirc \mathcal{B}_j).$$

Notice that here $\otimes$ is a formal symbol for parallel composition, and not the usual tensor product of linear spaces. There is a special system type I, the *trivial system*, such that AI = IA = A. The tests with input system I and output A are called *preparation-tests* of A, while the tests with input system A and output I are called *observation-tests* of A. Preparation-events of A are denoted by the symbols $|\rho)_\mathrm{A}$ or $(\rho\overset{\mathrm{A}}{\text{—}}$, and observation-events by $(c|_\mathrm{A}$ or $\overset{\mathrm{A}}{\text{—}}c)$. Note that the words *preparation-test* and *observation-test* have an intrinsic causal connotation (usually one observes something that has been prepared, and not vice versa), however, here the two words should be taken only as technical terms. The two terms recover their usual meaning in a causal theory—the commonly studied case—and our abuse of terminology is for the sake of limiting temporary technical words.

An arbitrary complex test obtained by parallel and sequential composition of box diagrams is called a *circuit*. A circuit is *closed* if its overall input and output systems are the trivial ones. Figure 1 is an example of closed circuit. Given a circuit we say that an event $\mathcal{H}$ *is immediately connected to the input of* $\mathcal{K}$, and write $\mathcal{H} \prec_1 \mathcal{K}$, if there is an output system of $\mathcal{H}$ that is connected with an input system of $\mathcal{K}$; e.g. referring to the circuit in figure 1 $\mathcal{A}_{i_2} \prec_1 \mathcal{D}_{i_5}$. We can moreover introduce the transitive closure $\prec$ of the relation $\prec_1$, and we say that $\mathcal{H}$ *is connected to the input of* $\mathcal{K}$ if $\mathcal{H} \prec \mathcal{K}$ (e.g. $\mathcal{B}_{i_3} \prec \mathcal{E}_{i_6}$). The two relations $\prec_1$ and $\prec$ can be trivially extended from events to tests.

A theory is *probabilistic* if every closed circuit represents a probability distribution; e.g. the closed circuit in figure 1 represents the probability $\Pr[i_1, i_2, \ldots i_8 \mid \Psi, \mathcal{A}, \ldots, \mathbf{G}]$ of outcomes $i_1, i_2, \ldots i_8$ conditioned by the choice of tests $\Psi, \mathcal{A}, \ldots, \mathbf{G}$[3]. In probabilistic theories we can quotient the set of preparation-events of A by the equivalence relation $|\rho)_{\mathrm{A}} \sim |\sigma)_{\mathrm{A}} \Leftrightarrow$ *the probability of preparing* $|\rho)_{\mathrm{A}}$ *and measuring* $(c|_{\mathrm{A}}$ *is the same as preparing* $|\sigma)_{\mathrm{A}}$ *and measuring* $(c|_{\mathrm{A}}$ *for every observation-event* $(c|_{\mathrm{A}}$ *of* A (and similarly for observation-events). The equivalence classes of preparation-events and observation-events of A will be denoted by the same symbols as their elements $|\rho)_{\mathrm{A}}$ and $(c|_{\mathrm{A}}$, respectively, and will be called *state* $|\rho)_{\mathrm{A}}$ and *effect* $(c|_{\mathrm{A}}$. For every system A, we will denote by $\mathrm{St(A)}$, $\mathrm{Eff(A)}$ the sets of states and effects, respectively. States and effects are real-valued functionals on each other, and then they can be naturally embedded in reciprocally dual real vector spaces, $\mathrm{St}_{\mathbb{R}}(\mathrm{A})$ and $\mathrm{Eff}_{\mathbb{R}}(\mathrm{A})$, whose dimension D_{A} is assumed here to be finite. The application of the effect $(c_i|_{\mathrm{A}}$ on the state $|\rho)_{\mathrm{A}}$ is written as $(c_i \mid \rho)_{\mathrm{A}}$ and corresponds to the closed circuit , denoting therefore the probability of the ith outcome of the observation-test $\mathbf{c} = \{(c_i \mid \mathrm{A}\}_{i \in \eta}$ performed on the state ρ of system A, i.e. $(c_i \mid \rho)_{\mathrm{A}} := \Pr[c_i \mid \boldsymbol{\rho}]$.

Any event with input system A and output system B induces a collection of linear mappings from $\mathrm{St}_{\mathbb{R}}(\mathrm{AC})$ to $\mathrm{St}_{\mathbb{R}}(\mathrm{BC})$, for varying system C. Such a collection is called *transformation* from A to B. The set of transformations from A to B will be denoted by $\mathrm{Transf(A, B)}$, and its linear span by $\mathrm{Transf}_{\mathbb{R}}(\mathrm{A, B})$. The symbols $\mathcal{A}$ and  denoting the event $\mathcal{A}$ will be also used to represent the corresponding transformation.

We now introduce a precise notion of determinism through the following definition [9].

[3] To be more precise the definition of probabilistic theory includes also the following formal rule for the composition of events of trivial systems $p_i \otimes p_j := p_i p_j =: p_i \bigcirc p_j$, stating the independence of closed circuits.

Definition 1 (ODT). An *operational deterministic theory* (ODT) is an OPT with all closed circuits having probabilities 0 or 1.

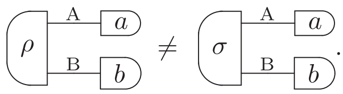
One cannot forbid the construction of the 'statistical' version of an ODT (as it happens for classical mechanics) by considering the OPT which is the convex closure of the ODT.

Given a set S the convex cone $\lambda\mathrm{S}$ is the conic hull of S, namely the set of all conic combinations of elements of S. With obvious notation we have the cones $\lambda\mathrm{St(A)}$, $\lambda\mathrm{Eff(A)}$, and $\lambda\mathrm{Transf(A, B)}$. The elements on the extremal rays of the cones are called *atomic*. In the following, we will use the Greek letters to denote states and Latin letters to denote effects. Moreover, in the rest of the paper we will not specify the system when it is clear from the context or it is generic.

An event $\mathcal{A}$ is *deterministic* if it belongs to a singleton test. We will denote respectively with $\mathrm{St}_1(\mathrm{A})$, $\mathrm{Eff}_1(\mathrm{A})$ and $\mathrm{Transf}_1(\mathrm{A, B})$ the set of deterministic states, effects and transformations for systems A and B, and we will often use the symbols $|\varepsilon)$ and $(e|$ to refer respectively to a deterministic state and effect. Note that in convex OPTs the sets $\mathrm{St}_1(\mathrm{A})$ and $\mathrm{Eff}_1(\mathrm{A})$ are convex. Deterministic transformations are also called *channels*.

Among the properties of OPTs, a relevant one is *local discriminability* [9], namely the possibility to discriminate multipartite states only through local measurement on the subsystems:

Definition 2 (Local discriminability). If $|\rho)_{\mathrm{AB}}, |\sigma)_{\mathrm{AB}} \in \mathrm{St}_1(\mathrm{AB})$ are states and $|\rho)_{\mathrm{AB}} \neq |\sigma)_{\mathrm{AB}}$, then there are two effects $(a|_{\mathrm{A}} \in \mathrm{Eff(A)}$ and $(b|_{\mathrm{B}} \in \mathrm{Eff(B)}$ such that



Local discriminability is equivalent to $\mathrm{St}_{\mathbb{R}}(\mathrm{AB}) = \mathrm{St}_{\mathbb{R}}(\mathrm{A}) \otimes \mathrm{St}_{\mathbb{R}}(\mathrm{B})$ [11], where now the symbol $\otimes$ denotes the usual tensor product of linear spaces. The analog condition also holds for the effects. An important consequence of local discriminability is that a transformation $\mathcal{T} \in \mathrm{Transf(A, B)}$ is completely specified by its action on $\mathrm{St(A)}$ [9]:

$$\mathcal{C} \mid \rho) = \mathcal{C}' \mid \rho) \quad \forall \mid \rho) \in \mathrm{St(A)} \;\Rightarrow\; \mathcal{C} = \mathcal{C}'.$$

We now introduce the definition of causality [11].

Definition 3 (Causal OPT). An OPT is *causal* if the probability for every preparation-test $\boldsymbol{\rho} = \{|\rho_i)\}_{i \in \eta}$ and any two observation-tests $\mathbf{a} = \{(a_j|\}_{j \in \chi}$ and $\mathbf{b} = \{(b_j|\}_{j \in \xi}$ one has $\sum_{j \in \chi}(a_j \mid \rho_i) = \sum_{k \in \xi}(b_k \mid \rho_i)$, $\forall\, i \in \eta$, namely the probability of the preparation is independent of the choice of observation.

Causality is equivalent to *no backward signaling* [12], namely within a closed circuit, the marginal probability of outcomes for a given test $\mathcal{H}$ do not depend on the choice of any test $\mathcal{K}$ not connected to the input of $\mathcal{H}$, i.e. $\mathcal{K} \nprec \mathcal{H}$. For example, in the circuit of figure 1 causality implies that

$$\Pr[i_5 \mid \Psi, \mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathbf{G}] = \Pr[i_5 \mid \Psi, \mathcal{A}, \mathcal{B}].$$

The present notion of causality is nothing but a rigorous definition of the so-called *Einstein causality*. Indeed, a corollary of *no backward signaling* is the *no-signaling without interaction* [9]. A crucial equivalent condition for causality of an OPT is the uniqueness of the deterministic effect [9].

The possibility of reversing the causal arrow (by defining *backward causality* or *retro-causality* as independence of observation on preparation) does not add anything new conceptually, since there is an isomorphism between any retro-causal theory and a causal one, upon exchanging the roles of input and output.

In the following we will take local discriminability for granted. We say that a linear map $\mathcal{T} \in \mathrm{Transf}_{\mathbb{R}}(\mathrm{A}, \mathrm{B})$ is *admissible* if it locally preserves the set of states $\mathrm{St}(\mathrm{AC})$, namely $\mathcal{T} \otimes \mathcal{I}_{\mathrm{C}}(\mathrm{St}(\mathrm{AC})) \subseteq \mathrm{St}(\mathrm{BC})$. In the following we will assume that every admissible map actually belongs to $\mathrm{Transf}(\mathrm{A}, \mathrm{B})$. We will refer to this last assumption as *no-restriction hypothesis*[4].

3. The deterministic non-causal theory

We now introduce an example of non-causal deterministic theory. The systems will be denoted by the symbols $\mathrm{n} \triangleright \mathrm{m}$, where n, m are positive integer numbers, and they enjoy the property that $\dim \mathrm{St}_{\mathbb{R}}(\mathrm{n} \triangleright \mathrm{m}) = \dim \mathrm{Eff}_{\mathbb{R}}(\mathrm{n} \triangleright \mathrm{m}) = n \cdot m$. Composition of systems is defined as $(\mathrm{n} \triangleright \mathrm{m})(\mathrm{n}' \triangleright \mathrm{m}') := \mathrm{x} \triangleright \mathrm{y}$, where $x = n \cdot n'$ and $y = m \cdot m'$, consistently with local discriminability. Notice that this definition is consistent with associativity and commutativity of parallel composition, as well as the existence of a trivial system $\mathrm{I} := (\mathrm{n} \triangleright \mathrm{m})$ with $n = m = 1$.

Denote by Γ_n the set of all the non-negative integer numbers less than n, i.e. $\Gamma_n := \{0, \ldots, n-1\}$. The set of states of the system $\mathrm{n} \triangleright \mathrm{m}$ is defined as $\mathrm{St}(\mathrm{n} \triangleright \mathrm{m}) := \{ |\alpha_{f\Xi}) | f\colon \Xi \to \Gamma_m \text{ and } \Xi \subseteq \Gamma_n \}$. The atomic states of $\mathrm{St}(\mathrm{n} \triangleright \mathrm{m})$ are the elements $|\alpha_{f,\{i\}})$ with $f\colon \{i\} \to \Gamma_m$, $i \in \Gamma_n$. In the following we will use a special notation for the atomic states: $|\alpha_{i\,j}) = |\alpha_{f\{i\}})$ with $f(i) = j$. The number of different atomic states for $\mathrm{n} \triangleright \mathrm{m}$ is $n \cdot m$, i.e. the same as the dimension of $\mathrm{St}_{\mathbb{R}}(\mathrm{n} \triangleright \mathrm{m})$. For $\Xi, \Upsilon \subset \Gamma_n$ with $\Xi \cap \Upsilon = \varnothing$, the states of $\mathrm{n} \triangleright \mathrm{m}$ enjoy the property $|\alpha_{f,\Xi}) + |\alpha_{g,\Upsilon}) \equiv |\alpha_{h,\Xi \cup \Upsilon})$, with $h\colon \Xi \cup \Upsilon \to \Gamma_m$, $h(i) := f(i)$ for $i \in \Xi$, and $h(i) := g(i)$ for $i \in \Upsilon$. Notice that for $\Xi \cap \Upsilon \neq \varnothing$, $|\alpha_{f\Xi}) + |\alpha_{g\Upsilon})$ is not a valid state. We have that a deterministic state is an element $|\varepsilon_f) := |\alpha_{f,\Gamma_n})$, hence the set of the deterministic states is $\mathrm{St}_1(\mathrm{n} \triangleright \mathrm{m}) = \{ |\varepsilon_f), \ f\colon \Gamma_n \to \Gamma_m \}$.

The set of states $\mathrm{St}(\mathrm{x} \triangleright \mathrm{y})$ for the bipartite system $\mathrm{x} \triangleright \mathrm{y} = (\mathrm{n} \triangleright \mathrm{m})(\mathrm{n}' \triangleright \mathrm{m}')$ is built up via the definition of bipartite atomic states as parallel composition of single-system atomic states $|\alpha_{(s,s')\,(t,t')}) := |\alpha_{s\,t}) \otimes |\alpha_{s'\,t'})$, with $\Gamma_x := \Gamma_n \times \Gamma_{n'}$ and $\Gamma_y := \Gamma_m \times \Gamma_{m'}$. It can be shown that this is the only possible definition of atomic state consistent with local discriminability (see propositions 1 and 2 in the appendix).

Under the no-restriction hypothesis we can easily build the set of effects for the system $\mathrm{n} \triangleright \mathrm{m}$ from the set $\mathrm{St}(\mathrm{n} \triangleright \mathrm{m})$. The atomic effects are the elements $(a_{s\,s'}|$ such that $(a_{s\,s'} | \alpha_{t\,t'}) = \delta_{st}\delta_{s't'}$ (see proposition 4 in the appendix). In general, it can be shown that $\mathrm{Eff}(\mathrm{n} \triangleright \mathrm{m}) := \{ (a_{\,v,E}| | v \in \Gamma_n \text{ and } E \subseteq \Gamma_m \}$, using the definition $(a_{\,v,E\cup F}| := (a_{\,v,E}| + (a_{\,v,F}|$ for $E \cap F = \varnothing$ (see proposition 5 in the appendix). The atomic effects are $(a_{\,s,\{s'\}}| \equiv (a_{s\,s'}|$. The deterministic effects are the elements $(e_v| := (a_{\,v,\Gamma_m}|$, and one can verify that $(e_v | \varepsilon_f) = 1$ for every $|\varepsilon_f) \in \mathrm{St}_1(\mathrm{n} \triangleright \mathrm{m})$. Indeed, one can check that $(a_{\,vE} | \alpha_{f\Xi}) := \chi_{\Xi}(v)\chi_E(f(v))$, with χ_S the indicator function of the set S, showing that for $E = \Gamma_m$, $\Xi = \Gamma_n$—i.e. for deterministic states and effects—$(e_v | \varepsilon_f) = (a_{\,v,\Gamma_m} | \alpha_{f,\Gamma_n}) = 1$. Notice that for a generic system $\mathrm{n} \triangleright \mathrm{m}$ there are n different deterministic effects; since an OPT is causal if and only if for every system there is just a single deterministic effect [9], we conclude that the presented theory is non-causal.

To complete the theory, we need to specify all possible transformations. The set of transformations $\mathrm{Transf}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$ is built up starting from the atomic elements $\mathcal{F}^{t\,t'}_{s\,s'}$ with $(s, s', t, t') \in \Gamma_n \times \Gamma_m \times \Gamma_p \times \Gamma_q$ defined as $\mathcal{F}^{t\,t'}_{s\,s'} |\alpha_{v\,v'}) := \delta^{v}_{s}\delta^{v'}_{s'} |\alpha_{t\,t'})$ (see propositions 6, 7, 8, and 9 in the appendix). The other transformations belonging to $\mathrm{Transf}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$ are the elements $\mathcal{T}^{f\,g}_{\Omega} := \sum_{(s',t)\in\Omega} \mathcal{F}^{t\,g(t,s')}_{f(t)\,s'}$ with $\Omega \subseteq \Gamma_p \times \Gamma_m$, $f\colon \Gamma_p \to \Gamma_n$, and $g\colon \Gamma_p \times \Gamma_m \to \Gamma_q$ (see propositions 10 and 11 in the appendix). Notice that $\mathcal{F}^{t\,t'}_{s\,s'} \equiv \mathcal{T}^{f\,g}_{\{s'\}\times\{t\}}$ with $f(t) = s$ and $g(t, s') = t'$. The channels from $\mathrm{n} \triangleright \mathrm{m}$ to $\mathrm{p} \triangleright \mathrm{q}$ are the elements $\mathcal{T}^{f\,g} := \mathcal{T}^{f\,g}_{\Gamma_n \times \Gamma_p}$. This completes the construction of the full theory, which is deterministic and non-causal.

We can give now an explicit example which shows the non-causal features of the presented theory. Let us consider a simple case with the system $2 \triangleright 2$ and the experimenter Alice. Alice wants to prepare the system $2 \triangleright 2$ by means of the preparation test $\{ |\alpha_{f,\Xi_i}) \}_{i=0,1}$, with $\Xi_i := \{i\}$ for $i = 0, 1$, and f arbitrary function from Γ_2 to Γ_2. She subsequently measures the system chosing one observation test between $\mathcal{D}_0 := \{ (a_{\,0,\Xi_i}| \}_{i=0,1}$ and $\mathcal{D}_1 := \{ (a_{\,1,\Xi_i}| \}_{i=0,1}$. It can be easily seen that the probability of preparing the state $|\alpha_{f,\Xi_i})$ depends on which observation Alice wants to perform. Indeed,

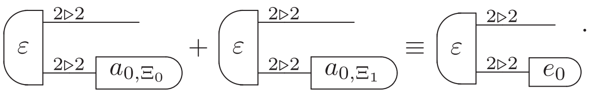
$$\begin{aligned}
\Pr[\alpha_{f,\Xi_0} \mid \mathcal{D}_0] &= (a_{\,0\Xi_0} | \alpha_{f,\Xi_0}) + (a_{\,0\Xi_1} | \alpha_{f,\Xi_0}) \\
&= (e_0 | \alpha_{f,\Xi_0}) = 1, \\
\Pr[\alpha_{f,\Xi_0} \mid \mathcal{D}_1] &= (a_{\,1,\Xi_0} | \alpha_{f,\Xi_0}) + (a_{\,1,\Xi_1} | \alpha_{f,\Xi_0}) = \\
&= (e_1 | \alpha_{f,\Xi_0}) = 0,
\end{aligned}$$

and similarly for the state $|\alpha_{f,\Xi_1})$.

We can moreover show how this deterministic non-causal theory violates the no-signalling without interaction, i.e.

[4] In previous literature [9] the same nomenclature has been used for the cone duality $\lambda\mathrm{Eff}(\mathrm{A}) = \lambda\mathrm{St}(\mathrm{A})^*$, which is a different concept.

by means of a bipartite deterministic state an experimenter Bob can communicate with Alice just with local measurements on his own subsystem. Let us suppose that both the systems of Alice and Bob are $2 \triangleright 2$, and that they share the bipartite deterministic state $|\,\varepsilon\,) \in \mathrm{St}(4 \triangleright 4)$. Keeping the same notation of the previous example, let us suppose that Bob can perform the two observation test $\mathcal{D}_0$, $\mathcal{D}_1$. It can be easily shown that, unlike in quantum theory, if $|\,\varepsilon\,)_{AB}$ is properly chosen the state Alice sees in her subsystem without knowing the outcome of the measurement performed by Bob (the so-called marginal state of Alice) will depend on the choice made by Bob. In this way Alice performing a local observation on her own subsystem can assess the choice of the measurement made on the other subsystem, getting therefore information from Bob. Indeed, if Bob performs the test $\mathcal{D}_0$ the marginal state of Alice will be



Let us choose as deterministic bipartite input the state $|\,\varepsilon\,) := \sum_{st} |\,\alpha_{s\,t}) \otimes |\,\alpha_{t\,s})$. We have that the marginal state of Alice when Bob performs the test $\mathcal{D}_0$ is

$$
\begin{aligned}
(e_0\,|_B|\,\varepsilon\,)_{AB} &= \sum_{st} |\,\alpha_{s\,t}) \otimes (e_0\,|\,\alpha_{t\,s}) \\
&= \sum_{st} \chi_{\Xi_t}(0)\,\chi_{\Gamma_2}(s)\,|\,\alpha_{s\,t}) = \sum_{s} |\,\alpha_{s\,0}),
\end{aligned}
$$

namely the deterministic state $|\,\varepsilon_{h_0}\,) \in \mathrm{St}(2 \triangleright 2)$ where h_0 is the function such that $h_0(x) := 0\ \ \forall\ x \in \Gamma_2$. Similarly, the marginal state of Alice when Bob performs the test $\mathcal{D}_1$ is $|\,\varepsilon_{h_1}\,)$—with $h_1(x) := 1\ \forall x \in \Gamma_2$. Alice can distinguish between the two marginal states $|\,\varepsilon_{h_0}\,)$, $|\,\varepsilon_{h_1}\,)$ by means of the test $\mathcal{D}_0$, assessing the choice of Bob.

The presented deterministic non-causal theory can also be built in a constructive way [14]. It is done in two steps. The first one consists in building a non-causal OPT through the addition of a non-causal shell around an internal causal core corresponding to the classical OPT, thanks to a construction analogous to that of quantum combs in the case of quantum theory [15, 16]. Then the resulting two-shell theory is constrained to be deterministic. An interesting result is that every transformation of the probabilistic non-causal core+shell theory can be implemented just using elements of the core causal theory [14].

4. Discussion

In summary, in this paper we have established the logical independence of the two notions of causality and determinism, which play a crucial role in physics, and stay at the core of the debate about foundations of quantum theory and relativity. As a legacy of classical physics the two concepts have been often merged into a single one that is unfit for quantum theory, thus leading to misconceptions, such as the J von Neumann argument in *Mathematical Foundations of Quantum Mechanics* (1932) that the transformation of *a quantum state … under the action of an energy operator … is purely causal*, while, *on the other hand, the state … which may measure a given quantity … undergoes in a measurement a non-causal change*. Here by determinism we simply mean that the probabilities of all events are either 0 or 1, whereas for causality we mean the usual Einstein notion, namely no-signalling from the future. We have proved that not only are the two notions formally independent, but also this distinction is not vacuous, since there are indeed both counterexamples of a theory which is causal and non-deterministic and, conversely, a theory which is deterministic and non-causal. Quantum theory provides the first example, while the second one has been introduced in the present paper (retro-causal deterministic theories are not a good example of non-causal deterministic theory, due to the mentioned isomorphism between causal and retro-causal theories).

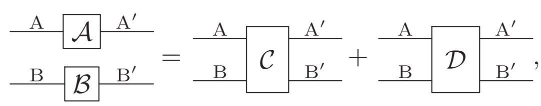
In conclusion of this paper, we want to comment about the relation between the notion of causality in definition 3 and the usual cause–effect relation and/or the physical causation in the philosophical literature, e. g. as in [1, 2]. After centuries of debate it may be said with a degree of confidence that an empirical notion of causality/causation is missing, and in all cases the cause–effect connection is of conterfactual nature. Causality should always be regarded as meaningless outside a theory. Within a theory definition 3 is the minimal requirement for the use of the term causality, and only if a theory is causal does it make sense to identify cause and effect, whereas in a non-causal theory the two words are non sensical. Obviously within definition 3 the preparation plays the role of the cause and the observation that of the effect. Finally, we want to stress that our definition is exactly the Einsteinian one in special relativity theory.

Appendix

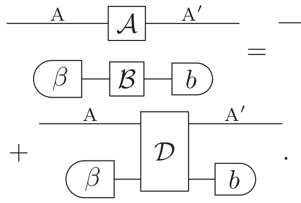
In this appendix we present all the technical results which ensure the consistency of the non-causal deterministic theory presented in section 3.

Proposition 1. *If local discriminability holds, the parallel composition of two atomic transformations is atomic.*

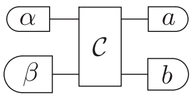
Proof. Let $\mathcal{A} \in \mathrm{Transf}(\mathrm{A}, \mathrm{A}')$, $\mathcal{B} \in \mathrm{Transf}(\mathrm{B}, \mathrm{B}')$ be atomic transformations between systems. Let us consider the transformation $\mathcal{A} \otimes \mathcal{B} \in \mathrm{Transf}(\mathrm{AB}, \mathrm{A}'\mathrm{B}')$. Let us suppose that $\mathcal{A} \otimes \mathcal{B}$ can be decomposed as follows

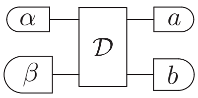


for a non-trivial couple of transformations $0 \neq \mathcal{C}, \mathcal{D} \in \mathrm{Transf}(\mathrm{AB}, \mathrm{A}'\mathrm{B}')$. For any state $|\beta) \in \mathrm{St}(\mathrm{B})$, and any effect $(b| \in \mathrm{Eff}(\mathrm{B}')$ such that $(b|\mathcal{B}|\beta) \neq 0$, we have

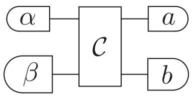


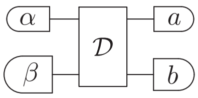
Since the transformation $\mathcal{A}$ is atomic we have that the transformations $(b|_{\mathrm{B}}\mathcal{C}|\beta)_{\mathrm{B}'}, (b|_{\mathrm{B}}\mathcal{D}|\beta)_{\mathrm{B}'} \in \mathrm{Transf}(\mathrm{A},\mathrm{A}')$ must be proportional to $\mathcal{A}$; in particular for any state $|\alpha) \in \mathrm{St}(\mathrm{A})$, and any effect $(a| \in \mathrm{Eff}(\mathrm{A}')$ such that $(a|\mathcal{A}|\alpha) \neq 0$, it must be

 $= \mu^{\mathcal{C}}_{b\beta}(a|\mathcal{A}|\alpha),$ (A.1)

 $= \mu^{\mathcal{D}}_{b\beta}(a|\mathcal{A}|\alpha),$ (A.2)

where $\mu^{\mathcal{C}}_{b\beta}, \mu^{\mathcal{D}}_{b\beta}$ are constants which can depend on the choice of $|\beta)$ and $(b|$. One can repeat a similar argument on the other subsystem, getting:

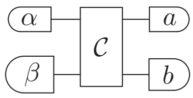
 $= \lambda^{\mathcal{C}}_{a\alpha}(b|\mathcal{B}|\beta),$ (A.3)

 $= \lambda^{\mathcal{D}}_{a\alpha}(b|\mathcal{B}|\beta),$ (A.4)

where $\lambda^{\mathcal{C}}_{a\alpha}, \lambda^{\mathcal{D}}_{a\alpha}$ are constants which can depend on the choice of $|\alpha)$ and $(a|$. Let us now suppose that $\lambda^{\mathcal{C}}_{a\alpha} = 0$. Then we have

$$\lambda^{\mathcal{C}}_{a\alpha}(b|\mathcal{B}|\beta) = \mu^{\mathcal{C}}_{b\beta}(a|\mathcal{A}|\alpha) = 0,$$

for all $(b|, |\beta)$. Since by hypothesis $(a|\mathcal{A}|\alpha) \neq 0$, we have $\mu^{\mathcal{C}}_{b\beta} = 0$ for all $(b|, |\beta)$, and finally this implies that

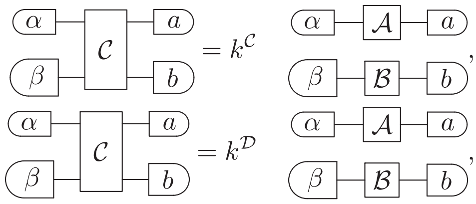
 $= 0,$

for all $(a|, (b|, |\alpha), |\beta)$, namely, by local discriminability, $\mathcal{C} = 0$, contrarily to the hypothesis. By similar arguments we can then prove that the coefficients $\lambda^{\mathcal{C}}_{a\alpha}, \lambda^{\mathcal{D}}_{a\alpha}, \mu^{\mathcal{C}}_{b\beta}$, and $\mu^{\mathcal{D}}_{b\beta}$ are all positive.

Comparing equation (A.1) with equation (A.3), and equation (A.2) with equation (A.2) one obtains:

$$\frac{\lambda^{\mathcal{C}}_{a\alpha}}{(a|\mathcal{A}|\alpha)} = \frac{\mu^{\mathcal{C}}_{b\beta}}{(b|\mathcal{B}|\beta)} > 0, \qquad \frac{\lambda^{\mathcal{D}}_{a\alpha}}{(a|\mathcal{A}|\alpha)} = \frac{\mu^{\mathcal{D}}_{b\beta}}{(b|\mathcal{B}|\beta)} > 0.$$

The previous relations show that all the ratios are independent of the choices of $|\alpha)$, $|\beta)$, $(a|$, $(b|$, i.e. $k^{\mathcal{C}} := \lambda^{\mathcal{C}}_{a\alpha}/(a|\mathcal{A}|\alpha) = \mu^{\mathcal{C}}_{b\beta}/(b|\mathcal{B}|\beta)$ and $k^{\mathcal{D}} := \lambda^{\mathcal{D}}_{a\alpha}/(a|\mathcal{A}|\alpha) = \mu^{\mathcal{D}}_{b\beta}/(b|\mathcal{B}|\beta)$. Using these definitions for $k^{\mathcal{C}}$ and $k^{\mathcal{D}}$ in equations (A.1), (A.2) one gets



for all $|\alpha), |\beta), (a|, (b|$. By local discriminability this implies $k^{\mathcal{C}}\mathcal{A} \otimes \mathcal{B} = \mathcal{C}$, and $k^{\mathcal{D}}\mathcal{A} \otimes \mathcal{B} = \mathcal{D}$, namely $\mathcal{A} \otimes \mathcal{B}$ is atomic. $\square$

Proposition 2. *Let* $\{|\alpha_{s\,t})\}_{(s,t)\in\Gamma_n\times\Gamma_m} \subset \mathrm{St}(\mathrm{n} \triangleright \mathrm{m})$ *the atomic states of the system* $\mathrm{n} \triangleright \mathrm{m}$*; similarly let* $\{|\alpha'_{s'\,t'})\}_{(s',t')\in\Gamma_{n'}\times\Gamma_{m'}} \subset \mathrm{St}(\mathrm{n}' \triangleright \mathrm{m}')$ *the atomic states of the system* $\mathrm{n}' \triangleright \mathrm{m}'$. *Then, the atomic states of the composite system* $\mathrm{x} \triangleright \mathrm{y} := (\mathrm{n} \triangleright \mathrm{m})(\mathrm{n}' \triangleright \mathrm{m}')$ *are the elements* $|\alpha_{s\,t}) \otimes |\alpha'_{s'\,t'})$.

Proof. By definition, the system $\mathrm{x} \triangleright \mathrm{y}$ has $x \times y$ atomic states, and since $\mathrm{x} \triangleright \mathrm{y} = (\mathrm{n} \triangleright \mathrm{m})(\mathrm{n}' \triangleright \mathrm{m}')$ we have $x \times y = n \times m \times n' \times m'$. Since the states $|\alpha_{s\,t}) \otimes |\alpha'_{s'\,t'}) \in \mathrm{St}(\mathrm{x} \triangleright \mathrm{y})$ are atomic (see proposition 1), different from each other, and their cardinality is exactly $n \times m \times n' \times m'$, we conclude that they are the atomic states of $\mathrm{St}(\mathrm{x} \triangleright \mathrm{y})$. $\square$

Proposition 3. *A linear map* $\mathcal{T} \in \mathrm{Transf}_{\mathbb{R}}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$ *is admissible if and only if is locally admissible, i.e.* $\mathcal{T}(\mathrm{St}(\mathrm{n} \triangleright \mathrm{m})) \subseteq \mathrm{St}(\mathrm{p} \triangleright \mathrm{q})$.

Proof. First, let us recall that a map $\mathcal{T} \in \mathrm{Transf}_{\mathbb{R}}(\mathrm{A}, \mathrm{A}')$ is admissible if and only if $\mathcal{T} \otimes \mathcal{I}_{\mathrm{B}}(\mathrm{St}(\mathrm{AB})) \subseteq \mathrm{St}(\mathrm{A}'\mathrm{B})$ for every system B. Let us prove the equivalence for the deterministic non-causal theory in two steps.

$(\Rightarrow)$: this implication is trivial and it always holds, regardless of the theory involved; i.e. local admissibility can be derived from the admissibility taking the system B to be the trivial one I.

$(\Leftarrow)$: the linear map $\mathcal{T} \in \mathrm{Transf}_{\mathbb{R}}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$ is locally admissible by hypothesis, therefore for any atomic state $|\alpha_{s\,s'\in \mathrm{St}})(\mathrm{n} \triangleright \mathrm{m})$ we have $\mathcal{T}|\alpha_{s\,s'}) = |\alpha_{f^{ss'}\Xi^{ss'}}) \in \mathrm{St}(\mathrm{p} \triangleright \mathrm{q})$, where $f^{ss'}\colon \Xi^{ss'} \subseteq \Gamma_p \to \Gamma_q$. Notice that, since for $s_0 \neq s_1$ the state $|\alpha_{s_0\,s_0'}) + |\alpha_{s_1\,s_1'})$ is valid, then by local admissibility also $\mathcal{T}\,|\alpha_{s_0\,s_0'}) + |\alpha_{s_1\,s_1'}) = |\alpha_{f^{s_0s_0'}\Xi^{s_0s_0'}}) + |\alpha_{f^{s_1s_1'}\Xi^{s_1s_1'}})$ is a valid state, therefore we must have that

$$\Xi^{s_0s_0'} \bigcap \Xi^{s_1s_1'} = \varnothing \qquad \forall\, s_0', \forall\, s_1' \text{ and } s_0 \neq s_1. \tag{A.5}$$

For an arbitrary system $\mathrm{n}' \triangleright \mathrm{m}'$, let us choose freely the state $|\alpha_{g\Upsilon})$ of the composite system $\mathrm{x} \triangleright \mathrm{y} := (\mathrm{n} \triangleright \mathrm{m})(\mathrm{n}' \triangleright \mathrm{m}')$. It can be expanded on the atomic multipartite states $|\alpha_{s\,s'}) \otimes |\alpha_{t\,t'})$—with $|\alpha_{s\,s'\in \mathrm{St}})(\mathrm{n} \triangleright \mathrm{m})$,

$|\alpha'_{t\,t'\in\mathrm{St}})(\mathrm{n}'\triangleright\mathrm{m}')$—as $|\alpha_{g\Upsilon}) = \sum_{ss'tt'}\alpha_{ss'tt'}|\alpha_{s\,s'})\otimes|\alpha'_{t\,t'})$ with $\alpha_{ss'tt'} := \delta_{s'g_1(s,t)}\delta_{t'g_2(s,t)}\chi_\Upsilon(s,t)$, for a couple of functions $g_1\colon \Upsilon\to\Gamma_m$, $g_2\colon \Upsilon\to\Gamma_{m'}$ such that $g(s,t) = (g_1(s,t), g_2(s,t))$. On such an arbitrary multipartite state the map $\mathcal{T}\otimes\mathcal{I}_{\mathrm{n}'\triangleright\mathrm{m}'}$ leads to a valid state of the composite system $\mathrm{x}'\triangleright\mathrm{y}' := (\mathrm{p}\triangleright\mathrm{q})(\mathrm{n}'\triangleright\mathrm{m}')$:

$$\begin{aligned}
&[\mathcal{T}\otimes\mathcal{I}_{\mathrm{n}'\triangleright\mathrm{m}'}]\,|\alpha_{g\Upsilon})\\
&= \sum_{ss'tt'}\alpha_{ss'tt'}\mathcal{T}|\alpha_{s\,s'})\otimes\mathcal{I}_{\mathrm{n}'\triangleright\mathrm{m}'}|\alpha_{t\,t'})\\
&= \sum_{ss'tt'}\alpha_{ss'tt'}|\alpha_{f^{ss'}\Xi^{ss'}})\otimes|\alpha_{t\,t'})\\
&= \sum_{ss'tt'vv'}\alpha_{ss'tt'}\delta_{v'f^{ss'}(v)}\chi_{\Xi^{ss'}}(v)|\alpha_{v\,v'})\otimes|\alpha_{t\,t'})\\
&= \sum_{ss'}\sum_{tt'vv'}\delta_{s'g_1(s,t)}\delta_{t'g_2(s,t)}\chi_\Upsilon(s,t)\delta_{v'f^{ss'}(v)}\chi_{\Xi^{ss'}}(v)\\
&\quad\times|\alpha_{v\,v'})\otimes|\alpha_{t\,t'}).
\end{aligned}\tag{A.6}$$

The most internal sum represents the valid state $|\alpha_{h^{ss'}\Delta^{ss'}})\in\mathrm{St}(\mathrm{x}'\triangleright\mathrm{y}')$ with $h^{ss'}\colon \Delta^{ss'}\to\Gamma_q\times\Gamma_{m'}$, where $\Delta^{ss'}\subseteq\Gamma_p\times\Gamma_{n'}$ is defined by $\chi_{\Delta^{ss'}}(x,y) := \delta_{s'g_1(s,y)}\chi_\Upsilon(s,y)\chi_{\Xi^{ss'}}(x)$, and $h^{ss'}(x,y) := (h_1^{ss'}(x,y), h_2^{ss'}(x,y))$, $h_1^{ss'}(x,y) := f^{ss'}(x)$, $h_2^{ss'}(x,y) := g_2(s,y)$. Hence the relation of equation (A.6) can be rewritten as $[\mathcal{T}\otimes\mathcal{I}_{\mathrm{n}'\triangleright\mathrm{m}'}]\,|\alpha_{g\Upsilon}) = \sum_{ss'}|\alpha_{h^{ss'}\Delta^{ss'}})$. This sum represents a valid state for $\mathrm{St}(\mathrm{x}'\triangleright\mathrm{y}')$ since the various $\Delta^{ss'}$ are disjoint: let us take two sets $\Delta^{s_0s_0'}$, $\Delta^{s_1s_1'}$, and evaluate $\chi_{\Delta^{s_0s_0'}\cap\Delta^{s_1s_1'}}\equiv\chi_{\Delta^{s_0s_0'}}(x,y)\chi_{\Delta^{s_1s_1'}}(x,y)$. If $s_0 = s_1$ we have

$$\begin{aligned}
&\delta_{s_0'g_1(s_0,y)}\delta_{s_1'g_1(s_0,y)}\chi_\Upsilon^2(s_0,y)\chi_{\Xi^{s_0s_0'}}(x)\chi_{\Xi^{s_0s_1'}}(x)\\
&\quad=\delta_{s_0's_1'}\delta_{s_0'g_1(s_0,y)}\chi_\Upsilon^2(s_0,y)\chi_{\Xi^{s_0s_0'}}(x)\chi_{\Xi^{s_0s_1'}}(x)
\end{aligned}$$

which is equal to zero when $s_0'\neq s_1'$—thanks to the first Kronecker's delta. On the other hand if $s_0\neq s_1$ we have

$$\begin{aligned}
&\delta_{s_0'g_1(s_0,y)}\delta_{s_1'g_1(s_1,y)}\chi_\Upsilon(s_0,y)\chi_\Upsilon(s_1,y)\chi_{\Xi^{s_0s_0'}}(x)\chi_{\Xi^{s_1s_1'}}(x)\\
&\delta_{s_0'g_1(s_0,y)}\delta_{s_1'g_1(s_1,y)}\chi_\Upsilon(s_0,y)\chi_\Upsilon(s_1,y)\chi_{\Xi^{s_0s_0'}\cap\Xi^{s_1s_1'}}(x),
\end{aligned}$$

which is always equal to zero thanks to equation (A.5), which implies $\chi_{\Xi^{s_0s_0'}\cap\Xi^{s_1s_1'}}(x) = 0$. □

From now on, all the admissibility proofs will be reduced to local admissibility, thanks to proposition 3.

Proposition 4. *Under the no-restriction hypothesis the atomic effects of* $\mathrm{n}\triangleright\mathrm{m}$ *are the elements* $(a_{s\,s'}|$ *of* $\mathrm{Eff}_{\mathbb{R}}(\mathrm{n}\triangleright\mathrm{m})$ *with* $(s,s')\in\Gamma_n\times\Gamma_m$ *such that* $(a_{s\,s'}|\alpha_{t\,t'}) = \delta_{st}\delta_{s't'}\ \forall\, s,t\in\Gamma_n$ *and* $\forall\, s', t'\in\Gamma_m$.

Proof. The proof goes in three simple steps: first we show that the elements $(a_{s\,s'}|$ are admissible. After showing that they are also linearly independent (therefore they span all the set $\mathrm{Eff}_{\mathbb{R}}(\mathrm{n}\triangleright\mathrm{m})$) we show that every effect $(c\,|$ for the system $\mathrm{n}\triangleright\mathrm{m}$ can be written as $(c\,| = \sum_{ij}c_{ss'}(a_{s\,s'}|$ with $c_{ss'}$ non-negative, proving that the set of atomic effects coincides with the set $\{(a_{s\,s'}|\}_{(s,s')\in\Gamma_n\times\Gamma_m}$.

The effects $(a_{s\,s'}|$ are locally admissible, since for every state $|\alpha_{f,\Xi})$

$$(a_{s\,s'}|\alpha_{f,\Xi}) = \sum_{tt'}\chi_\Xi(t)\delta_{t'f(t)}(a_{s\,s'}|\alpha_{t\,t'}) = \chi_\Xi(s)\delta_{s'f(s)},$$

which is an admissible probability $p\in\{0,1\}$. Thanks to proposition 3, the $(a_{s\,s'}|$ are admissible, and by the no-restriction hypothesis they belong to $\mathrm{Eff}(\mathrm{n}\triangleright\mathrm{m})$.

Now, let us show that a null linear combination of the elements $(a_{t\,t'}|$—say $(c\,| = \sum_{tt'}c_{tt'}(a_{t\,t'}|$—necessarily has $c_{tt'} = 0\ \ \forall\, t\in\Gamma_n,\ \forall\, t'\in\Gamma_m$. Indeed, for any atomic state $|\alpha_{s\,s'})$ we get

$$0 = (c\,|\alpha_{s\,s'}) = \sum_{tt'}c_{tt'}(a_{t\,t'}|\alpha_{s\,s'}) = c_{ss'},$$

for every s, s', i.e. all the $(a_{t\,t'}|$ are linearly independent. We have that the number of different effects $(a_{t\,t'\in\mathrm{Eff}}|(\mathrm{n}\triangleright\mathrm{m})$ is $n\cdot m$, as many as $\dim\mathrm{St}_{\mathbb{R}}(\mathrm{n}\triangleright\mathrm{m}) = \dim\mathrm{Eff}_{\mathbb{R}}(\mathrm{n}\triangleright\mathrm{m}) = n\cdot m$: we conclude that the effects $(a_{t\,t'}|\in\mathrm{Eff}(\mathrm{n}\triangleright\mathrm{m})$ span the whole linear space $\mathrm{Eff}_{\mathbb{R}}(\mathrm{n}\triangleright\mathrm{m})$.

The third step is easily proven noticing that an arbitrary effect $(c\,| = \sum_{tt'}c_{tt'}(a_{t\,t'}|$ is a $\{0,1\}$-functional over the states. Since $(c\,|\alpha_{i\,j}) = c_{ij}\ \forall\, i\in\Gamma_n,\ \forall\, j\in\Gamma_m$, we conclude that every effect is a conic combination of the elements $(a_{t\,t'}|$ with coefficients 0 or 1. Since linear combination with negative coefficients is forbidden we conclude that all the effects $(a_{t\,t'}|$ are atomic. For the same reason, there are no other atomic effects in $\mathrm{Eff}(\mathrm{n}\triangleright\mathrm{m})$. □

Proposition 5. *Under the no-restriction hypothesis the effects of the system* $\mathrm{n}\triangleright\mathrm{m}$ *are the elements* $(a_{v,E}| := \sum_{i\in E}(a_{v\,i}|$, *with* $i\in\Gamma_n$, $E\subseteq\Gamma_m$.

Proof. The proof proceeds in two steps. First of all we prove that the elements $(a_{v,E}|\in\mathrm{Eff}_{\mathbb{R}}(\mathrm{n}\triangleright\mathrm{m})$ are valid effects for the system $\mathrm{n}\triangleright\mathrm{m}$. Then we prove that there are no further effects in $\mathrm{Eff}(\mathrm{n}\triangleright\mathrm{m})$.

We only need to prove that the elements $(a_{vE}|\in\mathrm{Eff}_{\mathbb{R}}(\mathrm{n}\triangleright\mathrm{m})$ are locally admissible, and therefore they are admissible by proposition 3. Finally, this implies that they belong to $\mathrm{Eff}(\mathrm{n}\triangleright\mathrm{m})$ thanks to the no-restriction hypothesis.

The effects $(a_{v,E}|$ are locally admissible, since for every state $|\alpha_{f,\Xi})$ we have

$$(a_{v,E}|\alpha_{f,\Xi}) = \chi_E(f(v))\chi_\Xi(v),$$

which is an admissible probability $p\in\{0,1\}$.

Now let us prove that there are no other effects apart from $(a_{vE}|$. Given an effect $(c\,|\in\mathrm{Eff}(\mathrm{n}\triangleright\mathrm{m})$, thanks to proposition 4 we know it can be expanded over the atomic effects $(a_{t\,t'}|$ as $(c\,| = \sum_{tt'}c_{tt'}(a_{t\,t'}|$ with $c_{tt'} = 0, 1$, $t\in\Gamma_n$, and $t'\in\Gamma_m$. Suppose by contradiction that there exists a valid effect $(c\,| = \sum_{tt'}c_{tt'}(a_{t\,t'}|$ with $c_{ij} = c_{i'j'} = 1$ for some j, j' and

$i \neq i'$. Let us take the deterministic state $|\varepsilon_f) \in \mathrm{St}(\mathrm{n} \triangleright \mathrm{m})$ with $f(i) = j$ and $f(i') = j'$; we have that $(c|\varepsilon_f) \geqslant 2$, an absurd. $\square$

Proposition 6. *Under the no-restriction hypothesis, the linear maps* $\mathcal{F}^{t\,t'}_{s\,s'} \in \mathrm{Transf}_{\mathbb{R}}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$ *with* $(s, s', t, t') \in \Gamma_n \times \Gamma_m \times \Gamma_p \times \Gamma_q$ *such that* $\mathcal{F}^{t\,t'}_{s\,s'}|\alpha_{v\,v'}) = \delta_{sv}\delta_{s'v'}|\alpha_{t\,t'})$*, are valid transformations.*

Proof. We just need to check that the maps $\mathcal{F}^{t\,t'}_{s\,s'}$ are locally admissible, and then by proposition 3 and the no-restriction hypothesis, we conclude that they actually belong to $\mathrm{Transf}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$. Indeed, for every state $|\alpha_{f,\Xi})$, we have

$$\mathcal{F}^{t\,t'}_{s\,s'}|\alpha_{f\Xi}) = \chi_{\Xi}(s)\delta_{s'f(s)}|\alpha_{t\,t'}),$$

which is a valid state of $\mathrm{p} \triangleright \mathrm{q}$. $\square$

Proposition 7. *The transformations* $\mathcal{F}^{t\,t'}_{s\,s'} \in \mathrm{Transf}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$ *are linearly independent.*

Proof. Let us show that a null linear combination of the transformations $\mathcal{F}^{t\,t'}_{s\,s'} \in \mathrm{Transf}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$—say $\mathcal{A} = \sum_{ss'tt'} c_{ss'tt'}\mathcal{F}^{t\,t'}_{s\,s'}$—necessarily has $c_{ss'tt'} = 0$, for all $s \in \Gamma_n$, $s' \in \Gamma_m$, $t \in \Gamma_p$, $t' \in \Gamma_q$. Indeed, for any couple $|\alpha_{i\,i'}) \in \mathrm{St}(\mathrm{n} \triangleright \mathrm{m})$, $(a_{j\,j'} \in \mathrm{St}|(\mathrm{p} \triangleright \mathrm{q})$ we have

$$0 = (a_{j\,j'}|\mathcal{A}|\alpha_{i\,i'=c})_{ii'jj'},$$

for every $(i, i', j, j') \in \Gamma_n \times \Gamma_m \times \Gamma_p \times \Gamma_q$, i.e. the transformations $\mathcal{F}^{t\,t'}_{s\,s'} \in \mathrm{Transf}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$ are linearly independent. $\square$

Proposition 8. *The transformations* $\mathcal{F}^{t\,t'}_{s\,s'} \in \mathrm{Transf}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$ *are atomic.*

Proof. Let us suppose by contradiction that the transformation $\mathcal{F}^{t\,t'}_{s\,s'}$ is not atomic, namely $\mathcal{F}^{t\,t'}_{s\,s'} = \mathcal{A} + \mathcal{B}$ for some $\mathcal{A}, \mathcal{B} \in \mathrm{Transf}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$. For an arbitrary state $|\alpha_{f\Xi}) \in \mathrm{St}(\mathrm{n} \triangleright \mathrm{m})$ we have that $\mathcal{F}^{t\,t'}_{s\,s'}|\alpha_{f,\Xi}) = \chi_{\Xi}(s)\,\delta_{s'f(s)}\,|\alpha_{t\,t'})$, $\mathcal{A}|\alpha_{f,\Xi}) = |\alpha\mathcal{A}) = \sum_{ss'} c^{\mathcal{A}}_{ss'}|\alpha_{s\,s'})$, $\mathcal{B}|\alpha_{f,\Xi}) = |\alpha\mathcal{B}) = \sum_{ss'} c^{\mathcal{B}}_{ss'}|\alpha_{s\,s'})$, where we have expanded the states $|\alpha\mathcal{A}), |\alpha\mathcal{B}) \in \mathrm{St}(\mathrm{p} \triangleright \mathrm{q})$ over the atomic states $|\alpha_{s\,s'})$ of the system $\mathrm{p} \triangleright \mathrm{q}$. By hypothesis we have $\mathcal{F}^{t\,t'}_{s\,s'}|\alpha_{f\Xi}) = \mathcal{A}|\alpha_{f\Xi}) + \mathcal{B}|\alpha_{f\Xi})$, namely

$$\chi_{\Xi}(s)\delta_{s'f(s)}|\alpha_{t\,t'}) = \sum_{ss'} c^{\mathcal{A}}_{ss'}|\alpha_{s\,s'}) + \sum_{ss'} c^{\mathcal{B}}_{ss'}|\alpha_{s\,s'}).$$

Since the atomic states $|\alpha_{s\,s'})$ are linearly independent we have that the previous relation can be rewritten as

$$c^{\mathcal{A}}_{ss'} + c^{\mathcal{B}}_{ss'} = \chi_{\Xi}(s)\delta_{s'f(s)} \qquad \text{if } t = s t' = s'$$
$$c^{\mathcal{A}}_{ss'} + c^{\mathcal{B}}_{ss'} = 0 \qquad \text{otherwise}.$$

Since $c^{\mathcal{A}}_{ss'}, c^{\mathcal{B}}_{ss'} = 0, 1$, we conclude from the second relation that $c^{\mathcal{A}}_{ss'} = c^{\mathcal{B}}_{ss'} = 0$ if $t \neq s$ or $t' \neq s'$, while the first leads to $c^{\mathcal{A}}_{tt'} = \chi_{\Xi}(s)\,\delta_{s'f(s)}$ and $c^{\mathcal{B}}_{tt'} = 0$ (or the other way round). Since the initial state $|\alpha_{f\Xi}) \in \mathrm{St}(\mathrm{n} \triangleright \mathrm{m})$ is arbitrary we conclude that $\mathcal{F}^{t\,t'}_{s\,s'} = \mathcal{A} + 0$ (or $\mathcal{F}^{t\,t'}_{s\,s'} = 0 + \mathcal{B}$), i.e. $\mathcal{F}^{t\,t'}_{s\,s'}$ is atomic. $\square$

Proposition 9. *There are no atomic transformations in* $\mathrm{Transf}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$ *other than* $\mathcal{F}^{t\,t'}_{s\,s'}$.

Proof. Since the dimension of $\mathrm{Transf}_{\mathbb{R}}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$ is $\dim\mathrm{St}_{\mathbb{R}}(\mathrm{n} \triangleright \mathrm{m}) \times \dim\mathrm{St}_{\mathbb{R}}(\mathrm{p} \triangleright \mathrm{q}) = n \times m \times p \times q$, and the number of (linearly independent) atomic transformations $\mathcal{F}^{t\,t'}_{s\,s'}$ is $n \times m \times p \times q$ we conclude that such atomic maps span the entire space of linear transformations between the two linear spaces of states.

Now let us suppose by contradiction that there exists another atomic transformation $\mathcal{T} \in \mathrm{Transf}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$, different from any of $\mathcal{F}^{t\,t'}_{s\,s'}$. Since the maps $\mathcal{F}^{t\,t'}_{s\,s'}$ span all the space, we expand $\mathcal{T}$ over them:

$$\mathcal{T} = \sum_{ss'tt'} c^{ss'}_{tt'}\mathcal{F}^{t\,t'}_{s\,s'}.$$

Since $\mathcal{T}$ is atomic, it has to lie out of the cone built from the transformations $\mathcal{F}^{t\,t'}_{s\,s'}$; hence at least one of the coefficients $c^{ss'}_{tt'}$ is negative. Since $(a_{t\,t'}|\mathcal{T}|\alpha_{s\,s'}) = c^{ss'}_{tt'}$ is a probability, we have that $0 \leqslant c^{ss'}_{tt'} \leqslant 1$, i.e. there are no atomic transformations other than $\mathcal{F}^{t\,t'}_{s\,s'}$. $\square$

Proposition 10. *If the no-restriction hypothesis holds, the transformations* $\mathrm{Transf}_{\mathbb{R}}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q}) \ni \mathcal{T}^{f\,g}_{\Omega} := \sum_{(s',t)\in\Omega} \mathcal{F}^{t\,g(t,s')}_{f(t)\,s'}$ *with* $\Omega \subseteq \Gamma_p \times \Gamma_m$, $f\colon \Gamma_p \to \Gamma_n$, *and* $g\colon \Gamma_p \times \Gamma_m \to \Gamma_q$ *actually belong to* $\mathrm{Transf}(\mathrm{n} \triangleright \mathrm{m}, \mathrm{p} \triangleright \mathrm{q})$.

Proof. By proposition 3 and the no-restriction hypothesis, we just need to show that the linear maps $\mathcal{T}^{f\,g}_{\Omega}$ are locally admissible.

For an arbitrary state $|\alpha_{h,\Xi})$ we have

$$\begin{aligned}\mathcal{T}^{f\,g}_{\Omega}|\alpha_{h,\Xi}) &= \sum_{(s',t)\in\Omega} \mathcal{F}^{t\,g(t,s')}_{f(t)\,s'}|\alpha_{h,\Xi}) \\ &= \sum_{s'}\sum_{t}\chi_{\Omega}(s', t)\chi_{\Xi}(f(t))\delta_{s'h(f(t))}|\alpha_{t\,g(t,s')}).\end{aligned}$$

The internal sum represents the state $|\alpha_{g_{s'}\Upsilon_{s'}}) \in \mathrm{St}(\mathrm{p} \triangleright \mathrm{q})$ with $g_{s'}\colon \Gamma_p \to \Gamma_q$, $g_{s'}(x) := g(s', x)$ and the set $\Upsilon_{s'} \subseteq \Gamma_p$ defined by $\chi_{\Upsilon_{s'}}(x) := \chi_{\Omega}(s', x)\chi_{\Xi}(f(x))\delta_{s'h(f(x))}$. The whole sum $\sum_{s'}|\alpha_{g_{s'}\Upsilon_{s'}})$ represents a valid state of $\mathrm{p} \triangleright \mathrm{q}$, indeed for every

$s_0' \neq s_1'$ the sets $\Upsilon_{s_0'}$, $\Upsilon_{s_1'}$ are disjoint since

$$\begin{aligned}\chi_{\Upsilon_{s_0'}\cap\Upsilon_{s_1'}}(x) &= \chi_{\Upsilon_{s_0'}}(x)\chi_{\Upsilon_{s_1'}}(x)\\ &= \chi_\Omega(s_0', x)\chi_\Omega(s_1', x)\chi_\Xi^2(f(x))\delta_{s_0' h(f(x))}\delta_{s_1' h(f(x))}\\ &= \chi_\Omega(s_0', x)\chi_\Omega(s_1', x)\chi_\Xi^2(f(x))\delta_{s_0' h(f(x))}\delta_{s_0' s_1'},\end{aligned}$$

which is equal to zero when $s_0' \neq s_1'$ thanks to the last Kronecker's delta. □

Proposition 11. *All the elements of* Transf(n ⊳ m, p ⊳ q) *have necessarily the form:* $\mathcal{T}_\Omega^{f\,g} := \sum_{(s',t)\in\Omega}\mathcal{F}^{t\,g(t,\,s')}_{f(t)\,s'}$ *with* $\Omega \subseteq \Gamma_p \times \Gamma_m$, $f\colon \Gamma_p \to \Gamma_n$, *and* $g\colon \Gamma_p \times \Gamma_m \to \Gamma_q$.

Proof. Given a generic transformation $\mathcal{T} = \sum_{ss'tt'} c_{ss'tt'}\mathcal{F}^{t\,t'}_{s\,s'}$, we have that $(a_{j\,j'}|\mathcal{T}|\alpha_{i\,i'}) = c_{ii'jj'}$. Since $c_{ii'jj'}$ is a probability in a deterministic theory, we have $(c_{ss'tt'} = 0) \vee (c_{ss'tt'} = 1)$, $\forall\ (s, s', t, t') \in \Gamma_n \times \Gamma_m \times \Gamma_p \times \Gamma_q$.

By contradiction, let us suppose that the transformation $\mathcal{T} = \sum_{ss'tt'} c_{ss'tt'}\mathcal{F}^{t\,t'}_{s\,s'}$ with $c_{ii'jj'} = c_{ki'jl'} = 1$ with $i \neq k$, $j' \neq l'$. Let $h\colon \Gamma_n \to \Gamma_m$ with $h(x) = i'\ \ \forall\, x \in \Gamma_n$, then we have $(e_j|\mathcal{T}|\varepsilon_h) \geqslant 2$, i.e. an absurd.

In such a way we have not ruled out the case $\sum_{(s',t)\in\Omega}\mathcal{F}^{t\,g(t,\,s')}_{f(t,\,s')\,s'}$. A transformation of this last form must have a couple of coefficients such that $c_{ii'jj'} = c_{kk'jl'} = 1$ with $i \neq k$, $i' \neq k'$, $j' \neq l'$, otherwise the functional dependence of f on the variable s' would be trivial. Let $h\colon \Gamma_n \to \Gamma_m$ with $h(i) = i'$, $h(k) = k'$; then we have $(e_j|\mathcal{T}|\varepsilon_h) \geqslant 2$, i.e. again an absurd. □

References

[1] Salmon W 1998 *Causality and Explanation: Oxford Scholarship Online* (Oxford: Oxford University Press)
[2] Dowe P 2007 *Physical Causation: Cambridge Studies in Probability, Induction and Decision Theory* (Cambridge: Cambridge University Press)
[3] von Neumann J 1932 *Mathematische Grundlagen der Quantenmechanik* (Berlin: Springer) chapter 4
von Neumann J 1932 *Mathematical Foundations of Quantum Mechanics* (Princeton University Press) (Engl. transl.)
[4] Einstein A, Podolsky B and Rosen N 1935 *Phys. Rev.* **47** 777–80
[5] Hardy L 2001 quant-ph/0101012
[6] Planck M 1941 *Der Kausalbegriff in der Physik* (Mostbach: Physik Verlag)
[7] Jaeger G 2012 *AIP Conf. Proc.* **1424** 387–94
[8] Pearl J 2000 *Causality: Models, Reasoning and Inference* vol 29 (Cambridge: Cambridge University Press)
[9] Chiribella G, D'Ariano G M and Perinotti P 2010 *Phys. Rev.* A **81** 062348
[10] Russell B 1912 *Proc. Aristotelian Soc.* **13** 1–26
[11] Chiribella G, D'Ariano G M and Perinotti P 2011 *Phys. Rev.* A **84** 012311
[12] D'Ariano G M 2010 Probabilistic theories: what is special about quantum mechanics? *Philosophy of Quantum Information and Entanglement* ed A Bokulich and G Jaeger (Cambridge: Cambridge University Press) chapter 5
[13] D'Ariano G M 2007 *AIP Conf. Proc.* **962** 44–55
[14] D'Ariano G M, Manessi F and Perinotti P 2013 in preparation
[15] Chiribella G, D'Ariano G M and Perinotti P 2008 *Europhys. Lett.* **83** 30004
[16] Chiribella G, D'Ariano G M and Perinotti P 2009 *Phys. Rev.* A **80** 022339
[17] Plotnitsky A 2009 *AIP Conf. Proc.* **1101** 150–60