ISSN 2282-7765 [online]
ISSN 2282-7757 [printed]

Volume 10, Number 2, December 2022

# Science & Philosophy

Journal of Epistemology, Science and Philosophy



Accademia Piceno-Aprutina dei Velati in Teramo

**SCIENCE & PHILOSOPHY - JOURNAL OF EPISTEMOLOGY, SCIENCE AND PHILOSOPHY**

ISSN 2282-7757; eISSN 2282-7765.

Science & Philosophy is an International, peer-reviewed, open access journal, published every six months (June-December). Science & Philosophy aims to publish original research articles but also short communication and review papers of general significance that are written clearly and well organized.

The journal is a source of information for professionals in a wide range of scientific discipline. Indeed, Science & Philosophy is a multi-disciplinary journal that covers fundamental and applied research in various areas related to Science, Epistemology and Philosophy, including Mathematics, Statistics, and Social Science.

Both English and Italian languages are accepted for publications. In the case of an article in Italian, it is necessary to provide the title, the abstract, and the keywords also in English.

The journal is oriented on the scientific analysis of social phenomena, both by developing research with quantitative and qualitative methods and by interpreting with philosophical criticism. Science & Philosophy has its focus in the study of social phenomena both macro and micro, using the interdisciplinary approach of the social sciences, from economics to law, from politics to sociology, from history to social statistics. The journal focuses on the analysis of both phenomena and trends, for designing local developments and systems ecology. Scientific analysis is supported by an ethic of sociability.

**Note on Peer-Review**
All manuscripts are subjected to a double-blind review process. The reviewers are selected from the editorial board, but they also can be external subjects. The journal's policies are described at: http://eiris.it/ojs/index.php/scienceandphilosophy/about/submissions#authorGuidelines



**Publisher**: **APAV** - Accademia Piceno-Aprutina dei Velati in Teramo
**Tax Code** 92036140678, **Vat Id. Number** 02184450688
**Registered Office Address**: via del Concilio, 24 - 65121 Pescara
**Operational Office**: via Chiarini, 191 - 65126 Pescara

Gardaphe Frederick - Queens College, City University of New York, U.S. - fred.gardaphe@qc.cuny.edu;

Gatto Romeno - Department of Mathematics, Computer Science and Economics, University of Basilicata, Potenza, Italy - romano.gatto@unibas.it;

Gavrila Mihaela - Department of Communication and Social Research - Sapienza University of Rome, Italy - mihaela.gavrila@uniroma1.it;

Gerla Giangiacomo - Department of Mathematics / DIPMAT, Salerno University, Salerno, Italy - gerla@unisa.it;

Hošková - Mayerová Šárka - Department of Mathematics and Physics, University of Defence, Brno, Czech Republic - sarka.mayerova@seznam.cz;

Iacono Mauro - Department of Mathematics and Physics, University of Campania Luigi Vanvitelli, Caserta, Italy - mauro.iacono@unicampania.it;

Innamorati Stefano - Department of Industrial and Information Engineering and Economics, L'Aquila University, L'Aquila, Italy - stefano.innamorati@univaq.it;

Ispas Cristina - Universitatea "Eftimie Murgu" Reşiţa, Romenia - c.ispas@uem.ro;

Madureira Livia - Universidade de Tras - os - Montes e Alto Douro (UTAD), Portugal - lmadurei@utad.pt;

Malizia Pierfranco - Department of Economics, Politics and Modern Languages of L.U.M.S.A., Rome, Italy - pfmalizia@yahoo.it;

Marconi Domenico - Faculty of Biosciences and Agro - Food and Environmental Technologies, Teramo University, Italy - dmarconi@unite.it;

Markovic Ljiljana - Faculty of Philology, Department of Oriental Studies, University of Belgrade, Belgrade, Serbia - liliana.markovic@gmail.com;

Marradi Alberto - University of Florence, Florence, Italy - alkmar@libero.it;

Mascella Raffaele, Faculty of Communication Sciences, University of Teramo, Italy

Migliorato Renato - Department of Mathematics, University of Messina, Messina, Italy - renato.migliorato@gmail.com;

Montesperelli Paolo - Department of Communication and Social Research, Sapienza University of Rome, Rome, Italy - paolo.montesperelli@uniroma1.it;

Nicotra Luca - Cultural Association of Art and Science, Rome, Italy - luca.nicotra1949@gmail.com

Palladino Nicla - University of Perugia, Perugia, Italy - nicla.palladino@unipg.it;

Paone Fiorella - Department of Business Administration, University G. d'Annunzio of Chieti - Pescara, Italy - fiorella.paone@gmail.com;

Petrovic Mina - Department of Sociology, Faculty of Philosophy, University of Belgrade, Serbia - mipetrov@ssb.rs;

Porreca Annamaria, Department of Economics, University G. d'Annunzio of Chieti-Pescara, Pescara, Italy - annamaria.porreca@unich.it;

Rosati Norton Delfico Mauro - Rosati Law Firm & Partners - International LEGAL Advisors - info.btsconsultant@gmail.com;

Rotondo Paolo - Mathesis, Italyn Society of Mathematical and Physical Sciences, Italy - paolo_rotondo@libero.it;

Ruggiero Christian - Sapienza University, Rome, Italy - christian.ruggiero@uniroma1.it;

Savarese Elisa - I.T.I. Renato Elia, Castellammare di Stabia, Italy - elisa - elsava2@gmail;

Participation in the editorial board is renewed annually on the basis of the actual contribution, participation in the journal, and direction strategies defined by the chief editors. No communication is expected for the annual changes to the content of the board.

**Science & Philosophy** Volume 10(2), 2022

# On the symmetries of electrodynamic interactions

Hernán Gustavo Solari* and Mario Alberto Natiello†

**Abstract**

The development of relational electromagnetism after Gauss appears to stop around 1870. Maxwell recognised relational electromagnetism as mathematically equivalent to his own formulae and called for an explanation of why so different conceptions have such a large part in common. We reconstruct relational electromagnetism guided by the No Arbitrariness Principle. Lorenz' idea of electromagnetic waves, together with the "least action principle" proposed by Lorentz are enough to derive Maxwell's equations, the continuity equation and the Lorentz' force. We show that there must be two more symmetries in electromagnetism: a descriptive one expressing source/detector relations, and another relating perceptions of the same source by detectors moving with different (constant) relative velocities. The Poincaré group relates perceived fields by different receivers and Lorentz boosts relate source/detector perceptions. We answer Maxwell's philosophical question showing how similar theories can be abduced using different inferred entities. Each form of abduction implies an interpretation and a facilitation of the theoretical construction.
**Keywords:** critical epistemology; rationalism; relational electromagnetism; Lorentz transformations; Doppler effect; [1]

*Departamento de Física, FCEN-UBA and IFIBA-CONICET; Pabellón I, Ciudad Universitaria (1428) - C.A.B.A - Argentina. email: solari@df.uba.ar Orcid: 0000-0003-4287-1878

†Centre for Mathematical Sciences, Lund University. Box 118, S 221 00 LUND, Sweden. email: mario.natiello@math.lth.se (corresponding author) Orcid: 0000-0002-9481-7454

[1]Received on August 4, 2022. Accepted on December 27, 2022. Published on December 31, 2022. doi: 10.23756/sp.v10i1.811. ISSN 2282-7757; eISSN 2282-7765. 

# 1 Introduction

The notion that science, and in particular physics, does not depend on philosophical or psychological factors is usually manifested by scientists and the society at large. However, this view confuses what science should be with how science is actually practised. Following Peirce we can say that research stops when doubt is appeased and a (temporary) belief is reached. The condition for the cessation of doubt might have psychological and philosophical components. During the late XIX century and the beginning of the XX century an abrupt change in this condition can be verified (Solari and Natiello, 2022a) finally leading to new physical understanding [2] and a new epistemology (Solari and Natiello, 2022b). There is a relation of precedence: psychological needs (such as the need for analogies or to incorporate learned habits) determine, in part, physical theories which in turn determine philosophy. Denying the existence of the first link we could claim that the Truth in physics forces upon us the acceptance of some epistemologies and the rejection of others. In contrast, for a critical philosophy such as Kant's (Kant, 1798) it is philosophy the science that surveils and, if necessary, corrects all other human activities. Thus, for critical philosophy the sequence must be: philosophy controls the sciences and the contributions by psychological needs of scientists have no place and must be eliminated.

The symmetries of electromagnetic interactions played a central role in the transformation underwent by physics, and with it by science, during that period. Expectations imported from Mechanics did not fit observations of electromagnetic phenomena, in particular the propagation of electromagnetic interactions and light. Two alternatives circulated around 1850, namely local propagation through some form of physical medium in space (the ether) against delayed action at a distance. The second alternative had faded away by the turn of the century, although it was never proved wrong. The introduction (and subsequent elimination) of the ether along with a second ingredient: the expectations posed by society on science reshaped the way physicists approached Nature. The progress of the industrial revolution expected science to be the support of technological development, a goal not necessarily identical to that of exploring Nature in order to understand it. The utilitarian view of science advanced at the beginning of the second industrial revolution in the Prussian empire proclaims its success some 60 years later. With it comes an a–critical epistemology that denies philosophy the right to examine the foundations of science (Beiser, 2014) as it is actually practised: the utilitarian, capitalist, science.

Is it the same physics resulting from both forms of construction? For the case of Mechanics most results coincide (Solari and Natiello, 2018), while founda-

[2]Meaning the acceptance of a theory by a community

tional issues regarding the concept of inertial systems drastically differ (Solari and Natiello, 2021).

The ether failed to provide a sound solution to these problems and Special Relativity was advanced in 1905, being today the accepted explanatory framework. However, already in 1867 Ludwig Lorenz suggested an ether-free description of electromagnetism. While the interpretation of electrodynamics in terms of special relativity must be rejected as an acceptable theory under a rational construction (Solari and Natiello, 2022b), the success obtained by applying this theory to observable problems and the absence of an alternative (at least) equally successful, consilient and coherent (Whewell, 1840, 1858b) prevented the criticism of its foundations.

The combination of motion and coordinate description of electromagnetic phenomena has several aspects. At least three elements are usually present: Observer, Source (Emitter, Primary circuit) and Receiver (Detector, Secondary circuit). However, not all motions are equally relevant. The No Arbitrariness Principle (NAP)(Solari and Natiello, 2018) (elaborating on the idea that no knowledge about nature depends on arbitrary decisions) suggests that the only motion that actually can influence results is that between Source and Receiver. Moreover, in a relational description, there is no other motion involved and the Observer is either absent or sorted out through a group of symmetry transformations between equivalent choices.

In this work we illustrate how these setups can be fully handled. We assemble Electromagnetic theory in terms of classical epistemology; hopefully achieving a better matching with experiments than current theories and higher "consilience" (Thagard, 1978) (see also (Whewell, 1840, p. XXXIX, Aphorism XIV)). First, we derive the set of equations of electromagnetism combining Lorenz' approach with an ether-free version of Lorentz' action integral, unifying and surpassing ideas that have not been fully investigated so far. Further, we relate the electromagnetic description for the case where source and receiver are at relative rest with the corresponding description in a situation of relative motion, showing also how potentially controversial concepts such as the "velocity of light" $C = (\mu_0\epsilon_0)^{-\frac{1}{2}}$ in different states of relative motion fit in this nineteenth century framework. From the concept of reciprocal action (which is in the philosophical basis of Newton's mechanics) we examine the arbitrariness that has to be removed in Electromagnetic theory and then, the symmetry groups that must be involved a-priori. This rational [3]theory of Electromagnetism does not require any change in space-time or epistemology.

[3]The rational epistemology was presented by William Whewell (Whewell, 1840, 1858b,a) and further developed by Charles Peirce (Peirce, 1994) and its fundaments where available by 1858 before the seminal works of Maxwell (Maxwell, 1865) and Lorenz (Lorenz, 1867).

We try to develop a method that allows all philosophers to grasp its contents, thus rescuing physics from elitism. If science is to help us to come into harmony with the universe, beginning with Planet Earth, a new perspective of exemplary science must be reached, one aiming at understanding and empathising with all living forms. Thus, the aim of this work is political, but yet it is philosophical as well as technical. If successful in our task (as we believe we are), we can claim that there is no need to abandon the goal of understanding nature and also that the utilitarian science aimed at "dominating nature" (a prediction technique whose value is given by predictive success), needs to be left behind if harmony in Planet Earth is our goal.

# 2 On symmetries

Physics sustains the idea that there is a world that reaches us through the senses and is independent of the observer: the sensed-real. Although every particular observation may depend on the observer, the collection of observations points towards a common idea that we call reality, or *the real*. Thus, the relation between the sensed-real and reality (the idealisation) plays a fundamental role. This starting point has been called "The fundamental antithesis of philosophy" (Whewell, 1858a, Ch. I). Going from the sensed-real to the real we must separate what belongs to reality from its circumstances that result in particularities, which quite often are the consequence of arbitrary decisions. Thus, we reserve the name of arbitrariness for the observational and descriptive decisions that we have to make when associating an ideal relation with an observable relation.

It would be desirable to present physical laws in pure abstract form, without any arbitrary element, but it would be desirable as well, for physical laws to be as accessible as possible to the mind. Since abstraction imposes difficulties in grasping the meaning of such laws, there is a trade-off that must be worked out between the two desires. This trade-off results in the introduction of some (usually small) set of arbitrary elements in the description, under the requirement that such arbitrary elements could be eventually suppressed from the presentation or, what is the same, that a change in the choice of arbitrary elements results in an equivalent presentation. These ideas lead immediately to the existence of a group of transformations relating different choices of arbitrary elements. The group structure is the result of the composition law of the transformation between presentations of the laws under different arbitrary decisions. This is the central idea under the "No arbitrariness principle" (NAP) (Solari and Natiello, 2018).

The introduction of an observer brings about the possibility of attaching to it a Cartesian space for the description of the real and at the same time it introduces the symmetries of the space (the arbitrary element).

Moving directly into electromagnetism, we observe that all its fundamental experiments reflect the influence of electromagnetic phenomena associated to a pair of bodies (one of them labelled primary circuit, source, emitter, etc., and the other secondary circuit, receiver, detector). In the same form that space is not a possible subject of experimental detection but spatial relations can be measured, electromagnetic fields can only be detected by their effects on measuring devices, i.e., detectors. If the action of a source on a receiver can be addressed with controlled degrees of influence from the rest of the universe, in the limit of no influence, the idealised law describing the universe of such relations must depend only on the relative position and motion of source and receiver. Such notions can be found all over the foundational work of Faraday (Faraday, 1839, 1844, 1855) and Maxwell (Maxwell, 1873).

Electromagnetic phenomena imply the motion of electricity (whatever electricity is, as Maxwell often said) and then, since what changes the motion of bodies has been called *forces*, we can associate forces with the action of an electromagnetic (EM) body onto another EM body. Actually, this use entails a generalisation of the concept of force, since Newtonian forces change the motional status of macroscopic bodies while microscopic (quantum) objects, such as electrons involved in conduction currents, are not what classical mechanics had in mind when Newton developed its laws. Moreover, if we envisage EM-forces as Lorentz did, by adopting Weber's view of electrical atoms (Lorentz, 1892), such forces must be identically described by observers whose motions relate by Galilean coordinate transformations, and furthermore reciprocal action must be expressed as a symmetry in some privileged systems we call "inertial frames" (Thomson, 1884). For example, the symmetry inherent to Newton's third law is expressed as the equation $F_{12} + F_{21} = 0$ being invariant in front of Galilean changes of coordinates (where $F_{ij}$ is the force on body $j$ originated in the interaction with body $i$). Yet, we know at least since Poincaré (1900) (see (Solari and Natiello, 2018) as well) that Newton's "action and reaction law" is not compatible with delayed action at distance. As far as we know, the form this symmetry takes in EM has not been shown so far. We will display its effects in the present work.

When EM theory is moved from its original setting as an *interaction theory* into a *field theory*, some symmetry is broken since there are no longer two EM-bodies in reciprocal action but we are thereafter concerned with only one of them, most frequently the *source*. This presentation of EM may be called the *S-field*. With equivalent arbitrariness we could shift the focus to the receiver and consider an *R-field* description. Both descriptions refer to the same EM phenomena and are therefore related.

When the S-field, the field produced by the source, is perceived by the source itself or by any extended EM-body not moving with respect to the source, we call it *S-by-S-field*. When considering the same S-field as it is perceived by the receiver, we have the *S-by-R-field* description, see Figure 1 (see Figure 2 for the corresponding *R-field* description). The operation performed on the description of the phenomenon is to identify one body or the other with an extended EM-body in the reference frame of the observer. As both approaches describe the same action, a transformation, possibly dependent on the relative velocity between the EM-bodies, must relate their expressions.


Figure 1: Field of the source (a) as seen by the receiver (S-by-R-field) and (b) as seen by the source (S-by-S-field). Source to the left of each image. In blue: the device at rest with the observer.

For the case of multiple receivers we may want to consider the relation among the different S-by-R$_i$-field descriptions of each receiver. To connect $R_1$ with $R_2$ corresponds to the composition of the transformations between each receiver and the source, namely $R_1 \to S$ and (the inverse of) $R_2 \to S$. The composition of transformations yields a transformation between receivers, that will depend on the relative velocities of $R_1$ and $R_2$ with respect to the source. However, receiver-receiver transformations relate objects of equivalent character, they are automorphisms and must form a group as well.


Figure 2: Field of the receiver (a) as seen by the receiver (R-by-R-field) and (b) as seen by the source (R-by-S-field). Source to the left of each image. In blue: the device at rest with the observer.

The perceived fact that electromagnetic disturbances require some time to propagate between source and receiver is acknowledged by all existing theoretical frameworks of EM. To describe this fact, the concept of *delayed action at a distance* was advanced in an organised form by the Danish scientist Ludwig Lorenz (Lorenz, 1867) after preliminary attempts (Betti, 1867; Riemann, 1867; Neumann, 1868) from the Göttingen school originated by ideas of Gauss (bd.5 p. 627-629, Gauss, 1870).

Returning to relative motion, it must be noticed that even in the case where source and receiver are in constant relative motion, the transformation between the S-by-S-field and S-by-R-field will not be an inertial transformation (i.e., a Galilean coordinate change). Galilean transformations correspond to descriptive transformations that are not concerned with the observable relative motion of the bodies. The relative motion of source and receiver is a measurable part of the physics involved and not an arbitrariness (it is there independently of the observer). Consider the following experiment: a source is producing a signal sharply peaked around a given frequency, $\omega_0$ as perceived by a receiver not moving with respect to the source. A set of several, identically built and calibrated receivers are put in motion at various velocities, $v_i$, with respect to the source, see Figure3. How is the signal perceived by each receiver? Which is the perceived characteristic frequency $\omega_i$ ? Which is the relation between the signals registered by the various receivers?

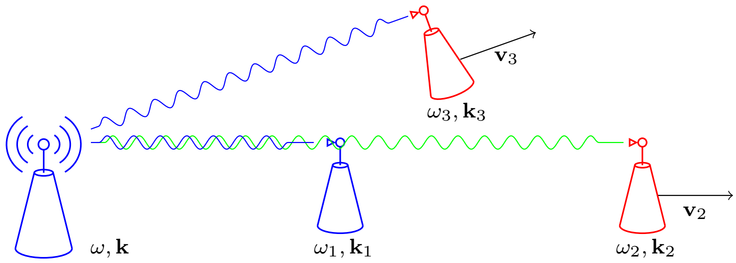



Figure 3: Sources and receivers. Blue receiver at rest relative to source, red receivers in relative motion with respect to the source.

# 3 Relational Electrodynamic Background

## 3.1 Interaction-based relational formulation.

In the presence of electromagnetic interactions, the observable effects of the interaction can be interpreted as the result of the action of the *Lorentz force* (Lorentz, 1892; Natiello and Solari, 2021) over the electrified particles that constitute matter.

The origins of the Lorentz force can be traced back to Maxwell and what he called the *Electromotive intensity* ([598], Maxwell, 1873). Similarly, Lorentz referred to Maxwell's electrokinetic and potential energies [630,631] and [634,635], Maxwell, 1873, combining them in an action integral and the principle of least action. These presentations take support in Maxwell equations,

$$
\begin{aligned}
B &= \nabla \times A & (1)\\
E &= -\frac{\partial A}{\partial t} - \nabla V & (2)\\
\epsilon_0 \nabla \cdot E &= \rho & (3)\\
\mu_0 j + \frac{1}{C^2}\frac{\partial E}{\partial t} &= \nabla \times B & (4)
\end{aligned}
$$

although their derivations some way or the other involved the ether in the argumentation: Maxwell when considering the "total current" of eq.(4) and Lorentz in the variational principle.

Ludwig Lorenz avoided to introduce the ether by acknowledging that light was a form of EM interaction and it corresponded with a transversal wave (Lorenz, 1861, 1863), later introducing retarded electromagnetic potentials (Lorenz, 1867) inspired in Franz Neumann(Neumann, 1846)[4],

$$(A, \frac{V}{C})(x,t) = \frac{\mu_0}{4\pi} \int \left( \frac{(j, \rho C)(y, t - \frac{1}{C}|x-y|)}{|x-y|} \right) d^3y, \tag{5}$$

as an expression based upon these observations, and also on Neumann's results and Kirchhoff results regarding EM waves in conductors which make ample use of the continuity equation, $\frac{\partial \rho}{\partial t} + \nabla \cdot j = 0$. The displacement equation (4) can be derived from Equation (5) and the continuity equation. It is everywhere assumed that the current-charge vanishes rapidly enough at infinity (so that the partial integrations usually present in EM theory can actually be performed).

In terms of differential equations, Eq. (1) and (2) are definitions of the magnetic and electric fields and the main constitutive equation reads

$$\Box (A, \frac{1}{C}V) = -\mu_0 (j, C\rho). \tag{6}$$

where $\Box = \Delta - \frac{1}{C^2}\frac{\partial^2}{\partial t^2}$ is the D'Alembert operator. This equation is satisfied also by:

$$(\widetilde{A}, \frac{\widetilde{V}}{C})(x,t) = \frac{\mu_0}{4\pi} \int \left( \frac{(j, \rho C)(y, t + \frac{1}{C}|x-y|)}{|x-y|} \right) d^3y. \tag{7}$$

The potentials $A, V$ describe the relation between current-density $j$ or a charge-density $\rho$ with their electromagnetic effect. The standard interpretation is that

[4]Maxwell's results have the same starting point in Neumann's work ([542], Maxwell, 1873)

$(j, C\rho)$ are the source (the primary circuit) of the EM action while the potentials are intermediate fields that indicate their action over the secondary circuit, corresponding to delayed action; this is, $(A, \frac{V}{C})$ are source fields, S-fields. A different association is possible for $(j, C\rho)$; they can be interpreted as those corresponding to the secondary circuit and in such case $(\widetilde{A}, \frac{\widetilde{V}}{C})$ are the R-fields that sense an EM perturbation away from the receiver and express its effect later in it, this is, they are advanced fields.

When relevant, we use the indices $1(2)$ for the source (receiver). It is possible to perform a derivation of the Lorentz force (Natiello and Solari, 2021) from the Principle of Least Action supported in Maxwell's energy considerations following Lorentz but using mathematical deduction at the few situations where Lorentz used arguments corresponding to the ether in (Lorentz, 1892). Let $\bar{x}(t)$ denote the distance between a reference point in the source and a reference point in the receiver. We will consider situations where source and receiver move as rigid bodies in relative motion (but not in relative rotation) as Lorentz did.

In what follows, $z$ denotes a "local" coordinate on body 2. We consider, following Lorentz, a collection of virtual displacements parametrised by time $\delta\bar{x}(t)$ [5]. The variation of charge and current densities $\rho_2(z,t), j_2(z,t)$ on the receiver can be expressed in the coordinates of eq.(5) as:

$$\begin{aligned} \delta\rho_2(x,t) &= (-\delta\bar{x}(t)\cdot\nabla)\,\rho_2(x,t) \\ \delta j_2(x,t) &= (-\delta\bar{x}(t)\cdot\nabla)\,j_2(x,t) + \delta\dot{\bar{x}}\rho_2(x,t) \end{aligned} \tag{8}$$

The latter relates the local expression of charge and current densities in the secondary circuit and the same physical object in terms of the coordinates associated to the primary circuit.

Maxwell considers the electrokinetic and potential energies, which Lorentz further combines in the action integral

$$\mathcal{A} = \frac{1}{2}\int dt \int (A_1(x,t)\cdot j_2(x,t) - \rho_2(x,t)\,V_1(x,t))\, d^3x \tag{9}$$

that here represents the interaction energy between a source or primary circuit labelled 1 and a receiver or secondary circuit labelled 2. The relation 6 is satisfied for fields and current-charge corresponding to the same index. The action integral in the present form corresponds to an S-by-S-field representation, namely that the

[5] As in the Lagrangian formulation, the collection of virtual displacements is differentiable, i.e., $\dot{\bar{x}}$ exists, and the variation is zero in the time extremes. Virtual displacements are not the same as time-dependent perturbations of the position, for the latter have other effects apart from the change of relative distances. Virtual displacements are closer to changes of initial conditions than to perturbations. In particular, during a virtual displacement, there is no wave progression.

fields of the source are evaluated at the position of the receiver in the coordinates $x$ of the source and time $t$.

We state the result as a theorem:

**Theorem 3.1.** *((Natiello and Solari, 2021)) Assuming that all of $|B|^2$,$|E|^2$, $A$, $j$, $V$, $\rho$ decrease faster than $\frac{1}{r^2}$ at infinity, assuming the action is given by eq. 9 and given the validity of the continuity equation $\frac{\partial \rho}{\partial t} + \nabla \cdot j = 0$, the electromagnetic force*

$$F_{em} = \int d^3x \left[ j_2(x,t) \times B_1(x,t) + \rho_2(x,t)\, E_1(x,t) \right]$$

*on the probe can be deduced from Hamilton's principle of minimal action ($\delta_{\bar{x}(t)}\mathcal{A} = 0$) using a virtual displacement $\delta_{\bar{x}}$ of the probe (which we indicate with subindex 2), eq.(8) with respect to the primary circuit producing the fields (subindex 1).*

## 3.2 Wave equation for the potentials

The wave equation for the potentials can be deduced from Equations (1-4).

**Lemma 3.1.** $A(x,t) = \frac{\mu_0}{4\pi} \int_U \left( \frac{j(y, t - \frac{1}{C}|x-y|)}{|x-y|} \right) d^3y \Rightarrow \Box A = -\mu_0 j$, *and similarly for* $\epsilon_0 \Box V = -\rho$, *where* $\Box \equiv \Delta - \frac{1}{C^2}\frac{\partial^2}{\partial t^2}$.

For a proof, see Appendix 3.1. Note that this result describes a property of eq.(5), independently of whether $A, V, j, \rho$ are the electromagnetic vector potential and current, etc., or not. We prove now that the result holds for the electromagnetic $A, V, j, \rho$, via a variation of the electromagnetic action 9:

**Theorem 3.2.** *Let* $(\mathsf{A}, \mathsf{V})$ *be the known values of the electromagnetic potentials in a piece of matter supported on a region of space with characteristic function $\chi$. Then, assuming that all of $|B|^2, |E|^2, A, j, V, \rho$ decrease faster than $\frac{1}{r^2}$ at infinity, Hamilton's principle of least action (Ch 3, 13 A p. 59, Arnold, 1989), $\delta\mathcal{A} = 0$, subject to the constraints given by* $(\mathsf{A}, \mathsf{V})$ *implies the relations*

$$\begin{aligned} \frac{1}{\mu_0}\nabla \times B - \epsilon_0 \frac{\partial E}{\partial t} &= \mu_0 j \\ \epsilon_0 \nabla \cdot E &= -\frac{\rho}{\epsilon_0}. \\ \nabla \cdot j + \frac{\partial \rho}{\partial t} &= 0 \end{aligned}$$

**Corollary 3.1.** *In the special case when the relation* $\nabla \cdot A + \frac{1}{C^2}\frac{\partial V}{\partial t} = 0$ *(the "Lorenz gauge") is satisfied, the manifestation of the potentials outside matter obeys the wave equation, eq.(6).*

We develop the proof in Appendix 3.2.

The theorem deserves to be named Lorenz-Lorentz theorem since in Lorenz conception light was associated to the EM activity inside matter (Lorenz, 1867) and Lorentz proposed the expression for the action based on Maxwell's energy considerations.

Recasting the potentials of eq.(5) as the convolution of charge and currents with the *Lorenz kernel* hereby defined:

$$K(x-y, s-r) = \frac{1}{|y-x|}\delta(s - r - \frac{1}{C}|y-x|),$$

namely

$$(A_1, \frac{V_1}{C})(x,s) = \frac{\mu_0}{4\pi}\int \left[\int_{-\infty}^{s} K(x-y, s-r)(j_1, C\rho_1)(y,r)\right] d^3y\, dr \qquad (10)$$

a fundamental symmetry between potentials and wave operators is expressed in the following

**Lemma 3.2.** *The action of the kernel* $K(x-y, s-r)$ *and the differential operator* $\Box$ *are reciprocally inverse of each other.*

*Proof.* We discuss the proof using $A$ to fix ideas, and write eq.(10) in shorthand as $A = \frac{\mu_0}{4\pi}K * j$ (where the star stands for convolution). Composition with $\Box$ gives:

$$\begin{aligned} \Box A &= \frac{\mu_0}{4\pi}\Box K * j = -\mu_0 j \\ K * \Box A &= -\mu_0 K * j = -4\pi A. \end{aligned}$$

Hence, in their respective domain of definition $\Box K = -4\pi Id$ (convolution identity) and $K * \Box = -4\pi Id$ (operator identity). $\Box$

## 3.3 Source/receiver symmetry of the action

Since the action (9) plays a fundamental role in this relational presentation we should devote some lines to consider its symmetries.

We first write the action in terms of definite integrals and the kernel $K(x-y, s-r)$

$$\mathcal{A} = \frac{1}{2}\frac{\mu_0}{4\pi}\int_{t_0}^{t} ds \int_{t_0}^{t} dr \iint K(x-y, s-r)\left(j_1 \cdot j_2 - C^2\rho_1\rho_2\right) d^3x\, d^3y \quad (11)$$

The form of the action in eq.(11) is almost symmetric in terms of exchanging primary and secondary circuits. Interchanging primary and secondary circuit, and $(x, s) \longleftrightarrow (y, r)$ the kernel changes into

$$K(x - y, s - r) = \frac{1}{|y - x|}\delta(s - r + \frac{1}{C}|y - x|) \tag{12}$$

Thus, the action considered is always the action of the primary circuit over the secondary circuit which can be written in two forms. In one of them, the S-field (the standard form), EM changes are propagated with delay by the potentials (and their derivatives, the EM-fields) at distances away from the source. The symmetry-related form, the R-field, associates an advanced field with the receiver. In this form, the field can be seen as a sensor that will carry disturbances to the receiver that will display changes at a later time.

The symmetry of the action has the immediate consequence that all lemmas and theorems of subsections (3.1) and (3.2) have an equivalent form under this symmetry operation. In particular, there is Lorentz-force where the S-fields, R-currents and R-charges are exchanged by R-fields, S-currents and S-charges. This relation is what corresponds to the action and reaction law for actions that propagate instantaneously, since in the limit $C \to \infty$ the S-field and the R-field of a given body/device coincide.

## 3.4 Detection/perception in relative motion

Let us consider the potentials $A, V$ originated in a source with current-charge $J = (j, C\rho)$ measured at (rest relative to) the source (with coordinate $y$). We consider further a detector extending over a variable $x$ with reference to a distinguished point in it. In the case of source and detector at relative rest, we write

$$(A, \frac{V}{C})(x, t) = \frac{\mu_0}{4\pi}\int d^3y \int ds \left(\frac{\delta((t - s) - \frac{1}{C}|x - y|)}{|x - y|}\right) J(y, s) \tag{13}$$

$$= \frac{\mu_0}{4\pi}\int d^3z \left(\frac{J(x - z, t - \frac{1}{C}|z|)}{|z|}\right) \tag{14}$$

These equations are formulated under the following premises: Coordinates $y$ and $x$ are described from the same spatial reference system $S$, whatever it is, and hence at a given time $t$, $x - y$ and in particular $|x - y|$ are objective invariant quantities. Moreover, since source and detector are in relative rest, these quantities are independent of $t$. In the present conception of electromagnetism there is another objective invariant quantity of relevance, namely the electromagnetic delay $\Delta_0 = t - s$. The index 0 highlights the situation of relative rest between source and detector. It is the state of point $y$ on the source at the previous time $s$, where $C(t-s) = |x-y|$ what connects with point $x$ of the detector at time $t$. Finally, the second row displays the change of variables $z = x - y$.

In order to address detection in relative motion we advance the following

**Conjecture 3.1.** *A detector recording solely electromagnetic information (e.g. an electromagnetic wave) cannot determine its relative velocity with respect to the source (assumed constant).*

Consequently, let us postulate that a detector in relative motion with velocity $v$ with respect to the source perceives an EM wave which cannot be distinguished from the one originating in some current-charge *at relative rest*. We would like to show something like:

$$(A, \frac{V}{C})_v(x,t) \quad = \quad \frac{\mu_0}{4\pi}\int d^3y \left(\frac{1}{|x-y|}\right) J_v(y, t-\Delta) \tag{15}$$

with $\Delta = \frac{1}{C}|x-y|$.

In this new situation we still have one reference frame $S$ to describe both source and detector. Again, $z = x - y$ is an objective quantity, only that now two differences arise: (a) $x - y$ depends on $t$ because of the relative motion and (b) the electromagnetic delay may be modified in order to take into account the relative motion. Throughout this discussion, $t$ is the (present) time when the electromagnetic interaction is detected, $(x-y)$ indicates the relative position of (points of) detector and source at time $t$, $\Delta_v = (t-s)_v$ is the electromagnetic delay and $(x-y)_v$ is the corresponding relative position at time $s$ when the electrical disturbance in the source took place, and the index $v \in \mathbb{R}^3$ indicates a situation of relative motion between source and detector. The index $v$ will be some function of the relative velocity $u$ between source and detector to be determined in what follows. Moreover, $(x-y)_v$ and $\Delta_v$ are objective and invariant quantities, independent of the choice of reference frame.

We intend to find the correspondence between disturbances in the primary circuit and actions on the secondary system. We begin by considering an infinitesimal velocity $\delta v$, with $\frac{d\delta v}{dt} = 0$. In this case we have

**Definition 3.1.** ***(Differential delayed interaction condition)*** *In the presence of relative motion with infinitesimal velocity $\delta v$, a disturbance originated at point $y$ and time $t - \Delta_{\delta v}$ produces an electromagnetic action at $(x,t)$, where*

$$C\Delta_{\delta v} \quad = \quad |x - y - \Delta_{\delta v}\delta v|.$$

For $\delta v = 0$ the condition reduces to $C\Delta_0 = |x-y|$, corresponding to Lorenz' potentials, eqs.(5) and (10)[6]. Note that $C$ enters in both expressions since we

[6]Letting $s = t - \Delta_{\delta V}$ we may read the definition as a consequence of: $(x-y)(s) = (x-y)(t) - (t-s)\delta v$.

postulate that the detector in relative motion registers an electromagnetic signal *as if the source were at relative rest*. This definition leads to the following

**Lemma 3.3.** *Let $(x-y)_v$ be the separation of source and detector at time $s$ when the electrical disturbance at the source took place in a situation of relative motion labelled by $v \in \mathbb{R}^3$ and $\Delta_v$ the corresponding electromagnetic delay, while $(x-y)_0, \Delta_0$ are the corresponding quantities for source and detector at relative rest. Then, for each $v$ the* ***delayed interaction condition*** *satisfies*

$$\begin{pmatrix} (x-y) \\ C\Delta \end{pmatrix}_v = \exp\left(-\begin{pmatrix} \mathbf{0} & \frac{v}{C} \\ \frac{v}{C}^T & 0 \end{pmatrix}\right)\begin{pmatrix} (x-y) \\ C\Delta \end{pmatrix}_0$$

*Proof.* To lowest order in $\delta v$ the difference in $\Delta$'s is:

$$\begin{aligned} C\left(\Delta_{\delta v} - \Delta_0\right) &= \sqrt{|x-y|^2 - 2\,(x-y)\cdot\delta v\Delta_{\delta v} + |\delta v|^2\Delta_{\delta v}^2} - |x-y| \\ &= -\frac{(x-y)}{|x-y|}\cdot\delta v\Delta_0 + O(\delta v^2) = -\,(x-y)\cdot\frac{\delta v}{C} + O(\delta v^2) \end{aligned}$$

In this limiting case the condition reads

$$\begin{aligned} \begin{pmatrix} (x-y)_{\delta v} \\ C\Delta_{\delta v} \end{pmatrix} &= \begin{pmatrix} (x-y) - \delta v\Delta_0 \\ C\Delta_0 - (x-y)\cdot\frac{\delta v}{C} \end{pmatrix} \qquad (16) \\ &= \left[\begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} - \begin{pmatrix} 0 & \frac{\delta v}{C} \\ (\frac{\delta v}{C})^T & 0 \end{pmatrix}\right]\begin{pmatrix} (x-y) \\ C\Delta_0 \end{pmatrix}. \end{aligned}$$

In other words, there exists an infinitesimal transformation on $\mathbb{R}^{3+1}$ connecting the condition for $v = 0$ with that for $\delta v$. By the Trotter product formula we obtain Lie's result for finite $v$ as a repeated composition of infinitesimal shifts,

$$\begin{aligned} TL(-v) &\equiv \exp\left(-\begin{pmatrix} \mathbf{0} & \frac{v}{C} \\ \frac{v}{C}^T & 0 \end{pmatrix}\right) \\ &= \lim_{n\to\infty}\left[\begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} - \frac{1}{n}\begin{pmatrix} \mathbf{0} & \frac{v}{C} \\ \frac{v}{C}^T & 0 \end{pmatrix}\right]^n \qquad (17) \end{aligned}$$

thus proving the statement. □

**Remark 3.1.** *Explicit formulae for the Lorentz transformations are shown in the Appendix B. The more familiar form $L(u)$ of the transformation is displayed in*

$$\begin{pmatrix} z_u \\ C\Delta_u \end{pmatrix} = L(u)\begin{pmatrix} z \\ C\Delta_0 \end{pmatrix} = \begin{pmatrix} z + (\gamma - 1)\hat{u}\,(\hat{u}\cdot z) + \gamma \frac{u}{C} C\Delta_0 \\ \gamma \left( C\Delta_0 + \frac{u \cdot z}{C} \right) \end{pmatrix}, \quad (18)$$

*where $u = C\hat{v} \tanh \left| \frac{v}{C} \right|$ and we use the shorthand $x - y = z$. There is a 1-to-1 correspondence in Lemma 3.3, between the two presentations of the Lorentz transformations, namely $TL(-v) \equiv L(-u)$. Hence, we will use only $u$ in the sequel. $u$ is interpreted as the relative velocity between source and detector.*
*The basis for the interpretation of $u$ as the relative velocity is as follows. Consider the vector space $\mathbb{R}^{3+1} \equiv \mathbb{R}^3 \times \mathbb{R}$ associated to relative positions and relative time. A Lorentz transformation (LT), eq.(18), as well as a Galilean transformation GT,*

$$\begin{pmatrix} Z' \\ T' \end{pmatrix} = \begin{pmatrix} \mathbf{1} & u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Z \\ T \end{pmatrix}$$

can be regarded as endomorphisms of $\mathbb{R}^{3+1}$ mapping a situation at relative rest onto a situation of relative motion. While the velocity $u$ in the GT has a mechanical origin, the parameter $u$ in LT is an abstract parameter used to classify transformations and a point of contact with the underlying physical problem is required to furnish a physical interpretation to the LT's. Considering lines on $\mathbb{R}^{3+1}$ associated to a fixed relative position, $Z$ and different time-intervals, we obtain for the Galilean transformation the (physical) relative velocity $u = \frac{Z'(T_1) - Z'(T_0)}{T'(T_1) - T'(T_0)}$ while in the case of the Lorentz transformation we obtain

$$\frac{z'(\tau_1) - z'(\tau_0)}{\tau'(\tau_1) - \tau'(\tau_0)} = \frac{\gamma u(\tau_1 - \tau_0)}{\gamma(\tau_1 - \tau_0)} = u.$$

While the GT preserves times and as such can be viewed as a transformation in relational-space only, the LT preserves $|z|^2 - (C\tau)^2$ and, as a particular case, the condition of being in electromagnetic contact, $|z|^2 - (C\tau)^2 = 0$. We may associate the same relational velocity to both GT and LT.

Eq.(17) displays the action of a Lorentz' boost (Gilmore, 1974) in the Lie algebra (rhs) and group (lhs). The generators of the Lorentz boosts plus the generators of the rotations constitute the basis of the Lie algebra which exponentiated gives the Poincaré-Lorentz group. While the spatial rotations form a subgroup of the Poincaré-Lorentz group, the Lorentz boosts do not. Any element of the Poincaré-Lorentz group can be written as a product: $P = L(u)R(\Omega)$ as well as $P = R(\Omega)L(u')$ being $\Omega$ a 3d-rotation and $u' = R(\Omega)u$ . These forms are known as left and right coset decompositions of the group (Hamermesh, 1962; Gilmore, 1974).

**Remark 3.2.** *By construction of the $LT's$,there is an upper limit for having electromagnetic contact amenable to be related with situations at relative rest. While there is no mechanical limit to relative velocity, the present theory describes electromagnetic interactions only for $|u| < C$ .*

**Remark 3.3.** *Eqs.16 and 18 for the detector and source points, $x, y$ which are in electromagnetic interaction at time t, display their relative position $(x-y)_u$ at the time $t - \Delta_u$ when the disturbance in the source took place. The ratio $\frac{|(x-y)_u|}{\Delta_u} = C$ is always satisfied by construction.*

Next, we note that the propagation kernel can be more properly written as

$$K = \begin{cases} 0, & (t-s) < 0 \\ \dfrac{\delta(t-s-\frac{1}{C}|x-y|)}{|x-y|}, & (t-s) \geq 0. \end{cases}$$

Hence, we have the following

**Lemma 3.4.** ***(Symmetric form of the propagation kernel)*** *Lorenz propagation kernel can be rewritten as*

$$K = \begin{cases} 0, & (t-s) < 0 \\ \dfrac{2}{C}\delta((t-s)^2 - \dfrac{1}{C^2}|x-y|^2), & (t-s) \geq 0. \end{cases} \tag{19}$$

*Proof.* In the distribution sense $K = \dfrac{2|x-y|}{C(t-s)+|x-y|}K$. By another distributional property, for any $g(s)$ such that $g(s_0) \neq 0$ it holds that $\dfrac{\delta(s-s_0)}{|g(s)|} = \delta(g(s)(s-s_0))$. In this case, $g(s) = t - s + \dfrac{1}{C}|x-y|$. Hence, we obtain the symmetric kernel expression of eq.(19). □

**Theorem 3.3.** *The Lorenz propagation kernel $K(x,t;y,s)$ has the following properties in relation to Lorentz transformations*

$$\begin{aligned} K(L_u(x,t); L_u(y,s)) &= K(x,t;y,s) \\ K(L_u(x,t); y,s) &= K(x,t;L_{-u}(y,s)) \\ \int d^3y\, ds\, [K(L_u(x,t);y,s)J(y,s)] &= \int d^3y\, ds\, [K(x,t;y,s)J(L_u(y,s))] \end{aligned}$$

*The last equation reads: the transformation of the potentials are the potentials associated to the transformations of the currents. We say then that the linear operator associated with $K$ commutes with the Lorentz transformation.*

*Proof.* It is straightforward to verify that the argument of the $\delta$-distribution in eq.(19) is invariant upon Lorentz transformations, namely that if $((x-y)_u, C(t-s)_u)$ satisfy eq.(18), then $(t-s)^2 - \frac{1}{C^2}|x-y|^2 = (t-s)_u^2 - \frac{1}{C^2}|(x-y)_u|^2$ and also $(t-s) \geq 0 \iff (t-s)_u \geq 0$. Thus,

$$K = \begin{cases} 0, & (t-s)_u < 0 \\ \frac{2}{C}\delta((t-s)_u^2 - \frac{1}{C^2}|\,(x-y)_u\,|^2), & (t-s)_u \geq 0. \end{cases}$$

is independent of $u$. Using the first property it follows that $K(L_u(x,t);y,s) = K(L_u(x,t;L_uL_{-u}(y,s)) = K(x,t;L_{-u}(y,s))$. The commutation relation is the result of integrating the kernel to produce a linear operator and changing integration variables $((y,s) \mapsto L_u(y',s')$. □

**Remark 3.4.** *The points that are in electromagnetic connection are characterised by* $(C(t-s))^2 - |x-y|^2 = 0$*. Calling* $\tau_u \equiv (t-s_0)_u$ *and* $\chi_u \equiv (x-y)_u$*, the interaction kernel is the convolution kernel of* $\delta(\tau_u^2 - (\chi_u/C)\hat{}2)$ *which can be split in two contributions, one for* $\tau_u \geq 0$ *and another for* $\tau_u \leq 0$*. But, if* $(0,0)$ *is influencing* $(\tau_0, \chi_0)$ *for* $\tau_0 \geq 0$*, it results that* $\tau_u > 0$ *(using that* $|u \cdot x/C^2| = \frac{|u\cdot\chi|}{|\chi||u|}\frac{|u||\chi|}{C^2} < \frac{|u||\chi|}{C^2}$*) hence the splitting is really in terms of* influencing*,* $\tau_u \geq 0$*, vs.* being influenced*,* $\tau_u \leq 0$*. This separates the sets in a form invariant with respect to* $u$*.*

### 3.4.1 Perceived fields and inferred currents-charges

Examining eq.(15), we note that it represents a convolution product with convolution kernel $\kappa(z,r)$, with $K(x,t;y,s) = \kappa(x-y, t-s)$ and that

$$(A, \frac{V}{C})_u = \frac{\mu_0}{4\pi}\kappa * J_u = \frac{\mu_0}{4\pi}J_u * \kappa$$

where the convolution is in time and space.

According to eq.(16), the arguments in the current are $(x-y, t-s)$, for $u = 0$. For $u \neq 0$ the points that are in electromagnetic relation according to Lemma 3.3 are $((x-y)_u, (t-s)_u)$, thus in $J_u * \kappa$, we propose

**Conjecture 3.2.** *The arguments of the effective current are* $((x-y)_u, (t-s)_u)$*, i.e.,* $J_u = L(-u)J(L(u)(x-y, t-s))$*, where* $J$ *is the current-charge measured by the source.*

At this point we must notice that there are three forms in which current-charge can be transformed to produce a new pair satisfying the continuity equation. Two of them are Galilean:

$$
\begin{aligned}
(j, C\rho)(x,t) &= (j - v\rho, C\rho)(x+vt, t) && (20)\\
(j, C\rho)(x,t) &= (j, C\rho - \frac{v}{C}\cdot j)(x, t + \frac{v\cdot x}{C^2}) && (21)\\
(j, C\rho)_u(x,t) &= L(-u)(j, C\rho)(L(u)(x,t)) && (22)
\end{aligned}
$$

In the third form, the leftmost $L$ acts on the charge-current $4D$-vector while the rightmost acts on the space-time coordinates.

If the form (20) is adopted, a theorem due to Maxwell ([602] Maxwell, 1873) shows that from the point of view of the receiver the transformation (21) must be applied to preserve the mechanical force but in such case the perceived potentials/fields are not waves. The empirical evidence has judged this view as not correct.

We propose to adopt eq.(22) as a definition of the inferred current. We insist at this point that the symmetry is not an a-posteriori observation of the formulae, but rather an a-priori demand of constructive reason as explained in (Solari and Natiello, 2018). The transformation of current-charge presents itself as a demand of reason to be later confronted with empirical results. That a charge density in motion can be perceived as a current is a belief firmly adopted since Weber's electrodynamic studies (Weber, 1846) and we are habituated to accept it, while that a neutral current in motion will be perceived as charge is not rooted in our beliefs in the same way, despite the fact that Maxwell's theorem already opened for that possibility.

**Remark 3.5.** *The symmetric form of $K$ is especially appealing when consider the backwards propagation kernel, as in the equation pairs (5)–(7) and (10)–(12). The backward propagation kernel is the result of inverting the time inequalities in 19.*

**Remark 3.6.** *Which is the meaning of a successive application of Lorentz' transformations to a current? The meaning we find apt is that if $J_u = L_{-u}J(L_u(x,t))$, then $J = L_u J_u(L_{-u}(x,t))$ (since Lorentz transformations have as inverse the transformation based on minus the velocity) and correspondingly $J_{u'} = L_{-u'}L_u J_u(L_{u'}L_{-u}(x,t))$. Since $L_{u'}L_{-u}$ is a general element of the Poincaré-Lorentz group, $L_{u'}L_{-u} = L_{u'\ominus u}R(u',u)$ with $u'\ominus u$ the coset addition of velocities, also known as Einstein's addition (Gilmore, 1974) and $R(u',u)$ a Wigner rotation[7] . Thus, the Poincaré-Lorentz group allows to convert between inferred*

[7]Wigner was not the first to study the group structure associated to Lorentz transformations or to mention the rotation. At least Silberstein (Silberstein, 1914, p. 167) in the published notes of his 1912-1913 course on Relativity at the University College, London, preceded Wigner, who acknowledged this precedence.

*currents or fields associated to different detectors in relative motion with respect to the same source. Notice that the relative velocity between both receptors is $u' - u$ but the correspondence of electromagnetic perceptions is not $L(u' - u)$ which might even not exist.*

Next, we explore the consequences of this proposal. Let us define operators acting on scalar or vector functions, $J$, of $(x, t)$ as

$$\begin{aligned} \widehat{K}\left[J\right](x,t) &\equiv \iint d^3z\, d\Delta\, K(z,\Delta) J(x-z, t-\Delta) \\ \widehat{L}_u\left[J\right] &\equiv J(L(u)(x,t)) \\ (\widehat{A}\circ\widehat{B})[J] &\equiv \widehat{A}\left[\widehat{B}[J]\right] \end{aligned} \tag{23}$$

The first line defines the action of the propagating kernel as a convolution, the second the action of a Lorentz transformation on the coordinates (recall that $u = Cv\tanh\left|\frac{v}{C}\right|$) while the third relation establishes notation.

**Lemma 3.5.** *According to the previous discussion, the perceived potentials read*

$$(A, \frac{V}{C})_u = \widehat{K}\left[J_u\right] \tag{24}$$

*In addition, we have the following identities*

$$\widehat{K}\left[J_u\right] = L(-u)\widehat{L}_u\left[\widehat{K}\left[J\right]\right]$$

*Proof.* Note that $L(-u)$ acts on the current-charge $J = (j, C\rho)$, while $\widehat{L}_u$ acts on the spatial/temporal arguments $x, Ct$. Eq.(24) is just eq.(15) rewritten through eq.(23). Recalling from eq.(22) that $J_u = L(-u)\widehat{L}_u\left[J\right]$ and from 3.3 that $\widehat{L}_u\circ\widehat{K} = \widehat{K}\circ\widehat{L}_u$ and finally that the matrix $L(-u)$ commutes with the scalar operator $K$ we obtain the result. □

### 3.4.2 The Doppler effect

The perception of wave frequencies in the case the waves are produced by a source in relative motion with respect to the receptor is known as *Doppler effect*. The EM Doppler effect plays a fundamental role in physics (Dingle, 1960; Mandelberg and Witten, 1962; Kaivola et al., 1985). The goal of this section is to show that the present theory provides an explanation for the experimental observations of the Doppler effect. To begin with, all Doppler experiments consist in comparing the waves perceived by a detector at rest with respect to the source against the perception of a detector moving at constant velocity (within acceptable experimental precision) relatively to the source.

In practice, the task is to obtain the Fourier transform of eq.(15). We will keep track of this process conceptually, and hence it is better to use the operator notation from Lemma 3.5. The Fourier transform of a function will be:

$$\mathcal{F}_{k,w}\left[\phi\right] = \frac{1}{(2\pi)^2}\iint d^3x\, dt\, \exp\left(-i(k\cdot x - wt)\right)\phi(x,t)$$

and is a function of $(k,w)$, where we have made an arbitrary choice in the election of the sign preceding $wt$ (that does not influence the conclusion). We will use the following known results:

$$\begin{aligned}
\mathcal{F}_{k,w}\left[\widehat{L}_u\phi\right] &= \mathcal{F}_{k',w'}\left[\phi\right] \text{ , with } (k',\frac{w'}{C}) = L(-u)(k,\frac{w}{C}) \\
\mathcal{F}_{k,w}\left[\widehat{K}\phi\right] &= \frac{1}{w^2 - C^2k^2}\mathcal{F}_{k,w}\left[\phi\right]
\end{aligned}$$

The first result is the immediate consequence of $L(u)$ being symmetric, while the second one can be obtained in various ways including direct integration. Applying these results to eq.((24)) we obtain

$$\begin{aligned}
\mathcal{F}_{k,w}\left[\widehat{K}[J_u]\right] &= \mathcal{F}_{k,w}\left[\widehat{K}\left[L(-u)\widehat{L}_u\left[J\right]\right]\right] \\
&= L(-u)\mathcal{F}_{k,w}\left[\widehat{L}_u\left[\widehat{K}\left[J\right]\right]\right] \\
&= L(-u)\mathcal{F}_{k',w'}\left[\widehat{K}\left[J\right]\right] \\
&= L(-u)\frac{1}{w'^2 - C^2k'^2}\mathcal{F}_{k',w'}\left[J\right] \\
&= L(-u)\frac{1}{w^2 - C^2k^2}\mathcal{F}_{k',w'}\left[J\right]
\end{aligned}$$

where $(k',\frac{w'}{C}) = L(-u)(k,\frac{w}{C})$. Thus, in terms of wave frequencies, the Fourier spectrum will have a peak at $w' = \gamma(u)(w - k\cdot u)$ associated with a source of frequency $w$. The primed quantities describe the characteristics of the wave as perceived by the detector while the unprimed refer to the source. When $k\cdot u = |k||u|$ the relative distance between source and detector increases, $w' < w$, and correspondingly the wavelength shifts towards higher values (red shift).

Hence, we have proved the following

**Theorem 3.4.** ***(Doppler effect)*** *A detector (observer) in relative motion with velocity $u$ with respect to an electromagnetic source emitting current-charge waves* *of wavelength and frequency* $(k, w)$ *detects electromagnetic waves of wavelength and frequency* $(k', \frac{w'}{C}) = L(-u)(k, \frac{w}{C})$.

**Remark 3.7.** *The symmetry* (25) *corresponds to expressing the action in terms of the inferred charge and currents by an observer. As such, it corresponds to a subjective view of EM.*

**Remark 3.8.** *The Galilean variation that allowed us to obtain the Lorentz force from the action, eq.*(8)*, indicates that the force experienced by the moving circuit takes the same form but the potentials to be used correspond to the perceived potentials of eq.*(15).

## 3.5 Mathematical presentation of the Lorentz transformation as a symmetry

Since Lorentz' transformations are well known in relation to electromagnetism, we consider their effect on the action and find their meaning in the present context.

Let $\mathcal{I}$ be the infinitesimal generator for the Lorentz transformation

$$\mathcal{I}_j = \left( Ct\frac{\partial}{\partial x_j} + \frac{x_j}{C}\frac{\partial}{\partial t} \right) \tag{25}$$

which together with the generators of the rotations

$$\mathcal{J}_i = \sum_{jk} \epsilon_{jki} \left( x_k \frac{\partial}{\partial x_i} - x_i \frac{\partial}{\partial x_k} \right)$$

(with $\epsilon_{jki}$ Kronecker's antisymmetric tensor) complete the Lie algebra of the Poincaré-Lorentz group (Gilmore, 1974).

**Theorem 3.5.** *The electromagnetic action* $\mathcal{A}$ (9) *transforms into an equivalent action* $\mathcal{A}'$ *when the infinitesimal transformations*

$$\hat{\delta} = \sum_i \left( \delta\theta_i \mathcal{J}_i + \delta v_i \mathcal{I}_i \right)$$

*operate on* $(j, C\rho)_{1,2}$ *simultaneously and* $\frac{d\delta v_i}{dt} = 0$.

*Proof.* The result follows from the observation that the kernel $K$ in (10) commutes with the six generators as a result of Theorem (3.3), and that, integrating by parts in space and time the action of $\hat{\delta}$ over $(j, C\rho)_2$ can be seen as an action over $(j, C\rho)_1$ preceded by a negative sign, and then, both actions compensate to first order. Thus, the infinitesimal action of any element of the Lie algebra acting on both subsystems (primary and secondary) corresponds to the identity. We have that

$$\mathcal{A} = \mathcal{A}' + \mathcal{F}(t)$$

with $\mathcal{F}(t)$ a functional of the potentials and currents evaluated at the time $t$. Since all variations are considered to be zero at the extremes of the time-interval, $\mathcal{F}(t)$ contributes to zero to the variational calculation. In terms of their variations, $\mathcal{A}$ and $\mathcal{A}'$ are equivalent. See the Appendix (A.3) for the algebraic details. □

The requirement for $\delta v$ to be constant in time is familiar to any one acquainted with Lorentz' transformations. It is interesting to mention that in the present context this requirement can be lifted by defining the variation as

$$\hat{\delta}_a = \sum_i \left( \delta\theta_i \mathcal{J}_i + \delta v_i \mathcal{I}_i + \frac{1}{2C} \frac{d\delta v_i}{dt} x_i \right) \tag{26}$$

# 4 Discussion and conclusions

From the point of view of pragmaticist epistemology (Peirce, 1955) all currents and charges are inferred. What we know about them are their effects, hence charge and current are "that what produces this and such effects", i.e., ideas, inferred entities, not directly accessible to our senses. However, charges and currents were originally associated to forces measured by a torsion balance and deflections of needles observed in galvanometers. Such primitive methods constitute the original definition of currents and charges and are available only for an observer at rest with the measuring apparatus since they are based upon material connections of circuits. In the text, we have restricted the use of "inferred" to those measurements that are performed based on action at a distance, i.e., without a "material" connection between the circuits (in particular, when this procedure is implied by the need of measuring while the detector is moving relative to the primary circuit, if the intricacies of a circuit continuously deforming are to be avoided). Thus, the scientist can perceive (measure) currents and charges using the original defining method if at rest relative to the source and the effects (as encrypted in forces, fields and potentials) of such events if the observer is at rest relative to the receiver. Currents and charges in the source are only inferred by the observer at rest with the receiver while forces, fields and potentials are inferred by the observer at rest with respect to the source.

We have shown that Electromagnetism can be formulated in terms of fields associated to sources as well as fields associated to receivers, this symmetry is broken in a construction that focuses exclusively in S-fields rather than R-fields.

Remark 3.5 exposes the relation between S- and R- fields, while the relation between source and receiver descriptions is given for example by the pair of equations (10) and (15). The restoration of this symmetry explains how the action-reaction law of Newton's mechanics is identically broken (Subsection 3.3) in the standard construction of Electromagnetism.

Most interestingly, the present approach based on eqs. (10), (22) and (15), is consistent with normal Electrodynamics and explains two fundamental concerns of the original theory, namely that electromagnetic waves propagate with the same electromagnetic parameter $C$ regardless of the state of relative motion between source and detector, and that the electromagnetic Doppler effect is acted by a Lorentz boost of parameter $u$, in agreement with the accepted description. These results are obtained within the original framework of the theory, in particular preserving the Euclidean character of the auxiliary space-time, $\mathbb{R}^{3+1} = \mathbb{R}^3 \times \mathbb{R}^1$, which fulfils the conditions imposed by spatial relations or relational space. In terms of interpretations, there is no need to regard the universal constant $C = (\mu_0 \epsilon_0)^{-\frac{1}{2}}$ as a velocity, nor to have something travelling between source and detector when considering electromagnetic interactions.

The complete set of equations of electromagnetism (Maxwell's equations, continuity equation and Lorentz' force) arise in the present form as the result of postulating Lorenz' delayed-action-at-a-distance 5 and Lorentz' action integral, to be used in the principle of least action 9. Lorenz' postulate has empirical basis while Lorentz' action is a (mathematical) organisation principle that has been considered fundamental by several authors as for example Poincaré (Poincaré, 1913). It is interesting to notice that before the irruption of the "second physicist" (Jungnickel and McCormmach, 2017), i.e., the theoretical physicist, theory in physics had a meaning close to "mathematically organised empirical observations". This is the spirit of Maxwell's work but it is as well the spirit in Newton, Ampere, Gauss and many others in the earlier times of physics. This epistemic position was heavily attacked by proponents of the ether such as Heaviside (Heaviside, 2011), Hertz (Hertz, 1893) and particularly Clausius (Clausius, 1869) who directly attacked Gauss' conception in the works by Riemann (Riemann, 1867), Betti (Betti, 1867) and Neumann (Neumann, 1868).

The present formulation addresses an issue recognised by Maxwell (1990, p. 228):

> ... According to a theory of electricity which is making great progress in Germany, two electrical particles act on one another directly at a distance, but with a force which, according to Weber, depends on their relative velocity, and according to a theory hinted at by Gauss, and developed by Riemann, Lorenz, and Neumann, acts not instantaneously, but after a time depending on the distance. The power with

> which this theory, in the hands of these eminent men, explains every kind of electrical phenomena must be studied in order to be appreciated [...]

And comparing with his preferred theory that "attributes electric action to tensions and pressures in an all-pervading medium" he writes:

> That theories apparently so fundamentally opposed should have so large a field of truth common to both is a fact the philosophical importance of which we cannot fully appreciate till we have reached a scientific altitude from which the true relation between hypotheses so different can be seen.

About one and a half century after Maxwell's conference we can discuss his philosophical inquire. Both theories are in perfect mathematical correspondence as they are with currently accepted electromagnetism but they differ in the abduction and interpretation as well as in the use of auxiliary concepts. Current electromagnetism relies heavily on the inferred idea of space-time and a mechanical analogy of the interaction. This approach is effective but leads us to embrace a new form of space-time, a necessary belief that not all of us are willing to admit. More precisely, current electromagnetism constructs first the space (relating it to Lorentz transformations) and only next spatial relations. In our view this order leads to logical inconsistencies (Solari and Natiello, 2022b) at the time of construction, despite the success achieved in terms of experimental comparisons. The present approach solves the problem by disposing of the subjective (auxiliary) space resting directly on spatio-temporal relations in such a way that rather than resting on just one transformation, a shared attribute of previous approaches, we find a harmonious coexistence of Galilean and Lorentzian transformations, a sort of reconciliatory mid-point. To achieve this views we had to accept first that space and time are not an a priori of knowledge as Kant thought (Kant, 1787) but rather a construction of the child as Piaget experimentally found (Piaget, 1999). Moreover, all these apparently conflicting approaches are needed for science to progress.

This view only uses (subjective) space and time as an auxiliary element when and if needed. The symmetry associated to the arbitrary decision of using a reference system is the one expressed by Galilean transformations (Solari and Natiello, 2018). In this context, inertial systems as auxiliary reference systems are constructed on the basis of the idea of free-bodies (Newton, 1687; Thomson, 1884; Solari and Natiello, 2021). This structure is underlying the work but not explicitly used as we have preferred to avoid reference frames.

The Lorentz transformations correspond in this construction to an endomorphism of relational space-time and are relevant only to the propagation of electromagnetic action. They are associated to the (unexpected) symmetry concerning perception and inferred charges and currents. While the full Poincaré-Lorentz group relates different but equivalent perceptions of the electromagnetic action, the particular set of Lorentz transformations relate the perception from a detector at rest with respect to the source to the perception of a detector in motion relative to the source with (invariant) velocity $u$. Such relation can be obtained only for $|u| < C$. We emphasise that the Poincaré-Lorentz group coexists with the Galilean symmetry of the description, although not all the equations, particularly not all differential equations, are transparent as expressions showing the symmetry. The integral presentation is in this respect more revealing.

We finally stress that symmetries as requirements of reason pre-exist physics and equations. They enter physics as a demand of reason in our quest to construct the cosmos, this is, to put in harmony our perceptions of the real-sensible.

The ether was the immediate consequence of attempts to understand electromagnetism by analogy to mechanical phenomena. Special relativity introduced an analogy of the forms, the Principle of relativity, without an understanding of the fundaments of the principle. It soon became evident that if analogies with mechanics would be preserved, the metric of space-time had to be changed. Yet, hiding the hypothesis the statement reads: electromagnetism imposes us to adopt a different metric of space-time than the Cartesian one used in its construction. Next, to accept this unmatching between the construction moment and the explanatory moment of science requires the exclusion of the first, leaving us with a science without understanding, supported only upon its predictive success, a technology of prediction, since success is the quality measure of any technology. Philosophers like Popper and Reichenbach considered their task to support the theories of scientists like Einstein. Consequently, they dropped all critical examination of matters, finally endorsing a program that was put forward by 1870, "physics must henceforth pursue the sole aim of writing down for each series of phenomena ... equations from which the course of the phenomena can be quantitatively determined; so that the sole task of physics consisted in using trial and error to find the simplest equations". Notice that even "trial and error", the method favoured by Popper, was already indicated. Such program is the instrumentalist program of science, aiming at dominating nature, a perfect mate of considering the Earth an infinite source of resources for the development of the capitalist society. We argue then that it has been forced upon us by social decisions that made nearly impossible the survival of the critical motion of reason. Conversely, by restating critical reasoning, we have been able to construct an electromagnetism that is more consilient than the received wisdom, it does not need to reform space-time and consequently makes no call for the abandonment of the construction moment. Reason can organise the chaos that reaches our senses, harmony is still an enticing possibility. If the child develops abstraction to understand the possible instead of being forced to accept the given, we need to put abstraction to work. We have been told that there is just one possible science, the given science, the science of capitalism. We have proved by presenting a counter example that the statement is wrong. Science is not only what scientists do (the given) but what humans can do as well, critical and ethical science, a science conscious of its ignorance. We close with ancient words by Chuang Tzu:

> Now you have come out beyond your banks and borders and have seen the great sea – so you realize your own pettiness. From now on it will be possible to talk to you about the Great Principle. (Chuang Tzu, 1968, Autumn Floods)

# Acknowledgements

We thank Alejandro Romero Fernández, Olimpia Lombardi and Federico Holik for valuable discussions.

# Declaration of Interest

The authors declare that there exists no actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could inappropriately influence, or be perceived to influence, their work.

# References

V I Arnold. *Mathematical Methods of Classical Mechanics. 2nd edition.* Springer, New York, 1989. 1st edition 1978.

Frederick C Beiser. *After Hegel.* Princeton University Press, 2014.

Enrico Betti. Sopra la elettrodinamica. *Il Nuovo Cimento (1855-1868)*, 27(1): 402–407, 1867.

Chuang Tzu. *The complete works of Chuang Tzu.* Columbia University Press, 1968. ISBN 0-231-03147-5.

R Clausius. Lxii. upon the new conception of electrodynamic phenomena suggested by gauss. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 37(251):445–456, 1869.

Herbert Dingle. The doppler effect and the foundations of physics (i). *The British Journal for the Philosophy of Science*, 11(41):11–31, 1960.

Michael Faraday. *Experimentl Researches in Electricity (Vol I)*. Richard and John Edward Taylor, 1839.

Michael Faraday. *Experimentl Researches in Electricity (Vol II)*. Richard Taylor and William Francis, 1844.

Michael Faraday. *Experimentl Researches in Electricity (Vol III)*. Richard and John Edward Taylor, 1855.

Carl Friedrich Gauss. *Carl Friedrich Gauss, Werke*, volume 5. K. Gesellschaft der Wissenschaften zu Göttingen, 1870. Digitizing sponsor University of California Libraries.

R Gilmore. *Lie Groups, Lie Algebras, and Some of Their Applications*. Wiley, New York, 1974.

M Hamermesh. *Groups Theory and its Application to Physical Problems*. Addison Wesley, Reading, MA, 1962.

Oliver Heaviside. *Electrical papers*, volume 2. Cambridge University Press (1894), 2011. Digitized by Google.

H Hertz. *Electric waves*. MacMillan and Co, 1893. Translated by D E Jones with a preface by Lord Kevin.

Christa Jungnickel and Russell McCormmach. *The Second Physicist (On the History of Theoretical Physics in Germany)*, volume 48 of *Archimedes*. Springer, 2017.

Matti Kaivola, Ove Poulsen, Erling Riis, and Siu Au Lee. Measurement of the relativistic doppler shift in neon. *Physical review letters*, 54(4):255, 1985.

Immanuel Kant. *The Critique of Pure Reason*. An Electronic Classics Series Publication, 1787. translated by J. M. D. Meiklejohn.

Immanuel Kant. The conflict of the faculties (the contest of the faculties). Der Streit der Fakultäten. Hans Reiss, ed., Kant: Political Writings, 2d ed. (Cambridge: Cambridge University Press, 1991), 1798. URL `la.utexas.edu/users/hcleaver/330T/350kPEEKantConflictFacNarrow.pdf`.

H A Lorentz. La théorie Électromagnétique de maxwell et son application aux corps mouvants. *Archives Néerlandaises des Sciences exactes et naturelles*, XXV:363–551, 1892. URL `https://www.biodiversitylibrary.org/item/181480#page/417/mode/1up`. Scanned by Biodiversity Heritage Library from holding at Harvard University Botany Libraries.

L Lorenz. Xlix. on the determination of the direction of the vibrations of polarized light by means of diffraction. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 21(141):321–331, 1861.

Ludvig Lorenz. Xii. on the theory of light. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 26(173):81–93, 1863.

Ludvig Lorenz. Xxxviii. on the identity of the vibrations of light with electrical currents. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 34(230):287–301, 1867.

Hirsch I Mandelberg and Louis Witten. Experimental verification of the relativistic doppler effect. *JOSA*, 52(5):529–535, 1962.

James Clerk Maxwell. A dynamical theory of the electromagnetic field. *Proceedings of the Royal Society (United Kingdom)*, 1865.

James Clerk Maxwell. *A Treatise on Electricity and Magnetism*, volume 1 and 2. Dover (1954), 1873.

James Clerk Maxwell. *The Scientific Letters and Papers of James Clerk Maxwell: Volume 1, 1846-1862*, volume 1. CUP Archive, 1990.

MA Natiello and HG Solari. Relational electromagnetism and the electromagnetic force, 2021. URL `https://arxiv.org/pdf/2102.13108`.

C. Neumann. Die principien der elektrodynamik. *Eine mathematische Untersuchung. Verlag der Lauppschen Buchhandlung, Tübingen*, 1868. Reprinted in Mathematischen Annalen, Vol. 17, pp. 400 - 434 (1880).

FE Neumann. Allgemeine gesetze der inducierten elektrischen strome. pogg. *Annalen der Physik (Poggendorf)*, 143(1):31–44, 1846.

Isaac Newton. *Philosophiæ naturalis principia mathematica ("Mathematical principles of natural philosophy")*. London, 1687. Consulted: Motte translation (1723) published by Daniel Adee publisher (1846). And the Motte translation revised by Florian Cajori (1934) published by Univ of California Press (1999).

Charles Peirce. *The Philosophical Writings of Peirce*. Dover Publications, 1955. Selected and edited by Justus Buchler.

Charles Peirce. Collected papers of charles sanders peirce. Charlottesville, Va. : InteLex Corporation,, electronic edition edition, 1994.

Jean Piaget. *The Construction Of Reality In The Child*. International Library of Psychology. Routledge, 1999. ISBN 0415210003,9780415210003.

H Poincaré. The theory of lorentz and the principle of reaction. *Archives Néerlandaises des sciences exactes et naturelles*, 5:252–278, 1900.

H Poincaré. *Science and Hypothesis*. The foundations of science. The science press, 1913. Translated by George B Halsted.

B. Riemann. Xlvii. a contribution to electrodynamics. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 34(231):368—-372, 1867. Translated from Poggendorff's Annalen, No. 6, 1867. Laid before the Royal Society of Sciences at Göttingen on the 10th of February 1858. Published posthumous.

Ludwik Silberstein. *The theory of relativity*. Macmillan, 2nd edition, 1924 edition, 1914.

H G Solari and M A Natiello. A constructivist view of newton's mechanics. *Foundations of Science*, 24:307, 2018. URL `https://doi.org/10.1007/s10699-018-9573-z`.

H. G. Solari and M. A. Natiello. On the relation of free bodies, inertial sets and arbitrariness. *Science & Philosophy*, 9(2):7–26, 2021. ISSN 2282-7757, eISSN 2282-7765. URL `http://eiris.it/ojs/index.php/scienceandphilosophy/article/view/669/851`.

H G Solari and M A Natiello. A critical assessment of scientific retroduction. 2022a. URL `https://philpapers.org/rec/SOLACA-2`.

H. G. Solari and Mario Natiello. Science, dualities and the fenomenological map. *Foundations of Science*, 2022b. URL `https://doi.org/10.1007/s10699-022-09850-4`.

Paul R Thagard. The best explanation: Criteria for theory choice. *The journal of philosophy*, 75(2):76–92, 1978.

James Thomson. On the law of inertia; the principle of chronometry; and the principle of absolute clinural rest, and of absolute rotation. *Proceedings of the Royal Society of Edinburgh*, 12:568–578, 1884.

Wilhelm Weber. Determinations of electromagnetic meassure; concerning a universal law of electrical action. *Prince Jablonowski Society (Leipzig)*, pages 211–378, 1846. Translated by Susan P. Johnson and edited by Laurence Hecht and A. K. T. Assis from Wilhelm Weber, Elektrodynamische Maassbestimmungen: Ueber ein allgemeines Grundgesetz der elektrischen Wirkung, Werke, Vol. III: Galvanismus und Electrodynamik, part 1, edited by H. Weber (Berlin: Julius Springer Verlag, 1893), pp. 25-214.

William Whewell. *The philosophy of the inductive sciences*, volume 1. JW Parker, 1840.

William Whewell. *The history of scientific ideas*, volume 1. JW Parker, 1858a. Third Edition.

William Whewell. *Novum organon renovastrum*. JW Parker, 1858b. Being the second part of The philosophy of Inductive sciences. Third edition.

# A Some Proofs

## A.1 Proof of Lemma 3.1

*Proof.* We perform the calculation in detail only for $A$, since the other one is similar. We use the shorthand $r = |x - y|$.

$$\begin{aligned}
\nabla_x A_i &= \frac{\mu_0}{4\pi} \int d^3y \left( j_i \nabla_x \frac{1}{r} - \frac{\frac{\partial}{\partial t} j_i \nabla_x \frac{r}{C}}{r} \right) \\
\Delta A_i &= \nabla_x \cdot \nabla_x A_i \\
&= \frac{\mu_0}{4\pi} \int d^3y \left( j_i \Delta \frac{1}{r} - 2 \left( \nabla_x \frac{1}{r} \right) \cdot \left( \frac{\partial}{\partial t} j_i \nabla_x \frac{r}{C} \right) \right) - \\
&\quad - \frac{\mu_0}{4\pi} \int d^3y \left( \frac{\frac{\partial}{\partial t} j_i \Delta \frac{r}{C}}{r} + \frac{\frac{\partial^2}{\partial t^2} j_i}{r} |\nabla_x \frac{r}{C}|^2 \right)
\end{aligned}$$

Moreover, standard vector calculus identities give

$$\begin{aligned}
\frac{\partial}{\partial t} j_i \left( 2 \nabla \frac{1}{r} \cdot \nabla \frac{r}{C} + \frac{\Delta \frac{r}{C}}{r} \right) &= 0 \\
|\nabla \frac{r}{C}|^2 &= \frac{1}{C^2}
\end{aligned}$$

and therefore

$$\Delta A_i(x,t) \;=\; \frac{\mu_0}{4\pi}\int d^3y\, j_i(y,t-\frac{r}{C})\Delta\left(\frac{1}{r}\right)+\left(\frac{1}{C^2}\right)\frac{\mu_0}{4\pi}\int d^3y\,\frac{\partial^2}{\partial t^2}\frac{j_i(y,t-\frac{r}{C})}{r}$$

The time derivative in the last term can be extracted outside the integral, thus yielding,

$$\begin{aligned}\Box A_i(x,t) &= \Delta A_i(x,t)-\left(\frac{1}{C^2}\right)\frac{\mu_0}{4\pi}\int d^3y\,\frac{\partial^2}{\partial t^2}\frac{j_i(y,t-\frac{r}{C})}{r}\\ &= \Delta A_i(x,t)-\left(\frac{1}{C^2}\right)\frac{\partial^2}{\partial t^2}A_i(x,t)\\ &= \frac{\mu_0}{4\pi}\int d^3y\, j_i(y,t-\frac{|x-y|}{C})\Delta\left(\frac{1}{r}\right)\\ &= -\mu_0 j_i(x,t)\end{aligned}$$

□

## A.2 Proof of Theorem 3.2

*Proof.* The result follows from the computation of the extremal action under the constraints

$$\begin{aligned}(V-\mathsf{V})\chi &= 0\\ (A-\mathsf{A})\chi &= 0\end{aligned}$$

Multiplying the constraints by the Lagrange multipliers $\lambda$ and $\kappa$ (the latter a vector), while we use the shorthand notations $B=\nabla\times A$ and $E=\left(-\frac{\partial A}{\partial t}-\nabla V\right)$, we need to variate the constrained electromagnetic action

$$\mathcal{A}=\frac{1}{2}\int dt\left(\int\left(\frac{1}{\mu_0}|B|^2-\epsilon_0|E|^2-\kappa\cdot(A-\mathsf{A})\chi+\lambda(V-\mathsf{V})\chi\right)d^3x\right).$$

Varying the integrand we obtain

$$\begin{aligned}\delta\mathcal{A} \;=\; &\int dt\left(\int\left(\frac{1}{\mu_0}(\nabla\times A)\cdot(\nabla\times\delta A)-\epsilon_0\left(\nabla V\cdot\nabla\delta V+\frac{\partial A}{\partial t}\frac{\partial\delta A}{\partial t}\right)\right.\right.\\ &\left.\left.-\epsilon_0\left(\nabla V\cdot\frac{\partial\delta A}{\partial t}+\nabla\delta V\cdot\frac{\partial A}{\partial t}\right)-\chi\kappa\,\delta A+\chi\lambda\,\delta V\right)d^3x\right)\end{aligned}$$

Partial integrations in time and standard vector calculus give the following identities:

$$\begin{aligned}
\int dt\,\frac{\partial A}{\partial t}\frac{\partial \delta A}{\partial t} &= \left[\delta A\cdot\frac{\partial A}{\partial t}\right]-\int dt\,\delta A\cdot\frac{\partial^2 A}{\partial t^2}\\
\int dt\,\nabla V\cdot\frac{\partial \delta A}{\partial t} &= [\delta A\cdot\nabla V]-\int dt\,\delta A\cdot\nabla\frac{\partial V}{\partial t}\\
(\nabla\times A)\cdot(\nabla\times\delta A) &= \nabla\times(\nabla\times A)\cdot\delta A-[\nabla\cdot((\nabla\times A)\times\delta A)]\\
\left(\nabla V+\frac{\partial A}{\partial t}\right)\cdot\nabla\delta V &= \left[\nabla\cdot\left(\left(\nabla V+\frac{\partial A}{\partial t}\right)\delta V\right)\right]-\delta V\nabla\cdot\left(\nabla V+\frac{\partial A}{\partial t}\right)
\end{aligned}$$

The terms in square brackets vanish in the variation either for the vanishing variation at endpoints or because of Gauss theorem applied to functions decaying fast enough at infinity. Hence,

$$\begin{aligned}
\delta\mathcal{A} \;=\; & \int dt\int d^3x\left(\frac{1}{\mu_0}(\nabla\times A)\cdot(\nabla\times\delta A)-\epsilon_0\left(-\delta V\nabla\cdot\left(\nabla V+\frac{\partial A}{\partial t}\right)\right.\right.\\
& \left.\left.-\delta A\cdot\frac{\partial^2 A}{\partial t^2}-\delta A\cdot\nabla\frac{\partial V}{\partial t}\right)-\chi\kappa\cdot\delta A+\chi\lambda\delta V\right).
\end{aligned}$$

Being $\delta A$ and $\delta V$ independent, we obtain

$$\begin{aligned}
\frac{1}{\mu_0}\nabla\times(\nabla\times A)+\epsilon_0\left(\frac{\partial^2 A}{\partial t^2}+\nabla\frac{\partial V}{\partial t}\right) &= \chi\kappa\\
-\epsilon_0\nabla\cdot\left(\nabla V+\frac{\partial A}{\partial t}\right) &= \chi\lambda
\end{aligned}$$

or equivalently

$$\begin{aligned}
\frac{1}{\mu_0}\nabla\times B-\epsilon_0\frac{\partial E}{\partial t} &= \chi\kappa\\
\epsilon_0\nabla\cdot E &= \chi\lambda
\end{aligned}$$

which allows us to identify $j=\chi\kappa$ (the density of current inside the material responsible for A) and $\rho=\chi\lambda$ (the density of charge responsible for V), thus proving the first result. Finally, the continuity equation follows from0

$$0=\nabla\cdot\left(\frac{1}{\mu_0}\nabla\times B-\epsilon_0\frac{\partial E}{\partial t}\right)+\frac{\partial}{\partial t}(\epsilon_0\nabla\cdot E)=\nabla\cdot j+\frac{\partial\rho}{\partial t}.$$

Note also that taking curl on the first equation we verify that $B$ satisfies a wave equation. Inserting $\nabla\times E$ in the time-derivative of the first equation and adding the gradient of the second equation, we obtain a wave equation for $E$.

Further, $\frac{1}{\mu_0}\nabla \cdot A + \epsilon_0 \frac{\partial V}{\partial t} = 0$ implies both that $\nabla \dfrac{\partial V}{\partial t} = -C^2 \nabla \left(\nabla \cdot A\right)$ and $\nabla \cdot \dfrac{\partial A}{\partial t} = -\dfrac{1}{C^2}\dfrac{\partial^2 V}{\partial t^2}$. Substituting each relation in the corresponding equation, we obtain eq(6), thus proving the Corollary. □

## A.3 Proof of Theorem 3.5

*Proof.* The rotational invariance is immediate, since for any rotation matrix $R$,the change of coordinates $x' = Rx$ (along with the corresponding change for $y$), keeps the distance $|R(x-y)| = |x-y|$ invariant. Hence, for the kernel in eq.(10) and any electromagnetic kernel depending on $|x-y|$ the action integral is invariant under rotations. Let $u$ be the velocity associated to a Lorentz transformation, which is constant by hypothesis. The proposed variation reads

$$
\begin{aligned}
\hat{\delta}\mathcal{A} &= \int_{-\infty}^{t} ds \int d^3x \, \hat{\delta} \left[A^1(x,s) j_2(x,s) - V^1(x,s)\rho_2(x,s)\right] \\
&= \int_{-\infty}^{t} ds \int d^3x \, \left(Cs\,\delta u \cdot \nabla_x + (\frac{x}{C} \cdot \delta u)\partial_s\right) \left[A^1 j_2 - V^1 \rho_2\right](x,s)
\end{aligned}
$$

where $\mathcal{I}(x,s) = Cs\,\delta u \cdot \nabla_x + (x \cdot \delta u)\partial_{Cs}$ is the Lorentz generator. By Gauss Theorem the following integral vanishes for any function $F$ inheriting the behaviour of $A, j$ at infinity:

$$
\int_K d^3x \, \delta u \cdot \nabla F(x) = \int_{\partial K} F(x)\delta u \cdot dS = 0
$$

Finally,

$$
\int_{-\infty}^{t} ds \, \frac{\partial}{\partial s} \int d^3x \, (x \cdot \delta u) G(x,s) = \mathcal{F}(t) - \mathcal{F}(t_0)
$$

for some function $\mathcal{F}$ depending only of $t$. However by the nature of the variational process, $\mathcal{F}$ does not contribute to the variation. □

# B The Lorentz transformation

The infinitesimal generator of the Lorentz transformation in eq.(16) reads

$$
\mathcal{I}_j = \begin{pmatrix} \mathbf{0} & \frac{v}{C} \\ (\frac{v}{C})^T & 0 \end{pmatrix}.
$$

The Lorentz transformation for finite $v$ is obtained by exponentiation (Gilmore, 1974), yielding the $4\times 4$ matrix expression $TL(v)$ for the transformation elements, where

$$TL(v) = \begin{pmatrix} W & X \\ X^{\dagger} & Y \end{pmatrix}$$

is formed by the 3-vector $X = \sinh\left(\left|\frac{v}{C}\right|\right)\dfrac{v}{|v|}$, the scalar $Y = \sqrt{1+|X|^2} = \cosh\left(\left|\dfrac{v}{C}\right|\right)$ and the $3\times 3$ matrix $W = Id + \left(\cosh\left(\left|\frac{v}{C}\right|\right) - 1\right)\dfrac{vv^{\dagger}}{|v|^2}$, where $v \in \mathbb{R}^3$ is a parameter classifying the different transformations. A better known expression for the Lorentz transformation arises from the change of variables $u = C\hat{v}\,\tanh\left|\dfrac{v}{C}\right|$ (Gilmore, 1974). In such terms,

$$L(u) = \begin{pmatrix} Id + (\gamma - 1)\hat{u}\hat{u}^{\dagger} & \gamma\dfrac{u}{C} \\ \gamma\dfrac{u^{\dagger}}{C} & \gamma \end{pmatrix}; \quad L\left(u(v)\right) \equiv TL(v),$$

where $\gamma\left(u(v)\right) = \dfrac{1}{\sqrt{1 - \left|\frac{u(v)}{C}\right|^2}} = \cosh\left(\left|\frac{v}{C}\right|\right)$. Note that for $\left|\dfrac{v}{C}\right| \ll 1$, we have that $\left|\dfrac{u}{C}\right| = \left|\dfrac{v}{C}\right| + O\left(|\dfrac{v}{C}|^3\right)$.

**Science & Philosophy** Volume 10(2), 2022

# An analysis of the concept of inertial frame in classical physics and special theory of relativity

Boris Čulina*

**Abstract**

The concept of inertial frame of reference in classical physics and special theory of relativity is analysed. It has been shown that this fundamental concept of physics is not clear enough. A definition of inertial frame of reference is proposed which expresses its key inherent property. The definition is operational and powerful. Many other properties of inertial frames follow from the definition, or it makes them plausible. In particular, the definition shows why physical laws obey space and time symmetries and the principle of relativity, it resolves the problem of clock synchronization and the role of light in it, as well as the problem of the geometry of inertial frames.

**Keywords**: inertial frame of reference; space and time symmetries; the principle of relativity; clock synchronization; physical geometry [1]

*University of Applied Sciences Velika Gorica, Velika Gorica, Croatia; boris.culina@vvg.hr

[1]Received on August 19th, 2022. Accepted on December 9th, 2021. Published on December 31th, 2022. doi: 10.23756/sp.v8i2.827. ISSN 2282-7757; eISSN 2282-7765. 

# 1 Introduction

The concept of inertial frame is a fundamental concept of physics. The opinion of the author is that not enough attention has been paid to such a significant concept, not only in textbooks, but also in the scientific literature. In the scientific and philosophical literature, many issues related to the concept of inertial frame have been addressed, but, as far as the author is aware, a systematic analysis of this concept has not been made. DiSalle's article (DiSalle [2020]) in the Stanford Encyclopedia of Philosophy gives an overview of the historical development of the concept of an inertial frame as an essential part of the historical development of physics. Thus, DiSalle's article is complementary to this article in its purpose and content. In this article, the concept of inertial frame of reference is analysed only within the framework of classical physics and special theory of relativity. This analysis could contribute to the analysis that has yet to be done: the analysis of the concept of inertial frame of reference in general relativity and especially in quantum physics.

The first part of this article identifies the basic properties of inertial frames in classical physics and special theory of relativity. The second part of the article gives a definition of inertial frame from which most other properties of inertial frames follow or this definition makes them plausible.

# 2 Analysis

## 2.1 Newton

In Philosophiæ Naturalis Principia Mathematica, one of Newton's goal is to describe absolute motion. This description also includes relative motion:[2]

> Absolute, true and mathematical time, of itself, and from its own nature flows equably without regard to anything external, and by another name is called duration: relative, apparent and common time, is some sensible and external (whether accurate or unequable) measure of duration by the means of motion, which is commonly used instead of true time ...
>
> Absolute space, in its own nature, without regard to anything external, remains always similar and immovable. Relative space is some movable dimension or measure of the absolute spaces; which our senses determine by its position to bodies: and which is vulgarly taken for immovable space ...
>
> Absolute motion is the translation of a body from one absolute place into another: and relative motion, the translation from one relative place into another ...

[2] English translation: Newton et al. [1846]

In this description Newton assumes that the geometry of absolute space is Euclidean geometry.

With his first law, the law of inertia, Newton describes the absolute motion of the body:

> Every body perseveres in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by forces impressed thereon.

The same is true for the other two Newton laws. However, Newton shows that these laws also apply to reference frames that move uniformly with respect to the absolute frame.

> The motions of bodies included in a given space are the same among themselves, whether that space is at rest, or moves uniformly forwards in a right line without any circular motion.

Today we call these frames inertial frames. Newton assumes that Euclidean geometry applies to them as well as to absolute space.

Newton's description of space and time provides a clear basis for his laws. These are absolute laws of absolute motion. But over time it has become clear that such an approach is untenable because it invokes "phantoms": absolute space and absolute time.[3] However, inertial frames remain as frames in which these laws apply. But how to define them when absolute space and absolute time are gone? Furthermore, from Newton we inherit the hypothesis that the centre of mass of the world rests in the absolute frame (Book 3 Hypothesis I), so that the centre of mass of the solar system, since it is far from other masses, moves uniformly relative to the centre of the world.[4] Thus, we can connect an inertial frame with the centre of mass of the solar system. This system can be well experimentally approximated by the requirement that fixed stars have a constant position in it. When we refer to the solar system as a reference frame below, we will mean this frame. Now, inertial frames can be defined as frames that move uniformly or are at rest relative to this frame. Experiments show that, with some limitations, Newton's laws as well as Euclidean geometry are valid in such frames.

## 2.2 Lange

The first constructive critiques of Newton's conception of inertial frame based on the concepts of absolute space and absolute time appear in the second half

[3] "In brief, Newton's absolute space is a phantom that should never be made the basis of an exact science." (Lange [1885])

[4] Immediately after Hypothesis I Newton makes a stronger claim, Proposition XI, that the centre of mass of the solar system also rests in absolute space. However, the assumptions stated in the proof are incomplete for such a conclusion.

of the 19th century. Lange (Lange [1885]) gives the following description of an inertial frame:[5]

> Definition I. An "inertial system"[6] is any coordinate system of the kind that in relation to it three points $P$, $P'$ , $P''$, projected from the same space point and then left to themselves – which, however, may not lie in one straight line – move on three arbitrary straight lines $G$, $G'$, $G''$ (e.g., on the coordinate axes) that meet at one point.
>
> Theorem I.[7] In relation to an inertial system the path of an arbitrary fourth point, left to itself, is likewise rectilinear.
>
> Definition II. An "inertial timescale" is any timescale in relation to which one point, left to itself (e.g., P), moves uniformly with respect to an inertial system.
>
> Theorem II. In relation to an inertial timescale any other point, left to itself, moves uniformly in its inertial path.

Lange defines a coordinate inertial frame as a frame in which three free particles released from a single point move in non-collinear straight lines. His definition assigns an experimentally verifiable condition but is not constructive in the sense that it does not give how to construct such a frame. Lange then postulates that all free particles in such a frame move in straight lines. Inertial time is defined as the time at which such a particle travels the same distance at the same time. This is the global time of an inertial frame and requires a measure in the geometry of the space of the inertial frame. Lange assumes Euclidean geometry. Again, the definition gives an experimentally verifiable condition but does not give a construction of such a time. Lange postulates that any other free particle travels the same distance in the same inertial time. The premise of the whole description is the existence of free particles and our ability to identify them.

Lange gives a successful analysis of the assumptions of Newton's first law. However, the basis of his approach is to single out the frame of measuring space and time according to how things will look in it, as a frame in which the motion of a free particle is the simplest – it is a uniform motion along a straight line. As Wheeler would say: "Time is defined so that motion looks simple."(Misner et al. [1973]). Although Lange assumes the concept of a straight line and Euclidean geometry, we could add to his analysis that space is defined so that a free particle moves along a straight line.

In the same spirit is another analysis of the concept of inertial frame given by Thomson (Thomson [1884]). He defines an inertial frame as a frame in which the bodies affected by the forces move according to Newton's laws and expresses the

[5] English translation: Lange [2014]

[6] The terms "inertial system" and "inertial timescale" come from Lange.

[7] In Lange's text, the word *theorem* has the meaning of a postulate.

law of inertia as the assertion to the existence of such a frame. So, here too, an inertial frame is determined by how things will look in it.

## 2.3 Modern textbooks

Modern textbooks of classical mechanics (not including the special theory of relativity) generally define an inertial frame in one of the following ways that we can relate to Newton's and Lange-Thomson's approach.

1. *The empirical approach.* An inertial frame is a frame that moves uniformly with respect to the solar system. It is postulated that Newton's laws apply in this frame (it is sufficient to postulate this for one such frame).

2. *The convenient approach.* An inertial frame is a frame in which Newton laws apply. Most often only the first law of inertia is mentioned, and the others are postulated. It is also postulated that the solar system is such a frame, as well as frames that move uniformly relative to it (it is enough to postulate it only for the solar system).

Although the concept of inertial frame is a fundamental concept, as a rule it is not analysed in modern textbooks of classical mechanics – the textbooks start from the concept in the development of mechanics. The internal structure of an inertial frame is not analysed, especially the mechanism of measuring space and time in such a frame. It is simply assumed, more often implicitly than explicitly, that space is Euclidean, and time is global. The empirical approach does not analyse why Newton laws would be valid in an inertial frame but states it as an experimentally confirmed statement. In the convenient approach, Newton laws are valid by definition. However, this definition is practically useless because, for example, we should examine the motions of all free particles with all velocities in all directions to determine whether Newton's first law is valid. This definition of inertial frame makes the term empirically unverifiable and does not show us how to construct such a frame. Therefore, as far as inertial frames are concerned, modern textbooks are a step backwards compared to Newton and Lange. Newton, using the concepts of absolute space and time, explains why his laws apply in inertial frames (because these frames move uniformly relative to absolute space) and why the solar system is inertial (it moves uniformly with respect to the centre of the world which is the absolute frame). Lange, in addition to bringing to light the important concept of inertial time, gives an empirically verifiable definition of inertial frame including inertial time in it.

In addition to the assumptions about an inertial frame that its space is Euclidean, time is global and absolute, and that Newton's laws apply, modern textbooks of classical mechanics sometimes assume, more often implicitly than explicitly, that in an inertial frame space is homogeneous and isotropic and time homogeneous and directed. There is no explanation as to why this would be the case (often it is not explained clearly enough what that means). Among the exceptions, the well-known Landau-Lifshitz textbook (Landau and Lifshitz [1976]) should be singled out – they define inertial frame using symmetries. In search of a frame "in which the laws of mechanics take their simplest form" they opt for a frame "in which space is homogeneous and isotropic and time is homogeneous". Such a frame they call inertial. Apart from the claim that such a frame "can always be chosen" and that "there is not one but an infinity of inertial frames moving, relative to one another, uniformly in a straight line", it is not stated how to operationally find such a frame. Furthermore, they assume Euclidean geometry and global time in such a frame. From this definition they derive Newton's first law, Lagrangian of a free particle, restrictions on the form of Lagrangian of a closed system, and conservation laws, thus showing that such a definition of inertial frame is very powerful.

The Landau-Lifshitz approach, which emphasizes the symmetries of space and time in an inertial frame, also belongs to convenient approaches that characterize an inertial frame as the frame in which the laws of mechanics are the simplest. Unlike this type of definitions that determine an inertial frame by how mechanical processes look in such a frame, a definition can be found in textbooks according to which an inertial frame is defined by its inherent property: it is a frame such that there are no external forces acting on it. However, neither such a description is sufficiently precise nor are the corresponding consequences drawn from the definition. For a typical example, we can cite a passage from Wikipedia (Wikipedia contributors [2021]):

> In classical physics and special relativity, an inertial frame of reference is a frame of reference that is not undergoing acceleration. In an inertial frame of reference, a physical object with zero net force acting on it moves with a constant velocity (which might be zero) – or, equivalently, it is a frame of reference in which Newton's first law of motion holds. An inertial frame of reference can be defined in analytical terms as a frame of reference that describes time and space homogeneously, isotropically, and in a time-independent manner. Conceptually, the physics of a system in an inertial frame have no causes external to the system.

If we understand the first statement as a definition of an inertial frame, we have a typical situation in this approach: various properties of an inertial frame are listed, and they are in no way related to the definition.

## 2.4 The special theory of relativity

The special theory of relativity has brought key improvements in the conception of inertial frame. In his groundbreaking work (Einstein [1905]) Einstein starts from the established concept of inertial frame: it is a frame in which "the equations of mechanics hold good". Thus, he accepts all classical assumptions, first of all Euclidean geometry which he considers realized by means of an extended solid body and rigid rods. However, in this paper, Einstein introduces two essential innovations related to inertial frames. The first is the generalization of the principle of relativity: not only are the laws of classical mechanics the same in all inertial frames, but all the laws of physics are the same in all inertial frames. Another innovation is the analysis of the concept of time in an inertial frame. Einstein starts from the fact that time is measured locally – with the same clock. He assumes that in an inertial frame at each point in space we can have identical clocks that we need to synchronize to get the global time of the inertial frame. Einstein describes the synchronization of clocks at different places $A$ and $B$ in an inertial frame as follows:[8]

> We have so far defined only an "$A$ time" and a "$B$ time". We have not defined a common "time" for $A$ and $B$, for the latter cannot be defined at all unless we establish by definition[9] that the "time" required by light to travel from $A$ to $B$ equals the "time" it requires to travel from $B$ to $A$. Let a ray of light start at the "$A$ time" $t_A$ from $A$ towards $B$, let it at the "$B$ time" $t_B$ be reflected at $B$ in the direction of $A$, and arrive again at $A$ at the "$A$ time" $t'_A$.
>
> In accordance with definition the two clocks synchronize if
>
> $$t_B - t_A = t'_A - t_B.$$
>
> We assume that this definition of synchronism is free from contradictions, and possible for any number of points; and that the following relations are universally valid:—
>
> 1. If the clock at $B$ synchronizes with the clock at $A$, the clock at $A$ synchronizes with the clock at $B$.
> 2. If the clock at $A$ synchronizes with the clock at $B$ and also with the clock at $C$, the clocks at $B$ and $C$ also synchronize with each other.
>
> Thus with the help of certain imaginary physical experiments we have settled what is to be understood by synchronous stationary clocks located at different places, and have evidently obtained a definition of "simultaneous", or "synchronous", and of "time". The "time" of an event is that which is

[8]English translation: Lorentz [1952]

[9]This part of the translation is wrong and should read: "...and the latter can now be determined by establishing by definition..."(J.D.Norton).

> given simultaneously with the event by a stationary clock located at the place of the event, this clock being synchronous, and indeed synchronous for all time determinations, with a specified stationary clock.
>
> In agreement with experience we further assume the quantity
>
> $$\frac{2\mathrm{AB}}{t'_A - t_A} = c,$$
>
> to be a universal constant — the velocity of light in empty space.
>
> It is essential to have time defined by means of stationary clocks in the stationary system, and the time now defined being appropriate to the stationary system we call it "the time of the stationary system".

In short, Einstein a) gave the definition of synchronization of two clocks by light, b) postulated that all clocks of an inertial frame can be consistently synchronized in the sense that synchronization of two clocks is an equivalence relation with exactly one equivalence class, c) that once synchronized clocks remain synchronized, and d) that the two-way speed of light (the speed measured on the same stationary clock) is the universal constant of an inertial frame.

Having thus obtained the global time of an inertial frame ("stationary system", in Einstein's words), Einstein can define the concept of one-way velocity in an inertial frame and state his second postulate (the light principle):

> Any ray of light moves in the "stationary" system of co-ordinates with the determined velocity $c$, whether the ray be emitted by a stationary or by a moving body. Hence
>
> $$\text{velocity} = \frac{\text{light path}}{\text{time interval}}$$
>
> where time interval is to be taken in the sense of the definition in § 1.[10]

Although Einstein, with his generalized principle of relativity and the light principle based on the analysis of the concept of time in an inertial frame, revolutionized physics, some things remained insufficiently clarified in the key part of his article quoted above:

1. The problem of the conventionality of the definition of synchronization (Reichenbach [1928]). Every definition of synchronization which is of the form $t_\mathrm{B} = t_\mathrm{A} + \varepsilon(t'_\mathrm{A} - t_\mathrm{A})$, where $0 < \epsilon < 1$, is in accordance with the principle of causality. Is Einstein's choice $\varepsilon = \frac{1}{2}$ physically different from other choices or is it just a pleasant convention with no physical significance? This problem has generated controversy that is still present today (Anderson et al. [1998], Jammer [2006], Janis [2018]).

[10] Einstein refers here to "the time of the stationary system" previously described.

2. The problem of consistent synchronization (in Einstein's sense) of all clocks. What properties of light are needed to achieve this? In particular, what is the role of the postulate of the constancy of the two-way speed of light in clock synchronization?

3. The problem of possible circularity. Light signals are used for synchronization in order to express the light principle about light with the help of such synchronized clocks. For example, from the very definition of synchronization it follows that the one-way speed of light in opposite directions is the same. If we add to this the postulate of the constancy of the two-way speed of light, we get the light principle as a consequence of synchronization and not as an additional postulate. Thus, although Einstein introduces the clock synchronization procedure to articulate the light principle, in his work it remains unclear which properties of light are required for synchronization. The problem of circularity also occurs at a deeper conceptual level because Einstein uses clock synchronization to define the global time of an inertial frame. However, he describes an inertial frame as a frame in which "the equations of mechanics hold good". These laws contain the law of inertia, which presupposes the global time of an inertial frame in its formulation, and which Einstein's clock synchronization has yet to establish.

In Minkowski [1909], Hermann Minkowski gave the formulation of the special theory of relativity in terms of a certain structure in the space of events. In short, the Minkowski event space is a 4-dimensional affine space in which worldlines of free particles and light are special types of straight lines (timelike and lightlike straight lines), and in which the metric tensor is given that is directly related to light signalling and time measurement by means of free-moving clocks. It is an elegant mathematical reformulation of the special theory of relativity that does not introduce essentially new elements into the concept of an inertial frame. We can understand Minkowski space as the structure in the event space generated by the structure of an inertial frame in a way that is invariant to the choice of an inertial frame. The light principle and the principle of relativity are automatically built into this structure (the principle of relativity as a condition on the physical laws that they must be formulated in terms of Minkowski space). Conversely, inertial frames can be understood as decompositions of Minkowski space to which the structure of Minkowski space is isomorphically transferred. In such a decomposition the space of an inertial frame is still Euclidean and the decomposition itself corresponds to Einstein's clock synchronization (Malament [1977]). However, in Minkowski's formulation the inherent property of an inertial frame becomes more visible. Namely, the worldlines of free particles are timelike straight lines in that space, so each inertial frame is identified as a class of all mutually parallel timelike straight lines. If we imagine that each such straight line is a worldline of a free particle, and not a particle acted upon by forces in equilibrium, then an inertial frame is a class of all free material particles that are at rest with each other. Thus, an inertial frame in this space of events naturally appears as a frame by which we can identify all events and whose main feature is that it is free, that its elements do not enter any interactions.

## 2.5 The general theory of relativity and quantum physics

Here we will dwell only on some general observations on the possibility of extending the above analysis to general relativity and quantum physics.

As is well known, the general theory of relativity sets physical limits on the classical concept of an inertial frame. Regardless of how we describe an inertial frame, the essential concept is the concept of free particle, the concept that is incompatible with the ubiquity of gravity and must be reformulated into the notion of a free-falling particle. Thus, the classical concept of inertial frame can only be realized approximately, within a limited space and time. Nevertheless, it is the key idealization of the general theory of relativity, the "infinitesimal" element of which the whole theory is composed. Note that even in the general theory of relativity, inertial frames have a natural inherent description: they are free-falling frames.

Although inertial frames are an essential element of quantum description of the world, they are rarely explicitly mentioned in quantum physics textbooks. If they are mentioned, they are not analysed, but their properties from classical (non-quantum) physics are simply transferred. In Bohr's approach (Howard [1994], Tanona [2004]), they are a macroscopic element that is an integral part of the quantum description of the world and is usually related to macroscopic measuring instruments. In such an approach, the classical concept of an inertial system retains its importance. However, in other approaches the concept loses its meaning. For example, in the approach described in (Aharonov and Kaufherr [1984], Angelo et al. [2011]) an inertial frame itself must be a quantum mechanical system. Then some classical properties of an inertial frame must be reformulated. For example, the property that it is a frame in which free particles move uniformly in straight lines is transformed into the property that the expected value of the position of a free particle changes uniformly along a straight line – the property that is difficult to verify experimentally. On the other hand, the characterization that an inertial frame is a frame on which nothing acts still makes sense. Furthermore, in quantum physics, space and time symmetries can be attributed to an inertial frame, as well as the principle of relativity. However, the quantum mechanical properties of an inertial frame make the basic purpose of such a frame problematic – to identify when and where something happened. What kind of such identification does quantum physics enable, that is, what kind of structure does it bring into the space of events? In particular, what geometry does it introduce into the space of an inertial frame? These are the key unanswered questions (Penrose [1968]):

> I do not believe that a real understanding of the nature of elementary particles can ever be achieved without a simultaneous deeper understanding of the nature of space-time itself.

## 2.6 Properties of an inertial frame

After this review, let us summarize which properties are attributed to inertial frames:

1. A frame in which the centre of mass of the solar system is at rest and in which the fixed stars have a constant position is an inertial frame.
2. The space and time of an inertial frame are such that free particles in it move uniformly in straight lines. In general, it is a frame in which Newton's laws apply.
3. An inertial frame is a frame in which the laws of physics have the simplest form.
4. An inertial frame is a frame on which there are no external forces. Or even more restrictively, it is a frame composed of free particles (there are no external or internal interactions) that are at rest with each other.
5. The space of an inertial frame is Euclidean.
6. The space of an inertial frame is homogeneous and isotropic, and time is homogeneous and directed.
7. The time of an inertial frame is local, the local times can be synchronized and so the global time of an inertial frame can be obtained.
8. Frames which move uniformly in straight lines relative to an inertial frame are inertial frames and there are no other inertial frames.
9. The principle of relativity: The laws of physics are the same in all inertial frames.
10. The light principle: The speed of light in vacuum is the same in all inertial frames, regardless of the mode of light formation.

# 3 Definition and consequences

## 3.1 Definition of the concept of inertial frame

For a successful definition of a property, it is not enough that, in addition to formal correctness, the definition is only extensionally correct – objects that have the defined property are precisely those objects that we want to single out from some multitude of objects. Such is, for example, Plato's definition of man as a two-legged animal without feathers. The most important criterion that the definition should meet is to be intensionally correct – to single out objects according to some of their essential properties. Unlike the first two criteria, we are not yet able to give this third most important criterion a sufficiently precise form.[11] But this does not mean that in particular cases we cannot distinguish better from worse definitions. Of course, in the choice of a definition, the criteria of precision (how precise the terms we use to define a new term) and effectiveness (how effectively we can examine whether an object has a defined property) are important, too. Of the properties listed in the previous section, inertial frames are characterized in an extensional sense by properties 1) together with 8), and properties 2), 3) 4) and 6).

Criterion 1) plus 8) is an experimental determination. Thus, its meaning is poor, and we cannot relate it to other properties of an inertial frame. We can only postulate them independently.

Criteria 2), 3) and 6) identify an inertial frame by how physical processes look in it. These are external characterizations of an inertial frame that cannot explain its other properties. In addition, these criteria are not operational – they do not show how to find such a frame. Criterion 3), in addition to being imprecise (what does it mean to have the simplest form?), provides no basis for identifying such frames. Since we do not know all the laws of physics, we cannot know in which frame they have the simplest form. Moreover, it is possible that some laws have the simplest form in one type of frame and other laws in another type of frame. Criterion 2) is clear because it is limited to Newton's laws. But accepting this criterion would mean an unnecessary limitation of the concept of an inertial frame to classical Newtonian physics. The necessary universality can be obtained only if criterion 2) is limited to the description of the motion of a free particle. The main purpose of the reference frame is to identify where and when something happened, and the requirement that in an inertial frame a free particle moves rectilinearly and uniformly, is precisely the requirement for the space and time of the frame. But such a requirement is too weak to be related to other properties of an inertial frame. If we want to reinforce it with other requirements for space and time, primarily space and time symmetries, then we come to criterion 6). How-

[11] The criterion of extensional correctness has as precise a form as it is clear to us on an extensional level which objects we want to single out.

ever, an inertial frame must have inherent physical characteristics that affect what the laws of physics look like in it and not to be adjusted so that in it those laws have a certain form. Furthermore, this second approach only makes sense if we can formulate physical laws independently of the concept of a reference frame, which is operationally questionable. This is true for space and time symmetries, too. Such an external characterization only makes sense if we can describe this space and time structure independently of the concept of reference frame. For example, in Brehme [1985] inertial frames are defined as frames that are isomorphic to the Minkowski space. However, the structure of Minkowski space is operationally derived from the structure of inertial frames, so this definition is only an elegant mathematical solution until we give Minkowski space a direct physical interpretation. This interpretation must explain why physical laws must be formulated in Minkowski space, that is, why they must have space and time symmetries, as well as satisfy the principle of relativity. However, even if we were to achieve such a definition of an inertial frame, structurally we would obtain a characterization of an inertial frame that it is a frame that (due to an isomorphism) has the structure of a Minkowski space. But again, it is an external characterization that does not tell us why an inertial frame would have such a space and time structure. Also, the definition would not be operational.

Criterion 4) is the only inherent criterion, a criterion that mentions the properties of the reference frame itself. While the aforementioned characterizations identify an inertial frame by how we describe physics in it, this characterization determines an inertial frame by what happens to the frame itself. Thus, in terms of intensional correctness, it is the best criterion. It is also an operational criterion, unlike criteria 2) 3) and 6). In addition, it is very powerful. When we clarify the basic idea that an inertial frame is a free frame, a frame on which nothing acts, we will get a definition of inertial frame from which almost all the remaining listed properties of an inertial frame can be derived or at least made plausible. For properties that cannot be related to the concept of inertial frame, it will be shown that there are good reasons why, by their nature, they do not fall under the concept. Therefore, we will take criterion 4) to define inertial frame.

We will call *reference frame* any frame that allows us to identify events spatially and temporally. The same reference frame can provide multiple coordinate systems for the identification. For example, the Euclidean space is a reference frame for determining position, and various coordinate systems for identification can be defined in it. Thus, we will distinguish a reference frame from the coordinate reference frame that can be built in it. An inertial frame will be a special type of a reference frame.

The condition that there are no external forces on an inertial frame is too weak. If we look at a solid body that is not affected by external forces, it can rotate. This rotation is registered by the appearance of internal tensions in the body. However, the condition that we do not allow internal forces in an inertial frame is too strong. Since an inertial frame must provide spatial and temporal determination of events, it must contain certain measuring instruments. Therefore, we will allow the existence of localized closed parts within which there is an interaction, but not the existence of a non-localized interaction, such as interactions caused by rotation, that could disrupt the symmetries of an inertial frame. This does not preclude the existence of large-scale solid bodies in an inertial frame, because in the absence of external forces and rotation we can ignore internal tensions in the body. There can only be isolated and localized interactions.

Since an inertial frame serves to determine the space and time coordinates of an event, it must also have the ability to determine that its parts are at rest relative to each other. Only localized deviations from rest in closed processes that serve to measure space and time are allowed.

Based on the above considerations, we define *inertial frame* as a reference frame such that the following holds:

1. There are no external forces on the frame.

2. Interactions within the frame are possible only in localized and closed parts of the frame.

3. Parts of the frame are at rest, except for possible localized deviations from rest.

The precision of this definition is limited by the precision of the terms used in it, but we will show that it is precise enough to be usable.

This definition does not follow from Newton's description of inertial frames as frames that move uniformly relative to the absolute frame. The law of inertia states that free frames move uniformly relative to the absolute frame, but the reverse is not true: frames that move uniformly relative to the absolute frame do not have to be free – these include frames that are affected by forces in equilibrium.

Likewise, this definition does not follow from the standard definition of an inertial frame as a frame in which free particles move uniformly in a straight line or are at rest. Parts of such a frame are at rest in the frame but this does not mean that they are free – this includes parts that are affected by external or internal non-localized forces that are in equilibrium.

This definition of inertial frame is one of the standard definitions of inertial frame, somewhat more precise here than usual. It is suggested by Newton's approach and by the standard definition through the observation of a free particle, but it is more restrictive than these definitions, as shown in the previous paragraphs. Such a definition occurs naturally from the aspect of event space (whether it has a Galilean structure or a Minkowski structure), as well as from the aspect of general relativity, where it corresponds to free-falling reference frames, if we localize them in space and time enough.

The most important term on which the definition of an inertial frame rests is the concept of interaction. Thus, for the definition to be operational, it assumes that we know what kind of interactions exist. However, other definitions are based on this same concept, too. For example, the same assumption lies behind the concept of a free particle in Newton's first law. Ultimately, we can understand this definition as a working definition, which changes every time we discover new interactions.

If we assume that the interactions decrease with distance then we can consider that any frame that is far enough away from other bodies and in which there are no non-localized interactions is approximately inertial. Thus, the solar system (a system in which the centre of mass is at rest, and which has a constant direction with respect to fixed stars) can be considered inertial. Since external actions as well as rotation cause non-localized tensions, we can experimentally check whether a reference frame is inertial with an appropriate system of accelerometers and gyroscopes. For Einstein, an inertial frame is tied to the extended rigid body. If there are no external actions on the rigid body and it does not rotate, then its parts are free and in a constant mutual position, so it determines an inertial frame. A free observer with a clock and theodolite, which sends light signals around and measures the time of sending and receiving signals also forms an inertial frame. Of course, the definition of inertial frame formulated here is an idealization in relation to which we can estimate how much the actual frame of reference corresponds to an inertial one. As already mentioned, the most significant restriction on the realization of such frames is set by the general theory of relativity, but also by quantum physics.

## 3.2 Space and time symmetries of an inertial frame

Since an inertial frame is free, the space of this frame is homogeneous and isotropic. Any inhomogeneity and non-isotropy would mean the existence of external forces or an unnecessary internal symmetry breaking (e.g., to choose a different unit of measure in each direction). Likewise, any time inhomogeneity would mean the presence of external forces or unnecessary internal symmetry breaking, so such a space is also time homogeneous. These inherent symmetries of an inertial frame can be extended to measurements of space and time but also to the description of all closed processes in such a frame. By definition, a *closed system* has no interaction with the environment, so the events in it are independent of the space and time in which it is located. When such a system is observed from an inertial frame in which all points, directions and time moments are equal, then in an inertial frame such a system can be described in such a way that it has the specified symmetries. Thus, we can set *the symmetry principle* of describing physical processes in an inertial frame: *In an inertial frame the laws of physics for closed systems have space homogeneity and isotropy as well as time homogeneity*. This principle derives not only from how physical processes take place but also from how we can describe them in an inertial frame. Note, we do not have to describe a closed physical process that way. But an inertial frame gives us the ability to describe them that way, and it is to be expected that such a description is the best in every respect. I would also note that it makes sense to state this principle even before we have an elaborate structure of space and time. Moreover, we can and must (if we want to exploit the advantages of an inertial frame) apply it to the very determination of the geometry of space and the structure of time in an inertial frame. The approach to the description of physical processes in which we try to preserve the original symmetries of an inertial frame I will call *the inertial frame approach*.

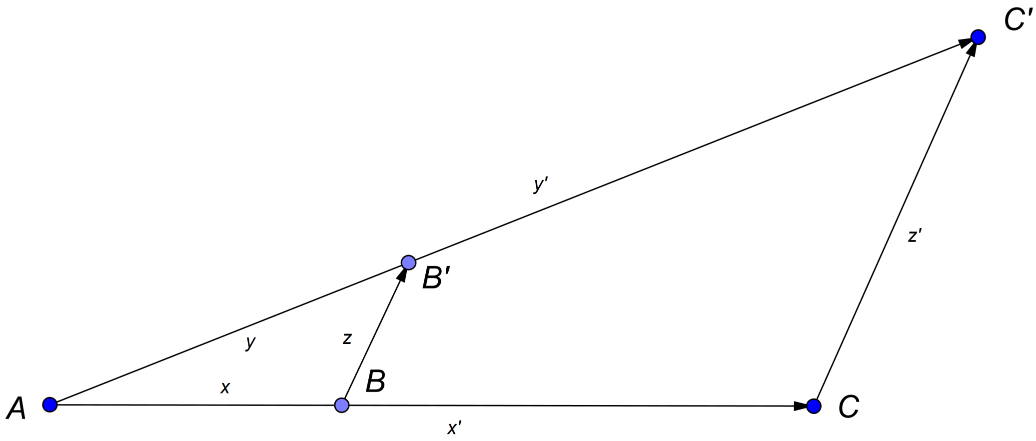
## 3.3 Space and time of an inertial frame

The simplest closed system is a system composed of one (massive) particle – it is a free particle. The ray of light, if we assume the absence of ether, is also a simple closed system.

Each closed periodic process, i.e., the process that returns to the initial state, including the initial position, determines the local measurement of time at that place. This is the general definition of *local clock*. It could be an atomic clock. It can also be a free particle or light that bounces off something and returns to its starting position (Langevin clock). For a unit of time we can take some standardized process, for example a free particle created in the standard way that bounces off something or a light particle created in the standard way that bounces off something, if we assume that it is a closed system (that there is no ether). Here we do not have to assume that no matter how we create light it always gives the same unit of time (the light principle). Due to space and time symmetries, all closed periodic processes must measure the same time up to the choice of the unit of measure and their operation is independent of position, orientation, and elapsed time. Let us show more precisely that this is so. Let us have two closed periodic processes (two clocks) $C_1$ and $C_2$ to measure time. In general, the relation between the times $t_1$ and $t_2$ of the duration of a process measured by the clocks $C_1$ and $C_2$ is a continuous function $f$: $t_2 = f(t_1)$. We will show that this function is a direct proportion: $t_2 = a \cdot t_1$, for some $a > 0$. For this purpose, we will consider the time $a$ of the duration of the periodic process $C_1$ measured on the clock $C_2$. Measured on the clock $C_1$ it is equal to 1. Thus $a = f(1)$. Measuring by clock $C_2$ the duration of $n$ consecutive $C_1$ processes, due to time homogeneity, will give $na = f(n)$.This means that if measuring any process with the clock $C_1$ yields fraction $\frac{n}{m}$ then the measurement with the clock $C_2$ will yield $\frac{n}{m}a = f(\frac{n}{m})$. Because of the continuity of the function $f$, this means that $x \cdot a = f(x)$, for any positive real number $x$. Thus, the function $f$ is a direct proportion. By adding spatial symmetries to this, we have shown that *all clocks at all points of an inertial frame are equal*.

Since free particles or rays of light, assuming the absence of ether, are the simplest examples of closed systems, their trajectories must be the simplest examples of curves in the space of an inertial frame. Due to the homogeneity and isotropy of space, such a trajectory must have the same spatial characteristics at each location. So, it is natural to take these paths to define *straight lines* in the space of an inertial frame. Distances can always be measured by the same standardized periodic process by which we locally define time – by means of a standardized free particle or a standardized light (assuming no ether). If since the sending of the standardized free particle (or light) from the point $A$, its rejection from the point $B$ and its return to the point $A$ the elapsed time $t$ is measured on the clock in $A$, then we can take that time for the measure of distance. But due to the isotropy of space, it is more natural to take half of that time to measure *distance*: $d(A, B) = \frac{t}{2}$. This does not change anything significantly because the measurement is always determined up to the multiplicative factor. Note that in this way we can also check an important element of the definition of inertial frame, that the parts of the frame are at rest. After the reflection of a particle or light, there is a displacement of the body from which the reflection is made. However, in an inertial frame, such localized deviations from rest are allowed by definition, provided that this shift is subsequently reversed. *Due to space and time symmetries, any choice of standardized free particles or standardized light (assuming no ether) gives the same geometry up to the unit of measure*. Locally, we can make these measurements more conveniently using rigid rods. Due to space and time symmetries, such a measuring instrument, as well as the measuring system generated by it, can be reproduced at any point and in all directions. And, due to the symmetries, this leads to the same geometry. The geometry of the space of an inertial frame, the geometry in which the paths of free particles and light rays are straight lines, and in which the distance measurement is based on the described measurement of elapsed local time, is homogeneous and isotropic – all points are equal and all directions are equal. Although scale symmetry is not generally valid for closed processes, we will show that it is valid for the geometry of an inertial frame. We will show that Thales's basic proportionality theorem holds: for the lengths of the segments marked in the figure below it holds that if $x' = \alpha x$ and $y' = \alpha y$ then $z' = \alpha z$.

This is due to the equality of all clocks in an inertial system. Let $x$, $y$ and $z$ lengths

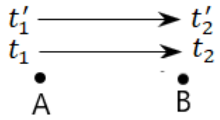



$$x = \overline{AB},\ y = \overline{AB'},\ z = \overline{BB'},\ x' = \overline{AC},\ y' = \overline{AC'},\ z' = \overline{CC'}$$

Figure 1: Thales's proportionality theorem

be measured using a periodic closed process $C_1$ to which a unit measure is equal to $c_1$: $x = t_x c_1$, $y = t_y c_1$ and $z = t_z c_1$. Imagine another periodic closed process $C_2$ such that its unit of measure is equal to $c_2 = \alpha c_1$. Measuring $x'$ and $y'$ with the clock $C_2$ gives the same numerical value as measuring $x$ and $y$ with the clock $C_1$: $x' = t_x c_2$ and $y' = t_y c_2$. Since the geometry is determined by measuring time, and all clocks are equal, just as the numbers measured with the clock $C1$ determine the triangle $x - y - z$, so the same numbers measured with the clock $C_2$ determine the triangle $x' - y' - z'$. That's why it has to be $z' = t_z c_2$. Therefore, $z' = \alpha z$. This proves Thales's theorem. In Čulina [2017] it is shown that from the assumptions of homogeneity, isotropy and scale symmetry of space axioms of Euclidean geometry can be obtained. This means that the geometry of an inertial frame is Euclidean geometry.

The symmetries of space and time solve both the problem of clock synchronization (including the question of conventionality) as well as the problem of possible circularity of the description. Using free particles or light, we can synchronize clocks in an inertial frame with the same procedure we used to determine the measurement of distances – we send a standardized free particle or light (assuming no ether) from one clock to the next and back. Symmetries give us the freedom to choose the means of synchronization. We can use any standardized free particle (we standardize the way of generating its motion) or a standardized ray of light, assuming that the light is a closed system (that there is no ether). If we assume that the motion of light is independent of the source of origin (the light




$$t_1' - t_1 = t_2' - t_2$$
$$t_2 - t_1 = t_2' - t_1'$$

Figure 2: time homogeneity of clocks

principle), then we do not have to standardize light at all. *Due to the above symmetries, whatever standardized process we use, we will always get the same clock synchronization* (if we synchronize them in one way, we will find for every other way, that it gives the same synchronization). If we synchronized the clock $B$ with the clock $A$ and that the synchronized clocks will remain synchronized, we will denote $A$ sinc $B$.

Even before synchronization, time homogeneity of an inertial frame tells us something about the connection of the time read by the clocks at the places $A$ and $B$. If we sent the standard signal from the clock at $A$ in the moments $t_1$ and $t_1'$, and the clock at $B$ received them in the moments $t_2$ and $t_2'$, then the difference in elapsed time is on both clocks same: $t_1' - t_1 = t_2' - t_2$. This is equivalent to the condition that the difference in signal travel time read on the $B$ clock on arrival and on the $A$ clock on departure is always the same: $t_2 - t_1 = t_2' - t_1'$. We will call this property of clocks *time homogeneity of clocks* in an inertial frame.

If we look at all possible synchronizations that are in accordance with the principle of causality, they are of the form $t_2 = t_1 + \varepsilon(t_1, A, B)(t_3 - t_1)$, $0 < \varepsilon(t_1, A, B) < 1$, where $t_1$ is the time read on the clock at $A$ when sending the signal from $A$, $t_2$ is the time read on the clock at $B$ when the signal arrives at $B$, and $t_3$ is the time read on the clock at $A$ when the signal returns to $A$. Due to the time homogeneity, $\varepsilon(t_1, A, B)$ must not depend on $t_1$ and due to space symmetries it must not depend on $A$ and on the direction towards $B$ – it must be the same number $\varepsilon$ for all points. In particular, it must be the same number to synchronize the clock at $A$ with the clock at $B$. We can get this synchronization by reflecting the previously described signal once again back to the point $B$ where its arrival will be read at the moment $t_4$ on the clock at $B$.
Synchronizing the clock at $B$ with the clock at $A$ gives

$$t_2 = t_1 + \varepsilon(t_3 - t_1)$$

To get the relationship $t_3$ with $t_2$ and $t_4$, we will eliminate $t_1$ in the above relationship using time homogeneity of the clocks:

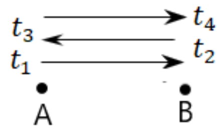



Figure 3: Synchronizing clocks

$$t_3 - t_1 = t_4 - t_2$$

We will get

$$t_3 = t_2 + (1 - \varepsilon)(t_4 - t_2)$$

Due to space isotropy, it must be $1 - \varepsilon = \varepsilon$, or $\varepsilon = \frac{1}{2}$. The conclusion is that the symmetries of an inertial frame require that the synchronization relation is symmetric, that is, that $\varepsilon = \frac{1}{2}$. Other choices would break the symmetry. Thus, the concept of inertial frame leads to Einstein's clock synchronization and not some other synchronization. This choice is not just a matter of convention, but it is part of the inertial frame approach (page 22) to the study of nature. The situation is the same as when setting symmetry conditions on physical laws. So here too, an inertial frame allows us to choose Einstein's synchronization, and we must certainly take advantage of this in the study of nature – to keep the symmetries of inertial frame, and so to choose Einstein synchronization. In what follows, we will mean by synchronization precisely this symmetrical Einstein synchronization.

It is not difficult to show that time homogeneity of clocks is equivalent to the condition that once synchronized clocks remain synchronized, and the symmetry of a synchronization relation is equivalent to the condition that $\varepsilon = \frac{1}{2}$.

With Einstein's synchronization, we can synchronize all clocks with one clock in a symmetrical way. However, due to space and time symmetries of an inertial frame, we will get the same result no matter what clock we take for the synchronizing clock. Thus, an inertial frame realizes Einstein's assumption of consistency of synchronization. We can show this in more detail. It follows from the isotropy of space that if we send a signal from point $A$ so that it comes to point $B$, it bounces to point $C$, from where it bounces back to point $A$, the time $t_{\leftarrow}$ to return to point $A$ (measured at the clock at $A$) will be equal to the time $t_{\rightarrow}$ it takes for the signal to go around these points in the opposite direction: from $A$ through $C$ and $B$ back to $A$ (measured at the clock at $A$). This condition was considered by Reichenbach and called the roundabout axiom (Reichenbach [1928]).

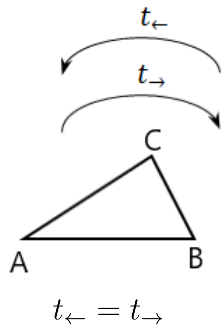



Figure 4: the roundabout axiom

It is easy to show (Reichenbach [1928], Macdonald [1983]) that, assuming time homogeneity of clocks, the circular isotropy of the synchronization signal is equivalent to the transitive property of Einstein synchronization. Thus, synchronization is also a transitive relation. So, we got that synchronization is an equivalence relation (reflexivity is trivial – each clock is synchronized with itself). Since we can synchronize all other clocks with one clock, this means that this equivalence relation has only one class, that is, that *we have a consistent synchronization of all clocks in the sense that every two clocks of an inertial frame are synchronized.*

We can now say that after the described synchronization procedure, all clocks of an inertial frame show the same time – the *global time of the inertial frame*. This time is an inertial time because it follows from the invariance of synchronization to the choice of a standardized particle or light for the synchronisation procedure that the time satisfies the Lange condition: a free particle travels the same distance at the same time.

Now that we have measures of space and time in an inertial frame, we can measure in it the (one-way) velocities of all free particles as well as the light produced in all possible ways.

Note that in this system of choice of units of space and time, a standardized free particle or standardized light has a velocity equal to 1 (both two-way and one-way velocity) – during time $t$ it travels the distance $t$. Of course, we can have another system of measurement of distances. If the system respects the symmetries of an inertial frame, we will get the same geometry. Only the unit of measurement of distance will be different. Such is, for example, the standard system of measurement with rigid rods.

We see that *Einstein synchronization of an inertial frame can be obtained without the use of light and so independently, without any circularity, the light postulate can be set.* Even if we choose one standardized light for the synchronization procedure (assuming that there is no ether, i.e., that the light is a closed system), this does not mean that every light, regardless of the conditions of its origin, has the same two-way speed as the standard light. The light principle, that every light has the same two-way speed $C$, is an additional postulate that goes beyond the concept of an inertial frame – it does not follow from it. The one-way light principle is then a consequence of the two-way light principle and clock synchronization. Of course, that speed is the same in all inertial frames. If we also use light to measure distance, this speed is equal to 1.

We see that a proper understanding of an inertial frame solves all the synchronization problems that arise in Einstein's article. A nice logical analysis of the synchronization problem and the role of light in it without assumptions of space and time symmetries can be found in Macdonald [1983], Minguzzi and Macdonald [2003].

Let us point out at the end that this concept of inertial frame says nothing about the direction of time. The existence of the direction of time is ubiquitous and inertial frames only inherits this property. Thus, this property is independent of the concept of inertial frame.

## 3.4 The principle of relativity

Since free particles move in an inertial frame uniformly in straight lines, and an inertial frame is a frame composed of free particles with a constant relative position, inertial frames also move uniformly in a straight line (in the sense that all its parts move uniformly in parallel straight lines while maintaining their mutual position) relative to a given inertial frame. Thus, all inertial frames are in relative uniform motion with each other. As already stated, the reversal is not valid. If a frame of reference moves uniformly in a straight line relative to an inertial frame, this does not mean that it is inertial, i.e., that its parts move freely. Such motion can also occur in the presence of external forces that are mutually in equilibrium. Their presence disturbs space and time symmetries of the frame and all previous analysis and the one that follows loses its basis. Here, this conception of inertial frame differs from the standard one according to which any frame that moves uniformly relative to an inertial frame is also an inertial frame.

The considerations we have applied to establish space and time symmetries of a closed physical process in an inertial frame, can also be applied to establish the possibility of the description in an inertial frame of a closed process that is invariant to the choice of an inertial frame. By definition, a closed system has no interaction with the environment, so the events in it are independent of an inertial frame from which we observe it. Since inertial frames do not differ in the way of observing a closed system then such a system can be described in a way invariant to the choice of an inertial frame. Thus, we can establish *the principle of relativity* of the description of physical processes in an inertial frame: *The laws of* *physics for a closed system are invariant to the transition from one inertial frame to another*. This principle, as well as the principles of space and time symmetries, derives not only from how physical processes take place but also from how we can describe them in an inertial frame. I repeat, we do not have to describe a closed physical process like that. But inertial frames give us the ability to describe them that way, and it is to be expected that such a description is the best in every respect. I think that with these considerations I have supplemented Geroch in Einstein et al. [2005], page 179, who says, among other things:

> The principle of relativity, then, hides within itself a subtle distinction—between what is and what is not taken as a law of physics. Indeed, it could be argued that a better perspective is to regard the principle of relativity, not as a general principle of nature at all, but rather as a guideline for distinguishing between those phenomena that are to be taken as "laws of physics" and those that are not. Phenomena that have the same description in every frame – that is, phenomena that are compatible with the principle of relativity – are to be accorded the status of physical laws, while phenomena that have different descriptions in different frames are to be regarded as merely specific phenomena. This is not a purely philosophical distinction: It can have consequences as to how physics is conducted.

It is not sufficiently known that the principle of relativity follows from space and time symmetries of inertial frames (see, for example, the proof in Rindler [2006], page 40). It follows directly from this result that the principle of relativity is founded in the same way as space and time symmetries.

# 4 Conclusion

Although the definition of the concept of inertial frame formulated here may seem insufficiently precise and "fragile", we see that it leads to a very robust and powerful properties of inertial frames. It ensures the existence of space and time symmetries of an inertial frame, as well as the principle of relativity. These symmetries together with the principle of relativity place certain restrictions on the possible laws of physics that guide us in finding them. Also, powerful conservation laws follow from them. Thanks to the symmetries of an inertial frame, all clocks are equivalent to each other (robustness!) – they define the same time up to a unit of measure. Symmetries also ensure time homogeneity of the clocks at various locations in an inertial frame, which is equivalent to the condition that once synchronized clocks remain synchronized. Also, symmetries ensure that we use free particles or light (assuming that once the light is emitted it is a closed system – no ether) to synchronize the clocks, and again in a robust way – no matter which procedure we choose to synchronize we will always get the same synchronization – Einstein clock synchronization. Likewise, no matter which procedure we choose to measure space, we always get the same space geometry – Euclidean geometry.

From the concept of inertial frame defined herein, we derived or at least made plausible all the enumerated properties of inertial frames with the following exceptions, which have been shown to remain clearly separated from the concept of inertial frame:

1. Although all inertial frames move uniformly with each other, it is not necessary that a frame of reference that moves uniformly relative to an inertial frame is also an inertial frame.
2. The speed of light in vacuum is the same in all inertial frames, regardless of how it is formed (the light principle).
3. Time has a direction.

# References

Y. Aharonov and T. Kaufherr. Quantum frames of reference. *Phys. Rev. D*, 30: 368–385, 1984.

R. Anderson, I. Vetharaniam, and G. Stedman. Conventionality of synchronization, gauge dependence and test theories of relativity. *Physics Reports*, 295: 93–180, 1998.

R. Angelo et al. Physics within a quantum reference frame. *Journal of Physics A: Mathematical and Theoretical*, 44(14):145304, 2011.

R. W. Brehme. On force and the inertial frame. *American Journal of Physics*, 53 (10):952–955, 1985.

B. Čulina. An Elementary System of Axioms for Euclidean Geometry Based on Symmetry Principles. *Axiomathes*, 2017.

R. DiSalle. Space and Time: Inertial Frames. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2020 edition, 2020.

A. Einstein. Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, 322(10): 891–921, 1905.

A. Einstein, R. Lawson, R. Geroch, and D. Cassidy. *Relativity: The Special and General Theory*. The masterpiece science edition. Pi Press, 2005.

D. Howard. What Makes a Classical Concept Classical? Toward a Reconstruction of Niels Bohr's Philosophy of Physics. In *Niels Bohr and Contemporary Philosophy*, pages 201–230. Kluwer Academic Publishers, 1994.

M. Jammer. *Concepts of Simultaneity: From Antiquity to Einstein and Beyond*. Johns Hopkins University Press, 2006.

A. Janis. Conventionality of Simultaneity. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition, 2018.

L. Landau and E. Lifshitz. *Mechanics*. Butterworth-Heinennan, third edition, 1976.

L. Lange. Über das beharrungsgesetz. *Berichte der Königlichen Sachsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-physische Classe*, 37:333–351, 1885.

L. Lange. On the law of inertia. *Eur. Phys. J. H.*, 39:251—-262, 2014. translation by Herbert Pfister of Lange [1885].

H. A. a. a. Lorentz. *The principle of relativity : a collection of original memoirs on the special and general theory of relativity*. Dover, 1952.

A. Macdonald. Clock synchronization, a universal light speed, and the terrestrial redshift experiment. *American Journal of Pysics*, 51:795–797, 1983.

D. Malament. Causal theories of time and the conventionality of simultaniety. *Noûs*, 11:293–300, 1977.

E. Minguzzi and A. Macdonald. Universal one-way light speed from a universal light speed over closed paths. *Foundations Of Physics Letters*, 16:593–604, 2003.

H. Minkowski. Raum und zeit. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 18:75–88, 1909.

C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W.H. Freeman and Co., 1973.

I. Newton, A. Motte, and N. W. Chittenden. *Newton's Principia: the mathematical principles of natural philosophy*. Daniel Adee, 1846.

R. Penrose. Structure of space-time. In C. DeWitt and J. Wheeler, editors, *Battelle Rencontres*, pages 121–235. Benjamin, 1968.

H. Reichenbach. *Philosophie der Raum-Zeit-Lehre*. Walter de Gruyter, 1928.

W. Rindler. *Relativity: Special, General, and Cosmological*. OUP Oxford, third edition, 2006.

S. Tanona. Idealization and Formalism in Bohr's Approach to Quantum Theory. *Philosophy of Science*, 71(5):683–695, 2004. doi: 10.1086/425233.

J. Thomson. On the law of inertia; the principle of chronometry; and the principle of absolute clinural rest, and of absolute rotation. *Proceedings of the Royal Society of Edinburgh*, 12:568–578, 1884.

Wikipedia contributors. Inertial frame of reference — Wikipedia, the free encyclopedia, 2021. [Online; accessed 18-March-2021].

# White Hole existence on the inverse universe

Shunji Mitsuyoshi[*], Eigo Shintani[†], Kosuke Tomonaga[‡], Yuichi Tei[§]

**Abstract**

The existence of White Hole (WH) has been suggested by Schwarzschild solution to the Einstein field equation as a time-reversed Black Hole (BH), besides there has not been observational evidence for their existence yet. Our idea of the "inverse universe", in which we introduce the time-reversed kinematics as another geometric state, can explain that WH should appear in such a geometry after a matter falls into a BH. In this work, we present a new operation for WH conversion from BH, and by using it the nearly infinity point on the universe, for instance the inside of BH, is geometrically connected to the inside of WH on the inverse universe. Such a conversion is useful to provide the simple solution to the problem of "information loss" in BH. Furthermore, we find another conversion point as the prior geometric state to the Big Bang, and we propose a new cosmology of cyclic universe.

**Keywords:** White hole; Black hole Information Loss; Dark Energy; Cyclic Universe.[**]

[*]Graduate School of Engineering, University of Tokyo, Hongo 7-3-1, Bunkyo, Tokyo, JAPAN;. mitsuyoshi@g.ecc.u-tokyo.ac.jp
[†]Graduate School of Engineering, University of Tokyo, Hongo 7-3-1, Bunkyo, Tokyo, JAPAN; eigo.shintani@gmail.com.
[‡]Graduate School of Engineering, University of Tokyo, Hongo 7-3-1, Bunkyo, Tokyo, JAPAN; tomonaga@bioeng.t.u-tokyo.ac.jp.
[§]Graduate School of Engineering, University of Tokyo, Hongo 7-3-1, Bunkyo, Tokyo, JAPAN; helixcm1@g.ecc.u-tokyo.ac.jp
[**] Received on September 15, 2022. Accepted on December 15, 2022. Published on December 15, 2022. DOI: 10.23756/sp.v10i1.869. ISSN 2282-7757; eISSN 2282-7765. 

# 1. Introduction

Black hole (BH) [1] has been known as the mysterious object; the general relativity predicts an object with extremely strong gravity that nobody can directly observe since a light can not escape from its inside (whose boundary is denoted as event horizon). We then expect BH existence from the observation of phenomena around BH, where strong gravitational force creates the extreme events, for instance the strong X-ray radiation from a matter captured by BH's gravity [2] [3]. The Schwarzschild solution [4], which is the exact solution of Einstein equation in the general relativity under a circumstance without rotation and charge, mathematically describe BH's geometry, but it does not explain a naïve question: "What happens to a matter fallen in the BH? Exactly speaking, where is matter going inside an event horizon?" [5] [6] [7]. Theoretically, inside an event horizon, any physical objects, even a light, are not allowed to go outside due to distorted geometry and eternally fall towards the singularity, in which the curvature becomes infinity. This implies that the kinematics behind an event horizon is compelled to an only direction to a singularity at timelike infinity [8]. On the other hand, BH will be eventually evaporated by Hawing radiation with the only light emission, and this is then consequent on a paradox, in which the initial information of matter, *e.g.* spin, is reduced to the three types of property, *i.e.* mass, electric charge and angular momentum, involving the Hawking radiation, seemingly the final matter violates a time-reversibility of the quantum mechanics.

Although a white hole (WH) [8] has been also theoretically predicted so far as the time-reversed BH solution of Einstein equation in which the physical object is compelled to outside of event horizon, no evidence of WH has been observed yet. We also come up with a question: "Why has a WH not been observed in our universe?". The general relativity is successful at the description of the gravitational geometry and establishes its validity with several observations and experiments below the Plank scale, except for a prediction of WH. This is also a cosmological issue.

In this article, we propose a new idea to address both issues: matter loss in BH and non-existence of WH, by introducing the "inverse universe" where its time-flow is reversed from that in our universe, like a mirror universe [9]. In our model, the matter fallen in BH is converted to anti-matter as a consequence of time-reversed operation, and it is then emitted from WH in the inverse universe. Therefore the total amount of matter including information is preserved in the whole universe, our universe and the inverse universe.

Furthermore considering another infinity point in the inverse universe as another conversion to the beginning of our universe, *i.e.* Big Bang, we can also provide the explanation of the fundamental issue of cosmology: "What is before Big-Bang?". Our cosmological model has an advantage to economically explain

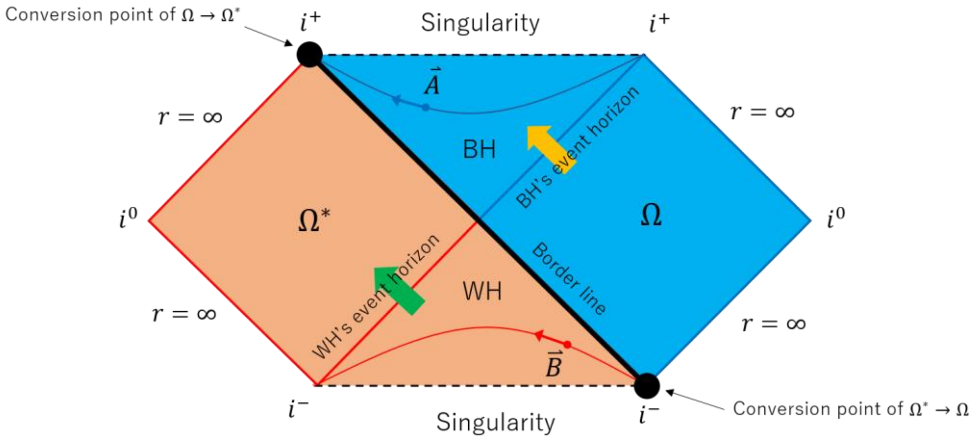



**Figure 1.** Penrose diagram, showing the maximal extension of Schwarzschild metric and border line between $\mathbf{\Omega}$ and $\mathbf{\Omega}^*$. There are two one-way directions such as from outside to inside of BH on $\mathbf{\Omega}$ and from inside to outside of WH on $\mathbf{\Omega}^*$ across the event horizon. In this metric, there are two singularities of BH and WH (top and bottom lines respectively). The timelike infinity point represents "$\mathbf{i}^+$" as end-point (infinite future) and "$\mathbf{i}^-$" as start-point (infinite past). There is a border line between $\mathbf{\Omega}$ and $\mathbf{\Omega}^*$ where the kinematical transformation is strictly restricted except its end-points at near infinity points, $\mathbf{i}^+$ and $\mathbf{i}^-$ called as conversion points explained in a text.

several cosmological problems over the other cyclic cosmology [10] [11] and mirror universe [9].

# 2. Conversion from Black Hole to White Hole

We first consider the kinematical transformation from $p$ to $q$ on the gravitational space-time (Minkowski geometry) Ω,

$$\Omega\colon p \Rightarrow q, \tag{1}$$

where physical object moves from outside to inside of BH's event horizon. This transformation happens in that physical object at timelike point $p$ on Schwarzschild metric is dragged by the gravitational force of BH into inside of Schwarzschild radius, in which the metric is changed as spacelike point $q$. According to the maximal extension of Schwarzschild metric (Kruskal-Szekeres coordinate [12] [13]), there is a region of negative time, defined as the extended geometry $\Omega^*$, wherein the kinematics advance to the past, taken by the time-reversed transformation. In a contrast to BH solution, for WH solution regarded as the time-reversed BH solution to Einstein equation, the kinematical transformation on $\Omega^*$ is

$$\Omega^*: q^* \Rightarrow p^*, \tag{2}$$

from inside $q^*$ to outside $p^*$ of WH's event horizon. The superposition of "*" represents state after the time-reversed operation is taken. In Figure 1, one can see that transformation $p \Rightarrow q$ shows the arrow crossing BH's event horizon on $\Omega$, and after that $q$ follows metric directing to timelike infinity $i^+$ inside BH. On the other hand, $q^* \Rightarrow p^*$ shows another arrow on $\Omega^*$ in which $q^*$ on WH follows metric directing to outside WH's event horizon.

No evidence of WH observation in our universe indicates that there is a border between $\Omega$ and $\Omega^*$ in which matter is mostly prohibited to cross its border, except the connection in nearly infinity points, $i^+$ and $i^-$(see below). This means the conversion from BH to WH happens inside BH's event horizon (see Figure 1), where we have no way to detect the existence of WH unless we enter the inside of BH's event horizon.

Here we also introduce the extremely large potential wall, but not infinity, in $\Omega$ and $\Omega^*$ boarder. Considering the quantum tunneling, due to the quantum fluctuation in the vacuum, the open threshold for the connection between $\Omega$ and $\Omega^*$ appears despite the extremely small probability. On the path to $i^+$ (triangle region of BH in Figure 1), however, this event should occur since inside the event horizon its time scale is infinitely extended. Observing the matter fallen in BH from the outside of event horizon, a matter seems to disappear, but in our model, whose information is transited into the inverse universe through WH.

Such an idea is useful for the solution to the problem of disappearance of matter and its information inside BH; we can interpret that those are taken by the geometrical transformation from $\Omega$ to $\Omega^*$ and information loss is independent of BH evaporation. As a result, total amount of matter and its information should be conserved on the two geometries, $\Omega$ and $\Omega^*$.

In the above, we consider the conversion from BH to WH, which is the connection of geometry at the nearly infinity point from $i^+$ on $\Omega$ to $i^-$ on $\Omega^*$ showing the left-upper corner to right-bottom corner of Penrose diagram (see

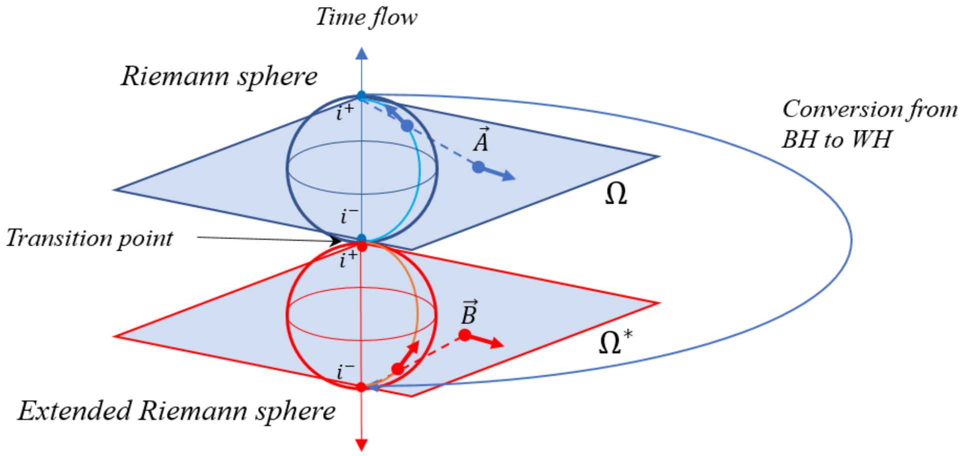



**Figure 2.** Riemann sphere and extended Riemann sphere. North pole of Riemann and extended Riemann spheres represent infinity point $\mathbf{i}^+$, and south pole is $\mathbf{i}^-$. The conversion from BH to WH can connect the north pole of Riemann sphere and south pole of the extended Riemann sphere. The time-flow showing vertical arrow has opposite direction since the extended Riemann sphere is a consequence of time reversed. The vector $\vec{\mathbf{A}}$ and $\vec{\mathbf{B}}$ represent kinematical transformations of physical object on geometry of $\mathbf{\Omega}$ and $\mathbf{\Omega}^*$, and those points can be projected onto Riemann and extended Riemann sphere. Assuming that $\vec{\mathbf{A}}$ is located inside the event horizon of BH, the corresponding transformation $\vec{\mathbf{B}}$, which is inside event horizon of WH appears through the conversion. The connected point of both Riemann spheres is transition point from the Big Crunch to the Big Bang.

Figure 1). The other side, *i.e.* the connection from $i^-$ on $\Omega^*$ to $i^+$ on $\Omega$, then regards a geometrical transformation from the end of the inverse universe (Big Crunch) to the beginning of universe (Big Bang). This is considered of a sort of the cyclic cosmology model [10] [11]. We discuss more details later.

# 3. Representation of the extended Riemann Sphere

Next we discuss the representation of gravitational space-time $\Omega$ with a complex plane, and define it in Riemann sphere by the stereographic projection; now the north pole corresponds to the infinity point $i^+$ (see Figure 2). In this picture, we can also figure out that there is a reversed sphere defined as the extended Riemann sphere [14] [15][††], which consists with the geometry of the inverse universe $\Omega^*$. So that the conversion from BH to WH can be interpreted

[††] We used a nomenclature "converse" in those references instead of "inverse".

as that the south pole on the extended Riemann sphere is connected to the north pole on the Riemann sphere. Consequently, the kinematics on the extended Riemann sphere regards the time-reversed kinematics on the Riemann sphere. Here we define "time-flow" on $\Omega$ as a time sequence, and it is then opposite flowing sequence on $\Omega^*$; the time-flow towards $i^+$ on the Riemann sphere corresponds to the time-flow towards $i^-$ on the extended Riemann sphere (see Figure 2). Considering the kinematics on those spheres, the directions are same, *i.e.* both vectors $\vec{A}$ and $\vec{B}$ in Figure 2 are transformed from south to north on surface of both spheres, however $\vec{B}$ is observed as opposite kinematics from $\vec{A}$ since the time-flow is opposite.

The path to enter the inverse universe is a way with conversion from BH to WH, and in this way the matter including energy is converted into the anti-matter in the extended Riemann sphere, which is a consequence of time-reversed operation; the particles are changed to the anti-particles, for instance, proton and electron are converted to anti-proton and positron respectively. So that, while in our universe the particles are dominated, in the inverse universe anti-particles are dominated, namely anti-hydrogen composed by positron and anti-proton is a fundamental atom. Furthermore the time-reversed operation for gravitational force makes a change of attractive force to the repulsive force, and it then results that the star can not be constructed in the inverse universe and particles are afloat without any cluster. The important ingredient is the dark energy [16] [17], which is currently considered as 70% dominant repulsive force in our universe to explain the acceleratingly expanding universe observed from *e.g.* type Ia supernova [18] [19], expected to be a stem of the cosmological constant, but have not been directly observed yet, is changed to interactive force in the inverse universe (here it is called "anti-dark energy") through BH/WH conversion. In the inverse universe, anti-dark energy plays an important role of shrinking the space-time to transition point, which is a cross point between Riemann and extended Riemann spheres in Figure 2. This ends up with the Big Crunch in the inverse universe. Our model of cosmology will be discussed in the next section.

# 4. Cosmology of the inverse universe

Now we can discuss a cosmology of the inverse universe. Starting from soon after the Big Bang in our universe, the first BH was created by collapsed massive star by the gravitational force and its amount increased following the time-flow. The matter including dark energy in Riemann sphere $\Omega$ is transformed into the extended Riemann sphere $\Omega^*$ through BH/WH conversion. Increasing the number of BH, the amount of matter and dark energy decreases in $\Omega$, and the speed of expansion of our universe will be reduced since the ingredient to control the scale of universe decreases. At the end, our universe will be totally

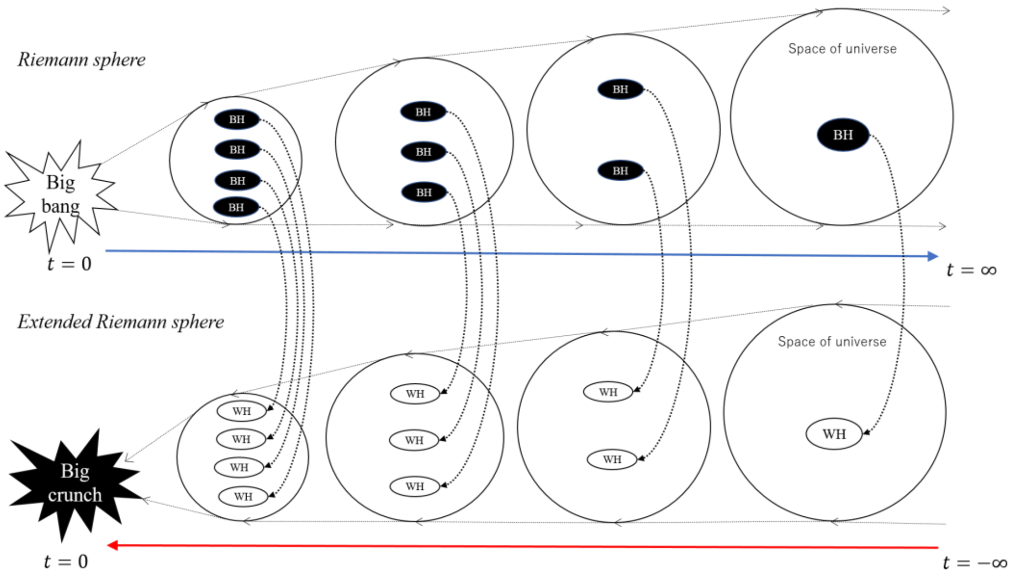



**Figure 3.** Time-flow of the universe and the inverse universe. On Ω, the BHs increasingly appear in the universe since the matter is rich. The space size has also been expanded by the dark energy. Expansion speed has been reduced because the amount of dark energy has also decreased due to increase of BHs. The matter and energy fallen in BHs have been transited onto $\Omega^*$ via WHs, and the amount of anti-matter on $\Omega^*$ has increased. The BHs are collapsed and disappear for a while, and eventually whole energy is transited onto $\Omega^*$ and nothing (exactly speaking a light appeared in BH evaporation is remnant) in the universe after all. On the other hand, on $\Omega^*$, first there is only space, and once the WHs appear, energy increases and space is shrunk by anti-dark energy. Eventually the universe ends up to Big Crunch.

filled by BH, and eventually the universe becomes void space, where a light created by Hawking radiation of BH evaporation [20] remains.

Since the time-flow of the extended Riemann sphere is reversed from that of Riemann sphere, in the inverse universe a WH first appears in void space, and anti-matter and anti-dark energy are emitted from WH. Here the universe shrinks due to the dominance of attractive force of anti-dark energy, and time-flow is then directed to the Big Crunch (see Figure 3). The matter including information fallen in the BH is emitted from the WH, and following time-flow, WH's size becomes small, and eventually disappears by Hawking radiation (see Figure 4).

Once the Big Crunch happens, assuming the transition from Big Crunch to Big Bang at the nearly infinity point $i^+$ in the inverse universe, anti-matter and anti-dark energy are transformed to normal ones as in our universe. Back to

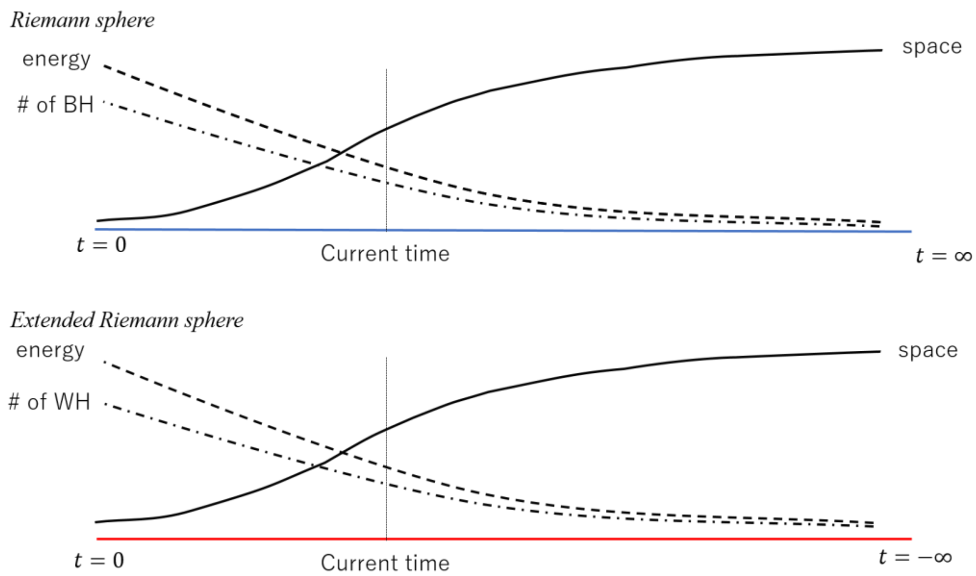



**Figure 4.** The amount of energy (or information), BH (top) and WH (bottom), and size of space on $\mathbf{\Omega}$ (top) and $\mathbf{\Omega}^*$ (bottom) as a function of time flow. Total amount of energy is invariant, while, on each side, the energy flows from $\mathbf{\Omega}$ to $\mathbf{\Omega}^*$ through BH and WH. A space of $\mathbf{\Omega}$ has been initially spread since repulsive force of dark energy is rich, but its speed is gradually reduced since the dark energy has been transformed by BH/WH conversion. On the other hand, a space of $\mathbf{\Omega}^*$ is decreased since an attractive force of anti-dark energy has been increase through BH. Current out position is relatively close to Big Bang, and the universe is still expanding due to rich dark energy.

Figure 2, one can see that the kinematics on the extended Riemann sphere approach infinity $i^+$ at the north pole, and through the transition point, the kinematics start from the south pole on Riemann sphere. This model provides the new cyclic cosmology. From the perspective of Riemann and the extended Riemann sphere, the matter circulates in two universes with different time-flow through the polarity of Riemann spheres.

In our model, total amount of matter is invariant on $\Omega + \Omega^*$; however individually it is changed along the time-flow. As shown in Figure 4, matter on $\Omega$ is transformed to $\Omega^*$ through BH/WH conversion, and it then decreases on $\Omega$ and increases on $\Omega^*$. The size of space, which is affected by the attractive force and repulsive force of matter (anti-dark energy) and dark energy (anti-matter) in our universe (inverse universe), respectively, is consistently changed on both $\Omega$ and $\Omega^*$. On $\Omega$, initially there filled the repulsive force of dark energy in the early universe and space was quickly spread. The number of BH increases after collapsed massive star, and simultaneously increasing the amount of matter and dark energy fallen in BH, the energy is reduced in $\Omega$. So that the size of space is asymptotically close to a constant when the amount of matter and dark energy in $\Omega$ is reduced to be zero. On $\Omega^*$, initially a WH appears and repulsive force of anti-gravitational force of anti-matter is filled in the inverse universe, and it then spreads a space. Increasing anti-dark energy which comes through BH in $\Omega$, the attractive force increases, while WHs disappear due to evaporating its energy. As a consequence, a space starts shrinking and disappears in Big Crunch, and its matter backs to a normal one starting from Big Bang.

# 5 Discussion and summary

In this article we propose the new idea to solve the cosmological problems, BH information loss, WH non-existence and before Big Bang, by introducing the inverse universe and BH/WH conversion. Our idea is that the matter behind the event horizon of BH is directed to the infinity point in which the time-reversed operation is taken and converted to the anti-matter, eventually emitted from WH in the inverse universe. Now the information in BH is just transformed onto that in the inverse universe via WH. In our model, WH should exist in the inverse universe as a pair of BH, and matters fallen in BH are converted to anti-matter in the inverse universe. Considering the representation of our model with Riemann sphere, we also derive a new cyclic cosmology in which Big Crunch in the inverse universe is connected to Big Bang in our universe at the nearly infinity point. Here the anti-dark energy plays an important role to shrink the universe with its attractive force, as an opposite role of dark energy. We comment that our solution to the BH information loss is retained in classical gravity and no assumption of unitarity violation of quantum mechanics. Behind the event horizon of BH, the reversibility of initial and final state is regarded as the BH/WH conversion, and ensemble of final state is transformed into the inverse universe. In addition, we argue that for its solution it is possible to separate between quantum effect and gravity, namely such an issue is nothing to do with the argument of holography [21].

Here one also has a question: “How can we experimentally observe the inverse universe?”. We consider such an observation will be the new phenomena of breaking time-reversal symmetry in our universe. In Figure 1, we assume a hard border of geometry between $\Omega$ and $\Omega^*$, however taking into account a quantum effect, besides extremely small probability, the extra particle possibly appears from the inverse universe by penetrating through such a hard border (like a quantum tunneling) even in the outside of event horizon. Such a phenomenon can be seen as asymmetry of particle and anti-particle in our universe. Note that to observe it, we experimentally need a high energy insertion to amplify the probability to access the inverse universe.

Now we propose that a reason of “why is particle richer in our universe than anti-particle?” (known as a baryon asymmetry problem [22]) is an emergence of the extra particle from the inverse universe due to quantum effect in the early universe. At the beginning of Big Bang, the number of particle and anti-particle was exactly same and annihilated each other, however its symmetry was accidentally broken by extra particle of the inverse universe. At last, after expanding and cooling down our universe, such an extra particle remains and constitutes stars and us (such an accidental event simultaneously happened and oppositely anti-particle became richer in the inverse universe.). In this scenario, an appearance of extra particle should be detected in the search of time-reversal symmetry breaking (same as CP violation) in the microscopic scale, *e.g.* non-zero electric dipole moment of particle [23]. Although there have been many models to explain the time-reversal symmetry breaking in high energy physics [24], those models need to assume the new particles, for instance supersymmetry particles, despite no evidence of those detection. On the other hand, our scenario does not assume any new particles even in high energy scale (over TeV scale), and simply such a breaking will occur within the framework of the standard model of particle physics. In a current situation of no discovery of the new particle beyond the standard model even in Large Hadron Collider experiment over 1 TeV energy [25], our scenario is rather reasonable compared to high energy models.

# References

[1] Misner, C. W., Thorne, Kip S., Wheeler, J. A., “Gravitation.” W. H. Freeman and Company, 1973.

[2] Webster, B. Louise, Murdin, P., “Cygnus X-1—a Spectroscopic Binary with a Heavy Companion?” Nature, 235 (5332), 37–38, https://doi.org/10.1038%2F235037a0, 1972.

[3] Bolton, C. T. , “Identification of Cygnus X-1 with HDE 226868.” Nature, 235 (5336), 271–273, https://doi.org/10.1038%2F235271b0, 1972.

[4] Schwarzchild, K., "Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie." Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften. 7, 189–196, 1916.

[5] Hawking S. W., “Breakdown of predictability in gravitational collapse.” Physical Review D 14, 2460, https://doi.org/10.1103/PhysRevD.14.2460, 1976.

[6] Hawking, S. W., “Information loss in black holes.” Physical Review D. 72

(8), 084013, https://doi.org/10.1103%2FPhysRevD.72.084013, 2005.
[7] Barbón, J. L. F., "Black holes, information and holography." Journal of Physics: Conference Series. 171 (1), 012009, https://doi.org/10.1088%2F1742-6596%2F171%2F1%2F012009, 2009.
[8] Hawking, S. W., Ellis, G. F. R., "The Large Scale Structure of Space-Time." Cambridge University Press, , 1975.
[9] Boyle, L., Finn, K., Turok, N., "CPT-Symmetric Universe." Physical Review Letters. 121 (25), 251301, https://doi.org/10.1103%2FPhysRevLett.121.251301, 2018.
[10] Penrose, R., "Before the Big Bang: An Outrageous New Perspective and its Implications for Particle Physics." in Proceedings of the EPAC 2006, 2759–2762, 2006.
[11] Gurzadyan, VG., Penrose, R., "On CCC-predicted concentric low-variance circles in the CMB sky." Eur. Phys. J. Plus. 128 (2), 22, https://doi.org/10.1140%2Fepjp%2Fi2013-13022-4, 2013.
[12] Kruskal, M. D., "Maximal extension of Schwarzschild metric." Physical Review. 119 (5), 1743–1745, https://doi.org/10.1103%2FPhysRev.119.1743, 1960.
[13] Szekeres, G. , "On the singularities of a Riemannian manifold." Publicationes Mathematicae Debrecen. 7, 285, https://doi.org/10.1023%2FA%3A1020744914721, 1960.
[14] Mitsuyoshi, M., Tomonaga, K., Hashimoto, M. and Tei, Y., "Utilization in Artificial Ego of Quantum Calculations Based on Anti-Einstein Field Hypothesis." DHU JPOURNAL Vol.8, 2021.
[15] Mitsuyoshi, S., Tomonaga, K., Hashimoto, M., Tei, Y. and Nakamura, T., "Some hypothesis to derive an anti-Einstein field." DHU JOURNAL Vol.6, 2019.
[16] Ellis, G. F. R., "Dark energy and inhomogeneity." Journal of Physics, Conference Series. 189, 012011, https://doi.org/10.1088%2F1742-6596%2F189%2F1%2F012011, 2009.
[17] Colin, J., Mohayaee, R., Rameez, M., Sarkar S., "Evidence for anisotropy of cosmic acceleration." Astronomy and Astrophysics. 631, L13, https://doi.org/10.1051%2F0004-6361%2F201936373, 2019.
[18] Perlmutter, S., Aldering, G., Valle, M. D., Deustua, S., Ellis, R.S., Fabbro, S., et al., "Discovery of a supernova explosion at half the age of the Universe." Nature. 391 (6662), 51–54, https://doi.org/10.1038%2F34124, 1998.
[19] Schmidt, B.P., Suntzeff, N.B., Phillips, M.M., Schommer, R.A., Clocchiatti, A., Kirshner, R.P., "The High-Z Supernova Search: Measuring

Cosmic Deceleration and Global Curvature of the Universe Using Type Ia Supernovae.” The Astrophysical Journal. 507 (1), 46–63, https://doi.org/10.1086%2F306308, 1998.
[20] Hawking, S. W., “Black hole explosions?” Nature. 248 (5443), 30–31, https://doi.org/10.1038%2F248030a0, 1974.
[21] Raju S., “Lessons from the information paradox.” Physics Reports 943, 1–80, https://doi.org/10.1016/j.physrep.2021.10.001, 2022.
[22] Sakharov, A. D., “Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe.” Pisma Zh.Eksp.Teor.Fiz. 5, 32-35, https://doi.org/10.1070/PU1991v034n05ABEH002497, 1967.
[23] Abel, C., Afach, S., Ayres, N. J., Baker, C. A., Ban, G., Bison, G., “Measurement of the Permanent Electric Dipole Moment of the Neutron.” Physical Review Letters. 124 (8), 081803, https://doi.org/10.1103%2FPhysRevLett.124.081803, 2020.
[24] Ibrahim, T., Itani, A., Nath, P., “Electron EDM as a Sensitive Probe of PeV Scale Physics.” Physical Review D. 90 (5), 055006, https://doi.org/10.1103%2FPhysRevD.90.055006, 2014.
[25] Aad, G., Abbott, B., Abdallah, J., Abdelalim, A. A., Abdesselam, A., Abdinov O., et al., “Search for Supersymmetry Using Final States with One Lepton, Jets, and Missing Transverse Momentum with the ATLAS Detector in $\sqrt{s} = 7$ TeV pp Collisions.” Physical Review Letters. 106 (13), 131802, https://doi.org/10.1103%2FPhysRevLett.106.131802, 2011.

# The authorship of the Principle of Inertia (La paternità del Principio d'Inerzia)

Luca Nicotra[1]

Parte II

La prima parte è pubblicata in «Science & Philosophy» Volume 10(1), 2022

**Abstract**

According to some currents of modern historiography, Galilei's propensity for circular motion would have led him to consider this and not rectilinear motion as "natural motion"; therefore the principle of inertia could not be fully attributed to Galileo, which he would never have formulated. The question of the authorship of the principle of inertia certainly weighs on both nationalistic elements and returns of antigaleleism, while the question of its not explicit formulation as a principle is due to ignorance of the type of organization that Galileo intended to give to the exposition of his physics. The author, after having hinted at possible prodromes of the principle of inertia and having reported the adverse opinions of illustrious historians of science (A. Koyré, I. B. Cohen, P. M. Duhem, P. Rossi, G. Holton), through a careful analysis of the Galilean writings, conducted on the digital versions with the help of text analysis programs, firmly reaffirms Galileo's authorship of the principle of inertia and the consequent principle of classical relativity.
**Keywords:** Inertia, natural motion, Principle of Classical Relativity.[2]

**Sunto**

Secondo alcune correnti della storiografia moderna, la propensione di Galilei per il moto circolare lo avrebbe portato a ritenere come "moto naturale" questo e non il moto rettilineo; quindi a Galileo non si potrebbe attribuire pienamente il

[1]Ingegnere meccanico. Presidente dell'A. P. S. "Arte e Scienza", Direttore responsabile di: «Periodico di Matematica», «Bollettino dell'Accademia di Filosofia delle Scienze Umane», «ArteScienza»; Rome, Italy; luca.nicotra1949@gmail.com.
[2]Received on November 28, 2022. Accepted on December 20, 2022. Published on December 30, 2022. doi: 10.23756/sp.v10i1.966. ISSN 2282-7757; eISSN 2282-7765. 

Principio d’Inerzia, che non avrebbe nemmeno mai formulato. Sulla questione della paternità del Principio d’Inerzia gravano certamente sia elementi nazionalistici sia ritorni di antigaleleismo, mentre la questione di una sua non esplicita formulazione come principio è dovuta all’ignoranza del tipo di organizzazione che Galileo intendeva dare all’esposizione della sua fisica. L’autore, dopo aver accennato a possibili prodromi del Principio d’Inerzia e aver riportato le avverse opinioni di illustri storici della scienza (A. Koyré, I. B. Cohen, P. M. Duhem, P. Rossi, G. Holton), attraverso un’attenta analisi degli scritti galileiani, condotta sulle versioni digitali con l’aiuto di programmi di analisi del testo, riafferma con decisione la paternità di Galileo del Principio d’Inerzia e del conseguente Principio di Relatività Classica.
**Parole chiave:** Inerzia, moto naturale, Principio di Relatività Classica.

# 1. Altre formulazioni del Principio d’Inerzia

Riprendiamo l’incipit della Giornata Quarta dei *Discorsi e dimostrazioni matematiche intorno a due nuove scienze attinenti alla meccanica ed i movimenti locali,*[3] considerato la formulazione più completa del Principio d’Inerzia da parte di Galilei,[4] essendo chiaramente presenti i suoi due elementi caratteristici: la rettilineità del moto e la velocità costante (moto equabile e perpetuo):

> *Mobile quoddam super planum horizontale proiectum mente concipio, omni secluso impedimento: iam constat, ex his quae fusius alibi dicta sunt, illius motum aequabilem et perpetuum super ipso plano futurum esse, si planum in infinitum extendatur; si vero terminatum et in sublimi positum intelligamus, mobile, quod gravitate praeditum concipio, ad plani terminum delatum, ulterius progrediens, aequabili atque indelebili priori lationi superaddet illam quam a propria gravitate habet deorsum propensionem, indeque motus quidam emerget compositus ex aequabili horizontali et ex deorsum naturaliter accelerato, quem proiectionem voco.*

In esso è chiaro il riferimento ad altre volte in cui “più diffusamente” («*ex his quae fusius alibi dicta sunt*») Galileo ha trattato lo stesso argomento, ovvero c’è l’allusione ad altre “applicazioni” o “formulazioni operative”, nelle sue opere, del Principio d’Inerzia. Si tenga presente che quella appena citata

[3] Nel seguito: *Discorsi.*

[4] Si tenga sempre presente quanto detto nella Prima Parte del presente scritto a proposito delle formulazioni “operative” del Principio d’Inerzia da parte di Galileo (Nicotra. 2022, pp. 85-86; Drago, 1997) ben diverse da formulazioni teoretiche quali quella data da Isaac Newton.

compare nei *Discorsi*, ultima opera di Galilei, pubblicata nel 1638. In effetti, è possibile trovare, in altre opere di Galileo, altri riferimenti al Principio d'Inerzia (Caffarelli, 2006), tuttavia non tutti ugualmente chiari e completi.

Oltre quello sopra ricordato, il più chiaro è contenuto nella Giornata Sesta dei *Discorsi* (Giornata aggiunta: *Della Forza della Percossa*[5]). Qui Galileo utilizza un apparato meccanico[6] in cui due masse uguali, sospese a una carrucola, si equilibrano, permettendo di realizzare un moto verticale con velocità costante (Favaro, 1890-1909, Vol. VIII, *Discorsi*, pp. 332-333):

> *... voglio che ci figuriamo un solido grave, per esempio di mille libre di peso, il quale posi sopra un piano che lo sostenti; voglio poi che intendiamo una corda a cotal solido legata, la quale cavalchi sopra una carrucola fermata in alto, per buono spazio, sopra detto solido. Qui è manifesto, che aggiungendo forza traente in giù all'altro capo della corda, nel sollevar quel peso si averà sempre una egualissima resistenza, cioè il contrasto di mille libre di gravità; e* ***quando da quest'altro capo si sospenda un altro solido egualmente pesante come il primo, verrà da essi fatto equilibrio****; e* ***stando sollevati****, senza che sopra alcuno sottoposto sostegno si appoggino,* ***staranno fermi, né scenderà questo secondo grave alzando il primo, salvo che quando egli abbia qualche eccesso di gravità.***

E più avanti ecco una nuova "formulazione" del Principio d'Inerzia, altrettanto importante perché afferma la sua validità anche nel caso del moto verticale, ponendo in evidenza i due casi di quiete e di moto rettilineo uniforme conseguenti all'annullamento della risultante delle forze esterne. Se già nella Giornata Quarta è evidente l'inerzia rettilinea, fuori di ogni dubbio lo è nella Giornata Sesta, dove l'unico moto preso in considerazione è quello rettilineo lungo i due bracci della carrucola (Favaro, 1890-1909, Vol. VIII, *Discorsi*, p. 336-337):

---

[5] Per molto tempo l'opera *Della Forza della Percossa* rimase inedita. L'ultimo e devoto discepolo di Galilei, Vincenzo Viviani (1622-1703) riuscì a copiarne un manoscritto posseduto dal figlio di Galileo, Vincenzio, in cui riconobbe la mano di Marco Ambrogetti, il sacerdote fiorentino di cui lo Scienziato pisano si servì come amanuense, per la dettatura delle sue lettere durante la reclusione ad Arcetri. Tuttavia Viviani non ritenne di inserirlo nella raccolta delle *Opere di Galileo Galilei* da lui curate per l'editore fiorentino Carlo Manolessi a Bologna nel 1655-56. Soltanto nell'edizione fiorentina delle *Opere* di Galileo del 1718 quello studio di Galilei sulla dinamica degli urti verrà inserito, come Sesta Giornata dei *Discorsi*, con il titolo *Della Forza della Percossa*. Gli esperimenti descritti da Galileo probabilmente furono eseguiti a Padova prima del 1610, ma Galileo scrisse la Sesta Giornata nel 1638. Alcuni esperimenti, però, potrebbero essere stati eseguiti ad Arcetri nel 1637-1639 (Caffarelli, 2010a).

[6] Detto "Macchina di Galileo", di cui 150 anni dopo fu costruita una versione simile da George Atwood nel 1784 (Macchina di Atwood).

> *E qui mi pare che accada per appunto quello che accade ad un mobile grave e perfettamente rotondo, il quale, se si porrà sopra un piano pulitissimo ed alquanto inclinato, da per sé stesso naturalmente vi scenderà, acquistando sempre velocità maggiore; ma se, per l'opposito, dalla parte bassa si vorrà quello cacciare in su, ci bisognerà conferirgli impeto, il quale si anderà sempre diminuendo e finalmente annichilando; ma* ***se il piano non sarà inclinato, ma orizontale, tal solido rotondo, postovi sopra, farà quello che piacerà a noi, cioè, se ve lo metteremo in quiete, in quiete si conserverà****,* ***e dandogli impeto verso qualche parte, verso quela si moverà, conservando sempre l'istessa velocità che dalla nostra mano averà ricevuta****, non avendo azione né di accrescerla né di scemarla, non essendo in tal piano né declività né acclività;* ***et in simile guisa i due pesi eguali, pendenti da' due capi della corda, ponendogliene in bilancio, si quieteranno, e se ad uno si darà impeto all'in giù, quello si andrà conservando equabile sempre.*** *E qui si dee avvertire che* ***tutte queste cose seguirebbero quando si movessero tutti gli esterni ed accidentari impedimenti****, dico di asprezza e gravità di corda, di girelle e di stropicciamenti nel volgersi intorno al suo asse, ed altri che ve ne potessero essere.*

Alessandre Koyré (1939, *Galilée et la loi de l'inertie*, parte III) contesta a Galileo di considerare sempre come forza soltanto la gravità, a differenza di Newton che ha saputo generalizzare il concetto di forza. Giustamente Roberto Vergara Caffarelli obietta: «Ma quali erano le altre forze in Newton, se non l'attrito nelle sue varie forme?» (Caffarelli, 2006). D'altra parte, quali altre forze a quell'epoca avrebbe potuto considerare Galilei?

Galileo in *Della Forza della Percossa* illustra ben quattro diversi esperimenti che si possono effettuare con la sua "Macchina", come sintetizza Vergara Caffarelli (2006):[7]

> *a) si dà una spinta a un peso:* ***i due pesi si muovono insieme in moto uniforme con la velocità impressa con la spinta.*** *Così Galileo dimostra la validità sperimentale del principio d'inerzia anche quando il moto è verticale, perché - spiega- è nulla la somma delle forze agenti su ogni massa;*
>
> *b) si appoggia uno dei due pesi uguali sopra un banchetto, si solleva l'altro ad un'altezza prefissata e poi lo si lascia cadere: quando la corda diventa tesa, si ha - come la chiama Galileo - una «strappata». L'altro peso, tirato dalla corda, impedisce al primo di continuare il moto accelerato.* ***I due pesi, insieme, procedono con moto uniforme, come nell'esperimento precedente****;*

[7] Il Principio di Inerzia è stato dimostrato sperimentalmente il 7 novembre 2005 con la Macchina di Galileo ricostruita da Roberto Vergara Caffarelli e collaboratori dell'Istituto Nazionale di Fisica Nucleare - sezione di Pisa (Caffarelli, 2010b).

> *c) si aggiunge un peso alla massa di quello appoggiato, si solleva l'altro peso e lo si lascia cadere, come prima: quando la corda diventa tesa, il peso maggiore sale fino ad un'altezza, che dipende dal peso aggiunto e dalla velocità massima raggiunta in caduta libera dal peso sollevato. Il loro moto, dopo la «strappata», è uniformemente ritardato, fino all'altezza massima, e poi accelerato in senso inverso.Calcolando la velocità raggiunta dal peso che cade, immediatamente prima della «strappata», e misurando la velocità dei due pesi, subito dopo la «strappata», **si può verificare la conservazione della quantità di moto**. Galileo non riuscì a dimostrare questa legge, che pure è ricordata spesso nei suoi scritti in varie maniere, perché, non potendo misurare la velocità massima raggiunta prima della strappata, fa ricorso ad un risultato teorico, che non è valido in quella situazione;*
>
> *d) Galileo aggiunge una piccola massa ad una delle due, porta questa massa in alto e la lascia cadere. Osserva un moto accelerato, ma molto lento perché c'è solo il peso della massa aggiunta a muovere tutte le masse: **prima chiara evidenza della differenza tra massa inerziale e massa gravitazionale**.*

Altri accenni (Caffarelli, 2006) al Principio d'Inerzia, ma molto meno espliciti e completi, sono presenti sia nel *De motu* (circa 1590) sia nel trattatello delle *Mecaniche* (1593).

Molto esplicito, invece, quello contenuto Nella *Istoria e dimostrazioni intorno alle macchie solari e loro accidenti* del 1613 (Favaro, 1890-1909, Vol. V, pp. 134-135):

> ***Imperò che mi par di osservare che i corpi naturali abbino naturale inclinazione a qualche moto**, come i gravi al basso, **il qual movimento vien da loro, per intrinseco principio e senza bisogno di particolar motore esterno, esercitato, qual volta non restino da qualche ostacolo impediti;** a qualche altro movimento hanno ripugnanza, come i medesimi gravi al moto in su, e però già mai non si moveranno in cotal guisa, se non cacciati violentemente da motore esterno; **finalmente, ad alcuni movimenti si trovano indifferenti,** come pur gl'istessi gravi al movimento orizontale, al quale non hanno inclinazione, poi che ei non è verso il centro della Terra, né repugnanza, non si allontanando dal medesimo centro: e però, rimossi tutti gl'impedimenti esterni, un grave nella superficie sferica, e concentrica alla Terra sarà indifferente alla quiete, ed ai movimenti verso qualunque parte dell'orizonte; ed in quello stato si conserverà, nel qual una volta sarà stato posto; cioè **se sarà messo in stato di quiete, quello conserverà, e se sarà posto in movimento, v.g. verso occidente, nell'istesso si manterrà**: e così **una nave per essempio avendo una sol volta ricevuto qualche impeto per il mar tranquillo, si moverebbe continuamente intorno al nostro globo senza cessar mai, e postavi con quiete, perpetuamente quieterebbe, se***

> ***nel primo caso si potessero rimuovere tutti gl'impedimenti estrinseci, e nel secondo qualche causa motrice esterna non gli sopraggiungesse.***

In quest'ultimo caso, nell'esempio della nave, effettivamente si potrebbe parlare di una "inerzia circolare": manca il moto rettilineo del Principio d'Inerzia, ma rimane affermata la costanza della velocità scalare in assenza di una forza netta.[8] Una "imprecisione" che però è ravvisabile anche nell'opera celeberrima di Isaac Newton (1687, 2001), che subito dopo aver formulato la *Lex I* (*Axiomata Sive Leges Motus*), ovvero il Principio d'Inerzia, così lo spiega con esempi:

> *Projectilia perseverant in motibus suis nisi quatenus a resistentia aeris retardantur & vi gravitatis impelluntur deorsum. Trochus,* [9] *cujus partes cohærendo perpetuo retrahunt sese a motibus rectilineis, non cessat rotari nisi quatenus ab aere retardatur. Majora autem Planetarum & Cometarum corpora motus suos & progressivos & circulares in spatiis minus resistentibus factos conservant diutius.*
>
> *(I proiettili persevererebbero nei loro moti, se la resistenza dell'aria non li ritardasse, e la forza di gravità non li spingesse verso il basso. Un cerchio le cui parti correndo perpetuamete si ritraggono l'un l'altra dal moto rettilineo, non cesserebbe di ruotare, se l'aria non lo ritardasse alquanto. I Pianeti e le Comete – che sono corpi più grandi muoventisi in spazi meno resistenti – conservano per maggior tempo i loro moti progressivi o circolari).*

Gli ultimi due esempi (il gioco del Trochus e le orbite planetarie) sono di moto circolare per illustrare il Principio d'Inerzia! Come mai non si è mai parlato di una "inerzia circolare" per Newton?

Nel *Dialogo sopra i due massimi sistemi del mondo tolemaico e copernicano* (1632)[10] c'è un altro riferimento incompleto al Principio d'Inerzia. Ritroviamo l'esempio, più volte utilizzato da Galileo, di un mobile posto su due diversi piani: uno inclinato e l'altro non inclinato. Quando il mobile è sul piano inclinato si muove di moto uniformemente accelerato verso il basso per effetto della forza peso, mentre per farlo risalire occorre applicare una forza che lo farà muovere di moto uniformemente ritardato. A questo punto Salviati chiede a Simplicio cosa accadrebbe se invece il mobile fosse posto su una superficie non inclinata: «*Ora ditemi quel che accaderebbe del*

[8] Si potrebbe, però, parlare ancora di moto rettilineo intrinseco riferito allo spazio bidimensionale delle superficie sferica: la retta in tale spazio è vista come circonferenza (o suo arco) nello spazio tridimensionale estrinseco.

[9] Il *trochus*, tradotto in "cerchio", cui allude Newton era un giocattolo (di cui parla anche Ovidio) simile ai cerchi utilizzati dai bambini ai nostri tempi.

[10] Nel seguito: *Dialogo*.

*medesimo mobile sopra una superficie che non fusse né acclive né declive*». Da notare che mentre prima Galileo, per bocca di Salviati, parla di un "piano" inclinato ora invece parla di una "superficie" non inclinata (*né acclive né declive*). E la superficie non inclinata che porta a far considerare a Simplicio è la superficie sferica della Terra e non il piano orizzontale ad essa tangente, perché è l'unica rispetto alla quale la forza peso del mobile (orientata verso il centro della Terra lungo il suo raggio) è ovunque bilanciata dalla reazione vincolare: «*Adunque una superficie che dovesse esser non declive e non acclive, bisognerebbe che in tutte le sue parti fusse egualmente distante dal centro. Ma di tali superficie ve n 'è egli alcuna al mondo*?» Ora Galileo dà all'inclinazione (*declive*, *acclive*) un significato non geometrico ma fisico! Se fosse geometrico dovrebbe considerare il piano tangente alla superficie terrestre ovvero il piano "orizzontale", mentre invece acquista il significato fisico di una superficie (non più un piano) lungo la quale permane l'annullamento della forza peso e della reazione vincolare e quindi lungo la quale il mobile non sarebbe soggetto né al moto accelerato nè a quello ritardato del caso del piano inclinato. E questa superficie «*né acclive né declive*», come correttamente risponde Simplicio, è «*quella del nostro globo terrestre***»**. Per illustrare cosa accadrebbe su una tale superficie Galileo, purtroppo, ricorre ancora una volta all'esempio della nave: «*Adunque una nave che vadia movendosi per la bonaccia del mare, è un di quei mobili che scorrono per una di quelle superficie che non sono né declivi nè acclivi, e però disposta, quando le fusser rimossi tutti gli ostacoli accidentarii ed esterni, a muoversi, con l'impulso concepito una volta, incessabilmente e uniformemente*». Dunque, un altro caso di "inerzia circolare", direbbe Koyré! Si potrebbe, però, ripetere l'osservazione della distinzione fra rettilineità estrinseca ed intrinseca.

Ecco di seguito l'intero brano tratto dal *Dialogo* (Favaro, 1890-1909, Vol. VII, *Dialogo*, pp. 173-174):

> SALV. Parmi dunque sin qui che voi mi abbiate esplicati gli accidenti d'un mobile sopra due diversi piani; e che nel piano inclinato il mobile grave spontaneamente descende e va continuamente accelerandosi, e che a ritenervelo in quiete bisogna usarvi forza ; ma sul piano ascendente ci vuol forza a spignervelo ed anco a fermarvelo, e che 'l moto impressogli va continuamente scemando, si che finalmente si annichila. Dite ancora di piu che nell'un caso e nell'altro nasce diversità dall'esser la declività o acclività del piano, maggiore o minore; si che alla maggiore inclinazione segue maggior velocità, e, per l'apposito, sopra 'l piano acclive il medesimo mobile cacciato dalla medesima forza in maggior distanza si muove quanto l'elevazione è minore. Ora ditemi quel che accaderebbe del medesimo mobile sopra una superficie che non fusse né acclive né declive.

SIMP. Qui bisogna ch'io pensi un poco alla risposta. **Non vi essendo declività, non vi può essere inclinazione naturale al moto, e non vi essendo acclività, non vi può esser resistenza all'esser mosso**, talché verrebbe ad essere indifferente tra la propensione e la resistenza al moto: **parmi dunque che e' dovrebbe restarvi naturalmente fermo**. Ma io sono smemorato, perché non è molto che 'l signor Sagredo mi fece intender che cosi seguirebbe.

SALV. Cosi credo, **quando altri ve lo posasse fermo; ma se gli fusse dato impeto verso qualche parte, che seguirebbe?**

SIMP. Seguirebbe il muoversi verso quella parte.

SALV. Ma di che sorte di movimento? di continuamente accelerato, come ne' piani declivi, o di successivamente ritardato, come negli acclivi?

SIMP. Io non ci so scorgere causa di accelerazione né di ritardamento, non vi essendo né declività né acclività.

SALV. Si. Ma se non vi fusse causa di ritardamento, molto meno vi dovrebbe esser di quiete: **quanto dunque vorreste voi che il mobile durasse a muoversi?**

SIMP. **Tanto quanto durasse la lunghezza di quella superficie né erta né china.**

SALV. **Adunque se tale spazio fusse interminato, il moto in esso sarebbe parimente senza termine, cioè perpetuo?**

SIMP. Parmi di si, quando il mobile fusse di materia da durare.

SALV. Già questo si è supposto, mentre si è detto che si rimuovano tutti gl'impedimenti accidentarii ed esterni, e la fragilità del mobile, in questo fatto, è un degli impedimenti accidentarii. Ditemi ora: quale stimate voi la cagione del muoversi quella palla spontaneamente sul piano inclinato, e non , senza violenza, sopra l'elevato?

SIMP. Perché l'inclinazion de' corpi gravi è di muoversi verso 'l centro della Terra, e solo per violenza in su verso la circonferenza; e la superficie inclinata è quella che acquista vicinità al centro, e l'acclive discostamento.

SALV. **Adunque una superficie che dovesse esser non declive e non acclive, bisognerebbe che in tutte le sue parti fusse egualmente distante dal centro**. Ma di ta li superficie ve n 'è egli alcuna al mondo?

SIMP. Non ve ne mancano: **ècci quella del nostro globo terrestre**, se però ella fusse ben pulita, e non, quale ella è, scabrosa e montuosa; ma vi è quella dell'acqua, mentre è placida e tranquilla.

SALV. **Adunque una nave che vadia movendosi per la bonaccia del mare, è un di quei mobili che scorrono per una di quelle superficie che non sono né declivi nè acclivi, e però disposta, quando le fusser rimossi tutti gli ostacoli accidentarii ed esterni, a muoversi, con l'impulso concepito una volta, incessabilmente e uniformemente.**

## 2. Il significato del "moto naturale" in Galilei

Naturale, riferito al moto, può avere due significati: uno esclusivamente fisico e l'altro metafisico. Nel primo caso naturali sono i moti che è dato osservare in natura, senza chiedersi la ragione ultima della loro esistenza limitandosi a comprenderne soltanto le cause efficienti. Nel secondo caso, invece, naturali sono i moti che avvengono per una causa prima, che è frutto di una convizione metafisica.

Per Galileo il significato di moto naturale è metafisico: i moti naturali sono quelli che la natura consente per mantenere i corpi nei loro luoghi naturali, stabiliti dalla Natura per realizzare un ordine finale. In tale ottica il moto rettilineo (o retto, come lo chiama Galileo) non può essere naturale, perché allontanerebbe i corpi dai loro luoghi propri, e in particolare non sarebbe possibile il moto rettilineo prolungato all'infinito (Favaro, 1890-1909, Vol. VII, *Dialogo*, pp. 43-44):

*... quello che si muove di moto retto, muta luogo; e continuando di muoversi, si va più e più sempre allontanando dal termine ond'ei si partì e da tutti i luoghi per i quali successivamente ei va passando; e se tal moto naturalmente se gli conviene, adunque egli da principio non era nel luogo suo naturale, e però non erano le parti del mondo con ordine perfetto disposte: ma noi supponghiamo, quelle esser perfettamente ordinate: adunque, come tali, è impossibile che abbiano da natura di mutar luogo, ed in conseguenza di muoversi di moto retto. In oltre, essendo il moto retto di sua natura infinito, perché infinita e indeterminata è la linea retta, è impossibile che mobile alcuno abbia da natura principio di muoversi per linea retta, cioè verso dove è impossibile di arrivare, non vi essendo termine prefinito…*

Soltanto nella fase di creazione dell'universo la Natura, forse, potrebbe aver utilizzato il moto rettilineo per permettere ai corpi di raggiungere, percorrendo la via più breve (quindi rettilinea), le loro posizioni finali, ovvero i luoghi naturali (Ivi):

> *... si potrebbe favoleggiare che fusse avvenuto del primo caos, dove confusamente ed inordinatamente andavano indistinte materie vagando, per le quali ordinare la natura molto acconciamente si fusse servita de i movimenti retti [...] nè altro è l'acquisto suo se non l'avvicinarsi al luogo desiderato, cioè dove l'inclinazion naturale lo tira; e là si condurrà egli per la più breve, cioè per linea retta ...*

Ma una volta raggiunti i luoghi naturali, il moto rettilineo non è più necessario, anzi sarebbe dannoso, perché li allontanerebbe da essi (Ivi):

> *... dopo l'ottima distribuzione e collocazione è impossibile che in loro resti naturale inclinazione di più muoversi di moto retto, dal quale ora solo ne seguirebbe il rimuoversi dal proprio e natural luogo, cioè il disordinarsi.*

Questo il brano completo (Favaro, 1890-1909, Vol. VII, *Dialogo*, pp. 43-44):

> *SALV. Stabilito dunque cotal principio, si può immediatamente concludere che, se i corpi integrali del mondo devono esser di lor natura mobili, è impossibile che i movimenti loro siano retti, o altri che circolari: e la ragione è assai facile e manifesta. Imperocchè* ***quello che si muove di moto retto, muta luogo; e continuando di muoversi, si va più e più sempre allontanando dal termine ond'ei si partì e da tutti i luoghi per i quali successivamente ei va passando****;* ***e se tal moto naturalmente se gli conviene, adunque egli da principio non era nel luogo suo naturale, e però non erano le parti del mondo con ordine perfetto disposte: ma noi supponghiamo, quelle esser perfettamente ordinate: adunque, come tali, è impossibile che abbiano da natura di mutar luogo, ed in conseguenza di muoversi di moto retto****.* ***In oltre, essendo il moto retto di sua natura infinito, perché infinita e indeterminata è la linea retta, è impossibile che mobile alcuno abbia da natura principio di muoversi per linea retta, cioè verso dove è impossibile di arrivare, non vi essendo termine prefinito****; e la natura, come ben dice Aristotile medesmo, non intraprende a fare quello che non può esser fatto, nè intraprende a muovere dove è impossibile a pervenire. E se pur alcuno dicesse, che se bene la linea retta, ed in conseguenza il moto per essa, è produttibile in infinito, cioè interminato, tuttavia però la natura, per così dire, arbitrariamente gli ha assegnati alcuni termini e dato naturali instinti a' suoi corpi naturali di muoversi a*

*quelli, io risponderò che ciò per avventura* ***si potrebbe favoleggiare che fusse avvenuto del primo caos, dove confusamente ed inordinatamente andavano indistinte materie vagando, per le quali ordinare la natura molto acconciamente si fusse servita de i movimenti retti****, i quali, sì come movendo i corpi ben costituiti gli disordinano, così sono acconci a ben ordinare i pravamente disposti; ma* ***dopo l'ottima distribuzione e collocazione è impossibile che in loro resti naturale inclinazione di più muoversi di moto retto, dal quale ora solo ne seguirebbe il rimuoversi dal proprio e natural luogo, cioè il disordinarsi. Possiamo dunque dire, il moto retto servire a condur le materie per fabbricar l'opera, ma fabbricata ch'ell'è, o restare immobile, o, se mobile, muoversi solo circolarmente****; se però noi non volessimo dir con Platone, che anco i corpi mondani, dopo l'essere stati fabbricati e del tutto stabiliti, furon per alcun tempo dal suo Fattore mossi di moto retto, ma che dopo l'esser pervenuti in certi e determinati luoghi, furon rivolti a uno a uno in giro, passando dal moto retto al circolare, dove poi si son mantenuti e tuttavia si conservano: pensiero altissimo e degno ben di Platone, intorno al quale mi sovviene aver sentito discorrere il nostro comune amico Accademico Linceo; e se ben mi ricorda, il discorso fu tale. Ogni corpo costituito per qualsivoglia causa in istato di quiete, ma che per sua natura sia mobile, posto in libertà si moverà, tutta volta però ch'egli abbia da natura inclinazione a qualche luogo particolare; chè quando e' fusse indifferente a tutti, resterebbe nella sua quiete, non avendo maggior ragione di muoversi a questo che a quello. Dall'aver questa inclinazione ne nasce necessariamente che egli nel suo moto si anderà continuamente accelerando; e cominciando con moto tardissimo, non acquisterà grado alcuno di velocità, che prima e' non sia passato per tutti i gradi di velocità minori, o vogliamo dire di tardità maggiori: perchè, partendosi dallo stato della quiete (che è il grado di infinita tardità di moto), non ci è ragione nissuna per la quale e' debba entrare in un tal determinato grado di velocità, prima che entrare in un minore, ed in un altro ancor minore prima che in quello; anzi par molto ben ragionevole passar prima per i gradi più vicini a quello donde ei si parte, e da quelli a i più remoti; ma il grado di dove il mobile piglia a muoversi è quello della somma tardità, cioè della quiete.* ***Ora, questa accelerazion di moto non si farà se non quando il mobile nel muoversi acquista; nè altro è l'acquisto suo se non l'avvicinarsi al luogo desiderato, cioè dove l'inclinazion naturale lo tira; e là si condurrà egli per la più breve, cioè per linea retta. Possiamo dunque ragionevolmente dire che la natura, per conferire in un mobile, prima costituito in quiete, una determinata velocità, si serva del farlo muover, per alcun tempo e per qualche spazio, di moto retto.*** *Stante questo discorso, figuriamoci aver Iddio creato il corpo, v. g., di Giove, al quale abbia determinato di voler conferire una tal velocità, la quale egli poi debba conservar perpetuamente uniforme: potremo con Platone dire che gli desse di* ***muoversi da principio di moto retto ed accelerato, e che poi,***

> ***giunto a quel tal grado di velocità, convertisse il suo moto retto in circolare, del quale poi la velocità naturalmente convien esser uniforme.***

Nei *Frammenti attenenti al Dialogo sopra i due massimi sistemi del mondo* Galilei, per bocca di Salviati, ribadisce, con vari esempi, che tutti i moti naturali non sono retti ma circolari (Favaro, 1890-1909, Vol. VII, in calce al *Dialogo*, p. 545):

> *SALV. Io dico che nissuna cosa si muove naturalmente di moto retto. Cominciamo a ricercar discorrendo: i moti di tutti i corpi celesti son circolari; le navi, i carri, i cavalli, gli uccelli, tutti si muovon di moto circolare intorno al globo terrestre; i moti delle parti degli animali son tutti circolari: ed in somma noi ci riduciamo a non trovare altro che "gravia deorsum et levia sursum"*[11] *sembrino muoversi rettamente; ma nè di questi siamo sicuri, se prima non si dimostra che il globo terrestre stia immobile.*

In particolare Galileo si sofferma, altrove, sui movimenti degli arti negli animali e nell'uomo che sono tutti circolari (Favaro, 1890-1909, Vol. VII, *Dialogo*, pp. 283-284):

> *SALV. .... Voi primieramente ammettete per vero che la natura abbia fatti gli articoli, le flessure e snodature a gli animali, acciocchè si possano muover di molti e diversi movimenti; ed io vi nego questa proposizione, e dico che* ***le flessioni son fatte acciocchè l'animale possa muovere una o più delle sue parti, restando immobile il resto, e dico che*** *quanto alle spezie e differenze de' movimenti, quelli* ***sono*** *di una sola, cioè* ***tutti circolari****: e per questo voi vedete, tutti i capi de gli ossi mobili esser colmi o cavi; e di questi, altri sono sferici, che son quelli che hanno a muoversi per tutti i versi, come fa nella snodatura della spalla il braccio dell'alfiere nel maneggiar l'insegna, e dello strozziere nel richiamar co 'l logoro il falcone, e tal è la flessura del gomito, sopra la quale si gira la mano nel forar col succhiello; altri son circolari per un sol verso e quasi cilindrici, che servono per le membra che si piegano in un sol modo, come le parti delle dita l'una sopra l'altra, etc. Ma senza più particolari incontri, un solo general discorso ne può far conoscer questa verità; e questo è, che di un corpo solido che si muova restando uno de' suoi estremi senza mutar luogo, il moto non può esser se non circolare: e perchè nel muover l'animale uno delle sue membra non lo separa dall'altro suo conterminale, adunque tal moto è circolare di necessità.*

---

[11] In latino e in corsivo nell'originale in volgare: i corpi pesanti verso il basso (*gravia deorsum*) e i corpi leggeri verso l'alto (*levia sursum*).

Ma Simplicio obietta che non tutti i moti degli animali sono circolari:

> SIMP. *Io non l'intendo per questo verso; anzi veggo io l'animale muoversi di cento moti non circolari e diversissimi tra loro, e correre e saltare e salire e scendere e notare e molt'altri.*

E Galileo risponde che è vero, ammettendo, questa volta, che non tutti i moti naturali sono circolari, poiché diversi sono i movimenti delle singole parti (circolari) e dell'intero corpo (rettilinei):

> SALV. *Sta bene: ma cotesti son moti secondarii, dependenti da i primi, che sono de gli articoli e delle flessure. Al piegar delle gambe alle ginocchia e delle cosce a i fianchi, che* ***son moti circolari delle parti, ne viene in conseguenza il salto o il corso, che son movimenti di tutto 'l corpo, e questi posson esser non circolari.*** *Ora, perchè del globo terrestre non si ha da muovere una parte sopra un'altra immobile, ma il movimento deve esser di tutto il corpo, non ci è bisogno di flessure.*

Certamente c'è una imprecisione da parte di Galileo nel voler identificare nel moto circolare qualunque moto avvenga in natura. Le orbite ellittiche dei pianeti ne sono una prova. Ma questa è un'altra questione, che si collega al particolare fascino che certamente Galileo nutriva per il cerchio, figura geometrica perfetta per lui, sia come matematico sia come filosofo sia come critico d'arte (Panofsky, 1956).

Tutta la controversia sulla supposta inerzia circolare di Galileo si basa sulla sua affermazione del moto circolare come moto naturale. Ma tale controversia svanisce nel nullo ove si rimuova l'attribuzione di "naturale" al moto rettilineo uniforme, peraltro non presente nella formulazione di Newton del Principio d'Inerzia, così come in qualunque altra formulazione successiva.

"Naturale" attribuito al moto rettilineo uniforme è un'arbitraria estrapolazione metafisica. La ragione, già detta, è che tale moto, avvenendo in assenza di una forza netta, viene interpretato come una "disposizione naturale" della materia. Ma questa attribuzione nasconde una interpretazione metafisica che dovrebbe essere estranea alla fisica, secondo quanto già affermato da Newton con la sua famosa frase «*Hypotheses non fingo*».[12] Il fisico deve limitarsi ad affermare che sperimentalmente è ragionevole affermare che in condizioni di assenza di una forza netta un corpo permane nel suo stato di quiete, se inizialmente vi si trova, o in quello di moto rettilineo uniforme se già è in moto. Questo è il Principio d'Inerzia e quanto asserisce "operativamente" Galileo nelle Giornate Quarta e Sesta dei *Discorsi*, che non è in contraddizione con le sue ripetute affermazioni che gran parte dei moti naturali sono circolari,

[12] Contenuta nello *Scholium Generale* della seconda edizione dei *Philosophiæ Naturalis Principia Mathematica* del 1713.

in quanto la qualifica di moto naturale a quello rettilineo uniforme non compete alla fisica, è un arbitrario sconfinamento nella metafisica, in contrasto con il monito newtoniano.

Pertanto, tutta la polemica di Koyré e altri, relativa alla supposta inerzia circolare di Galileo, è priva di senso.

Quanto poi alla stessa attribuzione metafisica del "naturale" nel significato di "spontaneo" al moto rettilineo uniforme, su quali argomentazioni poggia? Cosa sappiamo dell'origine dell'inerzia? Secondo la prima spiegazione, dovuta a Ernst Mach nel 1893 e poi formulata da Einstein come *Principio di Mach*, l'inerzia di ogni corpo sarebbe il risultato della sua interazione con il resto dell'universo e in particolare con le stelle (Nicotra, 2022, p. 88). Dunque non sarebbe una disposizione intrinseca della materia ma il risultato di una moltitudine enorme di interazioni che si equilibrano. C'è da chiedersi: sono proprio identiche le situazioni in cui un corpo non è sottoposto ad alcuna forza e, invece, lo è ad *n* forze la cui risultante sia nulla? La spiegazione di Mach sembra attribuire l'inerzia a quest'ultimo caso. D'altra parte, mentre è possibile ottenere sperimentalmente il moto rettilineo uniforme equilibrando *n* forze su di esso applicate e mantenerlo fin tanto che tale equilibrio permane, non è possibile ottenerlo con un corpo non sottoposto ad alcuna forza: il corpo rimarrebbe fermo. Ma poi è realmente possibile un tal caso? Cioè realizzare un ambiente fisico privo di alcuna interazione?

Da ultimo non sono forse da trascurare alcune riflessioni sulla presenza dell'infinito in fisica, concetto che è contenuto nella formulazione del Principio d'Inerzia. È ben noto che in fisica il concetto di infinito, pur ponendo molti interessanti interrogativi, non trova cittadinanza: ogni esperienza fisica è necessariamente limitata nel tempo e nello spazio. In tale spirito, ciò che noi giudichiamo moto rettilineo prolungato all'infinito nel tempo e nello spazio in realtà potrebbe essere una estrema approssimazione di un moto circolare con raggio di curvatura talmente grande da essere mascherato dagli inevitabili errori di misura della "rettilineità". Galileo accenna a tale possibilità. Inoltre, anche ammettendone la reale esistenza, è plausibile muovere nei riguardi del moto rettilineo perpetuo affermato nel Principio d'Inerzia, la stessa obiezione lecitamente posta al V postulato di Euclide (postulato delle parallele): nessun esperimento potrà mai né confermarlo né confutarlo, perché richiederebbe un processo infinito, quindi sperimentalmente irrealizzabile. Dunque, come non si potrà mai compiutamente dimostrare che esistono realmente rette parallele, perché ciò richiederebbe una verifica *ad infinitum* irrealizzabile, così chi ci può dare la certezza che veramente, in assenza di forza netta ad esso applicata, un corpo continui "sempre", *ad infinitum*, a muoversi secondo una retta e con velocità scalare costante? Non ha forse ragione Galilei, quando afferma (Favaro, 1890-1909, Vol. VII, *Dialogo*, pp. 43-44):

> *E se pur alcuno dicesse, che se bene la linea retta, ed in conseguenza il moto per essa, è produttibile in infinito, cioè interminato, tuttavia però la natura, per così dire, arbitrariamente gli ha assegnati alcuni termini e dato naturali instinti a' suoi corpi naturali di muoversi a quelli ...*

## 3. Galileo e il principio di relatività

Gli storici che attribuiscono a Galilei l'errore della circolarità, nel suo Principio d'Inerzia, disconoscono pure, per lo stesso motivo, la sua paternità nei riguardi del Principio di Relatività Classica,[13] espresso per bocca di Salviati nella Giornata Seconda del *Dialogo*, nel quale Galileo riprende, come laboratorio per le sue esperienze, la nave già utilizzata da Bruno (1584), che diventa il «*gran navilio*» (Favaro, 1890-1909, Vol. VII, *Dialogo*, pp. 212-213):

> *Riserratevi con qualche amico nella maggiore stanza che sia sotto coverta di alcun gran navilio, quivi fate d'aver mosche, farfalle e simili animaletti volanti; siavi anco un gran vaso d'acqua, e dentrovi de' pescetti; sospendasi anco in alto qualche secchiello, che a goccia a goccia vadia versando dell'acqua in un altro vaso di angusta bocca, che sia posto a basso: e stando ferma la nave, osservate diligentemente come quelli animaletti volanti con pari velocità vanno verso tutte le parti della stanza; i pesci si vedranno andar notando indifferentemente per tutti i versi; le stille cadenti entreranno tutte nel vaso sottoposto; e voi, gettando all'amico alcuna cosa, non piú gagliardamente la dovrete gettare verso quella parte che verso questa, quando le lontananze sieno eguali; e saltando voi, come si dice, a piè giunti, eguali spazii passerete verso tutte le parti. Osservate che avrete diligentemente tutte queste cose, benché niun dubbio ci sia che mentre il vassello sta fermo non debbano succeder cosí, fate muover la nave con quanta si voglia velocità; ché (***pur che il moto sia uniforme e non fluttuante in qua e in là) voi non riconoscerete una minima mutazione in tutti li nominati effetti, né da alcuno di quelli potrete comprender se la nave cammina o pure sta ferma***: voi saltando passerete nel tavolato i medesimi spazii che prima, né, perché la nave si muova velocissimamente, farete maggior salti verso la poppa che verso la prua, benché, nel tempo che voi state in*

[13] Questo asserisce che in un sistema di corpi non soggetto a forze esterne o soggetto a forze esterne a risultante nulla (sistema isolato) le leggi della meccanica rimangono invariate rispetto a qualunque sistema di riferimento in moto rettilineo uniforme o quiescente, e quindi non è possibile con nessuna esperienza meccanica decidere se il sistema si muove o è fermo. Nel principio di relatività classica si utilizza ancora il concetto di quiete assoluta, poi rimosso dalla Relatività di Einstein.

> *aria, il tavolato sottopostovi scorra verso la parte contraria al vostro salto; e gettando alcuna cosa al compagno, non con piú forza bisognerà tirarla, per arrivarlo, se egli sarà verso la prua e voi verso poppa, che se voi fuste situati per l'opposito; le gocciole cadranno come prima nel vaso inferiore, senza caderne pur una verso poppa, benché, mentre la gocciola è per aria, la nave scorra molti palmi; i pesci nella lor acqua non con piú fatica noteranno verso la precedente che verso la sussequente parte del vaso, ma con pari agevolezza verranno al cibo posto su qualsivoglia luogo dell'orlo del vaso; e finalmente le farfalle e le mosche continueranno i lor voli indifferentemente verso tutte le parti, né mai accaderà che si riduchino verso la parete che riguarda la poppa, quasi che fussero stracche in tener dietro al veloce corso della nave, dalla quale per lungo tempo, trattenendosi per aria, saranno state separate; e se abbruciando alcuna lagrima d'incenso si farà un poco di fumo, vedrassi ascender in alto ed a guisa di nugoletta trattenervisi, e indifferentemente muoversi non piú verso questa che quella parte.*

Galileo stesso "spiega" il Principio di Relatività proprio ricorrendo al Principio d'Inerzia, dal quale soltanto può derivare (Ibidem, p. 214):

> ***E di tutta questa corrispondenza d'effetti ne è cagione l'esser il moto della nave comune a tutte le cose contenute in essa ed all'aria*** *ancora, che per ciò dissi io che si stesse sotto coverta; ché quando si stesse di sopra e nell'aria aperta e non seguace del corso della nave, differenze piú e men notabili si vedrebbero in alcuni de gli effetti nominati: e non è dubbio che il fumo resterebbe in dietro, quanto l'aria stessa; le mosche parimente e le farfalle, impedite dall'aria, non potrebber seguir il moto della nave, quando da essa per spazio assai notabile si separassero; ma trattenendovisi vicine, perché la nave stessa, come di fabbrica anfrattuosa, porta seco parte dell'aria sua prossima, senza intoppo o fatica seguirebbon la nave, e per simil cagione veggiamo tal volta, nel correr la posta, le mosche importune e i tafani seguir i cavalli, volandogli ora in questa ed ora in quella parte del corpo; ma nelle gocciole cadenti pochissima sarebbe la differenza, e ne i salti e ne i proietti gravi, del tutto impercettibile.*

Certamente la spiegazione data («*E di tutta questa corrispondenza d'effetti ne è cagione l'esser il moto della nave comune a tutte le cose contenute in essa ed all'aria ancora*») non è esplicitamente ricondotta al Principio d'Inerzia ma lo è in maniera sottaciuta. Oggi diremmo che «*l'esser il moto della nave comune a tutte le cose contenute in essa*» non è altro che una conseguenza del fatto che anche su tutto ciò che è contenuto all'interno della nave (Galileo lo specifica puntualmente: «*per ciò dissi io che si stesse sotto coverta*») non agisce nessuna forza netta in grado di mutarne il moto rettilineo uniforme condiviso con la nave. Pertanto, non essendoci moti relativi fra gli oggetti all'interno della nave, le loro posizioni reciproche rimangono le stesse che nel caso della nave in quiete. Dunque, nessuna esperienza meccanica è in grado di distinguere se la nave sta ferma o si muove di moto rettilineo uniforme.

Secondo Paolo Rossi, tuttavia, Galilei non avrebbe mai formulato il principio di relatività classica (Rossi, 2006, p. 216):

> *Quel principio che è noto nei manuali come il principio della relatività galileiana (in base alle osservazioni meccaniche compiute all'interno di un sistema*[14] *non si può decidere se il sistema stesso sia in quiete o in moto rettilineo uniforme) non corrisponde a quello effettivamente formulato da Galilei che intendeva mostrare mediante quella sua dottrina l'impossibilità, per un osservatore collocato sulla Terra, di percepire il moto di rotazione della Terra medesima. […] Il moto «non fluttuante in qua e in là», nell'esempio galileiano della nave, vuol dire moto retto o diritto o procedente lungo il medesimo meridiano terrestre, ed è una forzatura tradurre «non fluttuante» con "rettilineo" (che è termine altrove e più volte impiegato da Galilei). La differenza non è lieve, perché il principio classico di relatività implica il concetto di un moto rettilineo uniforme e l'accettazione del principio d'inerzia (per il quale ogni corpo persevera nel suo stato di quiete o di moto rettilineo uniforme finché non intervenga una forza a modificare tale stato). Questo principio, che è alle radici della dinamica moderna, non fu mai formulato da Galilei proprio a causa dell'azione esercitata sulla sua fisica dalle sue convinzioni cosmologiche.*

Inoltre, osserva Paolo Rossi, Galilei per il «*gran navilio*» non avrebbe considerato un moto rettilineo e uniforme perché:

> *Il moto «non fluttuante in qua e in là», nell'esempio galileiano della nave, vuol dire moto retto o diritto o procedente lungo il medesimo meridiano terrestre, ed è una forzatura tradurre «non fluttuante» con "rettilineo" (che è termine altrove e più volte impiegato da Galilei).*

Tuttavia, l'interpretazione del termine «*non fluttuante*» data da Paolo Rossi non è anch'essa una ancor maggiore forzatura? Dove, Galilei, definisce il moto retto come «*procedente lungo il medesimo meridiano terrestre*»? D'altra parte, perché mai avrebbe dovuto considerare il caso particolarissimo di una nave che si muova lungo un meridiano? Inoltre, se avesse voluto significare "circolare", per indicare il moto lungo la superficie sferica terrestre, perché avrebbe dovuto usare «*non fluttuante*» in luogo di circolare, termine, già da lui numerose altre volte usato? Galilei usa il termine "moto circolare" 66 volte nel *Dialogo* e 4 volte ne *Il Saggiatore*.[15]

[14] Paolo Rossi dimentica di specificare che il sistema deve essere isolato.
[15] Vedi Appendice.

Rossi afferma che Galileo ha più volte impiegato altrove il termine "rettilineo", ma dove e in quali contesti? Nel *Dialogo,* "rettilineo" è utilizzato 3 volte, ma non riferito al moto.[16] Nei *Discorsi* non è presente. È certamente utilizzato in altre opere di Galileo (21 volte), ma sempre in scritti di geometria e mai riferito al moto.[17] Inoltre l'espressione "moto rettilineo" non figura mai sia nel *Dialogo* sia ne *Il Saggiatore* sia nei *Discorsi*.

Risultato negativo anche per le occorrenze di espressioni del tipo "moto lungo (o secondo) un meridiano"; la parola "meridiano" compare 32 volte nel *Dialogo,* ma non riferita al moto.[18]

Quanto all'espressione "moto retto" (che secondo Rossi significherebbe per Galilei "moto lungo un meridiano") essa è utilizzata 52 volte nel *Dialogo* e 25 nel *Saggiatore*[19] al posto di "rettilineo" nel senso di lungo una retta, come risulta chiaro in più contesti, per esempio nella Giornata 1.28 del *Dialogo*:

> *\SAGR.\... ma di già abitato da noi. Che se il *moto retto* è semplice per la semplicità della linea retta, e...*

Oppure ne *Il Saggiatore* - 10.5:

> *[...] manchevole la scusa del Sarsi, perché non solamente il *moto* veramente *retto* apparisce per linea retta.*

Dunque, Galileo, senza possibilità di alcun dubbio, utilizza l'espressione "moto retto" per indicare il moto "lungo una retta", e la retta cui si riferisce è certamente quella euclidea, in quanto non è pensabile che intendesse per retta un meridiano, perché altrimenti avrebbe anticipato di oltre un secolo la geometria non-euclidea ellittica! Ma sappiamo bene che Galileo non è un antesignano dei geometri non-euclidei.

# 4. Conclusioni

Certamente le "formulazioni" del Principio d'Inerzia che è possibile rintracciare nell'immensa opera galileiana non hanno la cristallina chiarezza e incisività della formulazione contenuta nei *Philosophiæ Naturalis Principia Mathematica* di Newton (1687, 2001), per le ragioni già dette nella prima parte di questo scritto (Nicotra, 2022, Drago, 1997) sulla impostazione delle opere di Galilei, orientata non alla esposizione sistematica delle ricerche

[16] Vedi Appendice.

[17] 10 volte nel Vol.II, 6 volte nel Vol. XVI, 5 volte nel Vol. XVIII dell'Edizione Nazionale (Favaro, 1890-1909).

[18] Vedi Appendice.

[19] Vedi Appendice.

secondo un ordine deduttivo proprio della trattatistica, bensì alla loro discussione e presentazione in una forma "costruttiva", propria della fase di ricerca. A tale scopo quale forma letteraria più adatta potrebbe esserci se non quella del dialogo, così magistralmente utilizzata da Galileo?

In conclusione si può dire che l'approccio di Galilei all'inerzia è quello di enunciati operativi e metodologici utili per la ricerca, per capire "nella sostanza" attraverso casi concreti, in contrapposizione all'approccio di enunciati "apodittici" posti al principio di una teoria deduttiva. Questo spiega la diversità delle formulazioni "operative" del Principio d'Inerzia presenti nelle opere galileiane, che rispecchiano le diverse fasi di astrazione maturate man mano nelle sue ricerche. In altri termini, l'inerzia di Galilei è l'inerzia della fase creativa di ricerca, mentre l'inerzia di Newton è quella della fase di sistemazione logica propria della trattatistica.

A riprova di quanto appena asserito è illuminante quanto rivelato nell'*Avvertimento*, a cura di Antonio Favaro, che precede i *Discorsi* nell'Edizione Nazionale, che fa riferimento alla riedizione da parte di Galileo dei trattati giovanili *De motu aequabili*, *De motu naturaliter accelerato*, *De motu proiectorum* come Terza e Quarta Giornata dei *Discorsi* nel 1638 (Favaro,1890-1909, Vol. VIII, *Discorsi*, p. 12):

> *Scorrendo pertanto il suo carteggio, troviamo che appena nell'ottobre del 1630 egli informa l'Aggiunti degli acquisti conseguiti nella dottrina del moto; e che nel settembre del 1632, in occasione di dolersi con Cesare Marsili per la indiscrezione commessa dal Cavalieri intorno alla linea dei proietti, accenna ad un libro di prossima pubblicazione, nel quale avrebbe trattato anche di questa materia; e un anno dopo, rispondendo ad una lettera d'Andrea Arrighetti, che gli inviava, in seguito a richiesta avutane, due proposizioni concernenti le resistenze dei solidi, scrive d'aver per le mani un trattato intorno a quest'argomento. Con maggior precisione sotto il 7 marzo 1634 scriveva ad Elia Diodati d' essersi trattenuto cinque mesi in casa dell'Arcivescovo di Siena, «trattato da padre da Sua Signoria Illustrissima, ed in continue visite della nobiltà di quella città; dove composi un trattato di un argomento nuovo, in materia di mecaniche, pieno di molte specolazioni curiose ed utili» ; e gli amici, venutine in cognizione, lo sollecitavano a pubblicarlo. Finalmente, pochi mesi dopo scriveva a Fra Fulgenzio Micanzio: «Il trattato del moto, tutto nuovo, sta all'ordine; ma* ***il mio cervello inquieto non può restar d'andar mulinando, e con gran dispendio di tempo, perchè quel pensiero che ultimo mi sovviene circa qualche novità, mi fa buttare a monte tutti i trovati precedenti****.*

Forse è importante notare le date dei riferimenti al Principio d'Inerzia contenuti nelle opere di Galilei: 1590 nel *De Motu*, 1593 nelle *Mecaniche*, 1613 nella *Istoria e dimostrazioni intorno alle macchie solari e loro accidenti*, 1632 nel *Dialogo* e infine 1638 per le due formulazioni "corrette" contenute nei *Discorsi (*Quarta e Sesta Giornata). Soltanto in queste ultime compare esplicitamente il moto rettilineo, mentre nelle altre o non è dichiarato o, al contrario, è esplicitamente dichiarato il moto circolare. Un ravvedimento, dunque, di Galileo giunto alla fine della sua carriera di scienziato? Non sarebbe l'unico caso, nella storia della scienza, di un cambiamento, anche sostanziale, nelle idee di uno scienziato. Si ricordino in proposito le stesse parole di Galileo nella lettera a Fulgenzio Micanzio del 19 novembre 1634 (Bibl. Marciana, CI. XI It., cod. XLVII, car. 1) riportate più sopra dal Favaro:

> *Il trattato del moto, tutto nuovo, sta all' ordine; ma il mio cervello inquieto non può restar d' andar mulinando, e con gran dispendio di tempo, perchè quel pensiero che ultimo mi sovviene circa qualche novità, mi fa buttare a monte tutti i trovati precedenti.*

Quello che conta è in ogni caso l'ultima versione di un risultato scientifico, in accordo con il continuo progredire e perfezionarsi della ricerca scientifica. Pertanto potremmo, con i *Discorsi*, parlare di un *Galileus ab omni naevo circuli vindicatus*!

Malgrado i diversi tentativi di delegittimare la paternità di Galilei, a lui vengono ancora attribuiti sia il Principio d'Inerzia sia il Principio di Relatività Classica in quasi tutti i testi scolastici e universitari di autori affermati e anche in molti autorevoli testi di divulgazione scientifica.

Etienne Klein (2006, p. 58) afferma che «la meccanica newtoniana si fonda sul cosiddetto principio di relatività enunciato per la prima volta da Giordano Bruno e ripreso da Galileo». Fra le testimonianze più autorevoli mi piace ricordare le numerose citazioni di Albert Einstein sulla paternità galileiana del principio d'inerzia, fra cui questa molto esplicita (Einstein, 1962):

> *Classical mechanics is based on Galileo's principle: A body is in rectilinear and uniform motion as long as other bodies do not act on it.*

# 5. Appendice

Le occorrenze dei termini "rettilineo", "meridiano", "moto circolare" e "moto retto", presenti nel *Dialogo* e nel *Saggiatore*, sono state rilevate con l'applicazione del programma di interrogazione DBT (Data Base Testuale) di E. Picchi (Consiglio Nazionale delle Ricerche), sulle versioni digitali di quelle opere pubblicate nel 4° CD di LIZ Letteratura Italiana Zanichelli (Manierismo e Barocco), 2001. Per ogni occorrenza è riportato il contesto associato. La ricerca per famiglie di parole è stata effettuata con una distanza massima di 10 parole fra i termini e una lunghezza del contesto di 10 parole. Ricerche effettuate con valori maggiori di tali parametri hanno prodotto un maggior numero di occorrenze, ma è evidente che, aumentando ulteriormente la distanza fra i termini, aumenta anche la probabilità di una perdita di nesso fra di essi.

Espressione *rettilineo*

3 occorrenze ne *Il Saggiatore*:

*1) \SALV.\... D o altro in infinito più acuto, ma però sempre *rettilineo*; ma la diminuzione degli spazii per li quali il mobile - Giorn. 2.471;*

*2) \SALV.\... contengono un angolo infinitamente più stretto ed acuto di qualsivoglia acuto *rettilineo*, quale sarà questo. Piglisi nella perpendicolare A C qualsivoglia - Giorn. 2.471;*

*3) \SALV.\... per minime che elle siano e comprese dentro ad angustissimo angolo *rettilineo* delle quali parallele le parti che restano tra l'arco e - Giorn. 2.471;*

Espressione *meridiano*

32 occorrenze nel *Dialogo sopra i due massimi sistemi del mondo*:

*1) \SALV.\... delle quali, quando la Luna è nel *meridiano*, guarda verso maestro, e l'altra - Giorn. 1.182*

*2) \SIMP.\... esser nulla il passare il Sole da un *meridiano* all'altro, alzarsi sopra questo orizonte, - Giorn. 2.38*

*3) \SAGR.\... orizontale e l'altro verticale, sotto il *meridiano*. Ma che più? ditemi, signor - Giorn. 2.751*

*4) \SALV.\... sendo la stella nuova collocata nel cerchio *meridiano* là verso settentrione, a uno che da - Giorn. 3.48*

*5) \SALV.\... elevazioni polari e le altezze della stella nel *meridiano*, tanto le minime sotto il polo, - Giorn. 3.80*

*6) \SALV.\... 51.18 m.p.: l'altezza della stella nel *meridiano*, presa daTicone, fu gr. 27.45 - Giorn. 3.83*

*7) \SALV.\... tempo brevissimo, come fanno le remote dal *meridiano*: e se questo è, sì come - Giorn. 3.117*

*8) \SALV.\... nel qual è la stella, dal *meridiano*; in oltre dopo notabile intervallo di tempo - Giorn. 3.117*

*9) \SALV.\... tanto mentre ell'era nell'infima parte del *meridiano*, quanto nella suprema; l'altra è - Giorn. 3.119*

*10) \SALV.\... mentre si trovi nella parte di sotto nel *meridiano*, che quando è nella superiore, come - Giorn. 3.121*

*11) \SALV.\... nell'ora che il Sole si trovava nel *meridiano*; ed accortici come il viaggio suo non - Giorn. 3.226*

*12) \SALV.\... per il quale passerebbe il piano del nostro *meridiano*, nel qual piano sarebbe ancora l'asse - Giorn. 3.228*

*13) \SALV.\... il finitore A B C D né il *meridiano* A C passi per i poli dell'asse - Giorn. 3.228*

*14) \SALV.\... del terminatore A B e la sezione del *meridiano* A C, il diametro del cerchio massimo - Giorn. 3.228*

*15) \SALV.\... è quando detti poli sono nella sezion del *meridiano*, la curvità è ridotta al sommo, - Giorn. 3.228*

*16) \SALV.\... altre volte poi trovarsi nel piano del *meridiano* del riguardante, in modo tale che l' - Giorn. 3.236*

*17) \SALV.\... e l'altro il variar altezze nel *meridiano*; che si tira poi in conseguenza il - Giorn. 3.300*

*18) \SAGR.\... Sole si alza e si abbassa nel nostro *meridiano* per un arco grandissimo, quasi di 47 - Giorn. 3.301*

*19) \SALV.\... H, che verrà insieme ad esser un *meridiano*, ed in esso pigliamo una stella fuori - Giorn. 3.310*

*20) \SALV.\... eclittica; ed anco la sua altezza nel *meridiano* sarà fatta maggiore nello stato B che nel - Giorn. 3.310*

*21) \SALV.\... se noi piglieremo un'altra stella nel medesimo *meridiano*, più remota dall'eclittica, qual sarebbe - Giorn. 3.310*

*22) \SALV.\... sfera, nella quale s'intenda descritto un *meridiano* D F C, che sarà eretto al - Giorn. 3.310*

*23) \SALV.\... perché, se noi segneremo un altro *meridiano* men lontano dalla Terra, qual sarebbe questo - Giorn. 3.310*

*24) \SAGR.\... o avvicinamento) è il mostrarcisi nel medesimo *meridiano* ora più elevate ed ora meno. Di - Giorn. 3.311*

*25) \SALV.\... suo coperto può segare ad angoli retti il *meridiano* di qualche abitazione posta nella pianura. Voglio - Giorn. 3.319*

*26) \SALV.\... stelle del Carro, nel passar per il *meridiano*, venga ascondendosi doppo la trave già collocata - Giorn. 3.319*

*27) \SALV.\... D e 'l polo austrino B il Sole *meridiano* è elevato oltre al lor vertice verso 'l - Giorn. 3.330*

*28) \SALV.\... per tanti gradi qualsivoglia stella fissa osservata nel *meridiano* apparirebbe essersi elevata o inclinata. - Giorn. 3.331*

*29) \SALV.\... tutti questi in tal giorno averanno il Sole *meridiano* sopra il vertice loro, ed il Sole - Giorn. 3.332*

*30) \SALV.\... ago magnetico non solamente il drizzarsi sotto un *meridiano* verso i poli con moto orizontale - Giorn. 3.357*

*31) \SALV.\... accidente di declinare (stando bilanciato sotto il *meridiano* già segnato sopra una sferetta di calamita), - Giorn. 3.357*

*32) \SAGR.\... suo asse, stante nel piano di un *meridiano*, verso la superficie della Terra, e - Giorn. 3.370.*

Espressione *moto circolare*

66 Occorrenze nel *Dialogo sopra i due massimi sistemi del mondo*:

*1) \SAGR.\... che Aristotile accenna, un solo esser al mondo il *moto circolare*, ed in conseguenza un solo centro, al quale - Giorn. 1.28*

*2) \SAGR.\... moti semplici e *moto* misto, chiamando semplici il *circolare* ed il retto, e misto il composto di questi - Giorn.1.28*

*3) \SALV.\... il quale sia ancora tanto più eccellente, quanto il *moto circolare* è più perfetto del *moto* retto: - Giorn. 1.33*

*4) \SALV.\... afferma egli esser proprie del corpo semplice e mobile di *moto circolare*; - Giorn. 1.33*

*5) \SALV.\... rivolti a uno a uno in giro, passando dal *moto* retto* al *circolare*, - Giorn. 1.35*

*6) \SALV.\... a quel tal grado di velocità, convertisse il suo *moto* retto in *circolare*, del quale poi la velocità - Giorn. 1.35*

*7) \SAGR.\... al corpo di Giove, subito creato, il suo *moto circolare*, con tale e tanta velocità. \SALV.\ Io - Giorn. 1.36*

*8) \SALV.\... è *moto circolare* intorno al centro: adunque il *moto circolare* non s'acquisterà mai naturalmente - Giorn. 1.77*

*9) \SALV.\... partirsi ed allontanarsi dal medesimo termine; e perché nel *moto circolare* il mobile sempre si parte - Giorn. 1.87*

*10) \SALV.\... non essendo altro, salvo che la quiete e 'l *moto circolare*,*

*atto alla conservazione dell'ordine. Ed io non - Giorn. 1.87*

*11) \SALV.\... di questi, in modo che, negato che il *moto circolare* sia solo de i corpi celesti, ed affermato - Giorn. 1.96*

*12) \SALV.\... o che malamente e con errore abbia Aristotile dedotti dal *moto circolare* quelli che ha assegnati a i corpi celesti. - Giorn. 1.96*

*13) \SIMP.\... corpo celeste non si può assegnar contrario, imperocché al *moto circolare* niun altro movimento è contrario, - Giorn. 1.99*

*14) \SIMP.\... né secondo alcuna sua propria parte. Che poi al *moto circolare* niuno altro sia contrario, lo prova Aristotile - Giorn. 1.99*

*15) \SALV.\... sua illazione resta nulla. Dicovi per tanto che quel *moto circolare*, che voi assegnate a i corpi celesti, conviene - Giorn. 1.100*

*16) \SAGR.\... che voi dite che, negato ad Aristotile che il *moto circolare* non sia della Terra, come degli altri corpi celesti - Giorn. 1.101*

*17) \SALV.\... il movimento *circolare* e la quiete; ma quanto al *moto* per linea retta, non veggo, che possa servire - Giorn. 1.123*

*18) \SALV.\... anco per far questa restituzione non si trovasse che qualche *moto circolare* fusse più accomodato. - Giorn. 1.123*

*19) \SALV.\... intero globo e sfera de gli elementi attribuiscasi o il *moto circolare* o una perpetua consistenza - Giorn. 1.123*

*20) \SALV.\... naturale, ed all'incontro chiamano a lor preternaturale il *moto circolare*, del quale incessabilmente - Giorn. 1.123*

*21) \SIMP.\ ...Dal mancar di contrari immediatamente, e mediatamente dal *moto* semplice *circolare*. - Giorn. 1.260*

*22) \SIMP.\... adunque etc. Secondariamente, tutti gli altri mobili di *moto circolare* par che restino indietro e - Giorn. 2.66*

*23) \SALV.\... presa da certa esperienza, ed è tale. Il *moto *circolare* ha facoltà di estrudere, dissipare e scacciar dal suo - Giorn. 2.96*

*24) \SIMP.\... naturale delle parti è il *moto* retto al centro dell'universo, onde il *circolare* non gli può naturalmente competere. - Giorn. 2.101*

*25) \SALV.\... Non dic'egli che 'l *moto circolare* alla Terra sarebbe violento? e però non eterno? - Giorn. 2.104*

*26) \SALV.\... eterno. Ma se noi faremo la Terra mobile di *moto circolare*, questo potrà esser eterno ad essa ed alle parti - Giorn. 2.106*

*27) \SALV.\... pur forza (se voi non le volete concedere il *moto circolare*) di mantenerle e difenderle l'immobilità. - Giorn. 2.118*

*28) \SALV.\... costituendola lontana dal mezo, la facessero andar con *moto circolare* intorno ad esso mezo: - Giorn. 2.126*

*29) \SALV.\... stimato impossibile, che 'l sasso potesse muoversi di un *moto* misto di retto e di *circolare*; - Giorn. 2.147*

*30) \SAGR.\... assicura l'esperienza, se noi non veggiamo mai altro *moto* che il composto delli due, *circolare* ed in giù? - Giorn. 2.295*

*31) \SALV.\... cadendo dalla sommità della torre C, venga movendosi del *moto* composto del comune *circolare* e del suo proprio retto. - Giorn. 2.306*

*32) \SALV.\... non si muove realmente d'altro che di un *moto* semplice *circolare*, sì come quando posava sopra la torre - Giorn. 2.308*

*33) \SALV.\... quando posava sopra la torre pur si muoveva di un *moto* semplice e *circolare*. - Giorn. 2.308*

*34) \SAGR.\... costituite, gli vien levato, ed assegnato pur al *moto circolare*. - Giorn. 2.311*

*35) \SIMP.\... Il *moto* del sasso sin che è nella cocca è *circolare* cioè va per un arco di cerchio, il cui - Giorn. 2.410*

*36) \SALV.\... la pietra scappa dalla canna, qual è il suo *moto*? séguit'ella di continuare 'l suo precedente *circolare*, - Giorn. 2.411*

*37) \SIMP.\... Io non intendo che 'l *moto* proietto sia a dirittura di tutto il *circolare*, ma di quell'ultimo punto dove terminò il moto - Giorn. 2.420*

*38) \SIMP.\... circolare, ma di quell'ultimo punto dove terminò il *moto circolare*. Io mi intendo dentro di me, ma non - Giorn. 2.420*

*39) \SALV.\...e se da questo *moto circolare* deve passar al *moto* retto, qual dovrà esser questa linea retta? - Giorn. 2.425*

*40) \SALV.\... Sin qui avete per voi stesso saputo che il *moto *circolare* del proiciente imprime nel proietto - Giorn. 2.437*

*41) \SALV.\...Talché il discostamento del proietto dalla circonferenza del precedente *moto *circolare* in su 'l principio è - Giorn. 2.439*

*42) \SIMP.\... fusse eguale alla velocità del suo *moto circolare* fatto nel cerchio massimo del concavo dell'orbe lunare, - Giorn. 2.559*

*43) \SALV.\... mentre si tratteneva nel concavo della Luna aveva il *moto circolare* delle ventiquattr'ore insieme con la Terra - Giorn. 2.617*

*44) \SIMP.\... come voi vedete, domandando da qual principio dependa questo *moto circolare* de' gravi e de' leggieri,- Giorn. 2.618*

*45) \SIMP.\... circolarmente, né anco il globo terrestre si muoverà di *moto circolare*; e così avremo l'intento. - Giorn. 2.622*

*46) \SALV.\... la Terra non abbia principio né esterno né interno al *moto circolare*, ma dico che non so qual de' dua ella - Giorn. 2.623*

*47) \SIMP.\... continua l'autor di domandar da qual principio dependa questo *moto circolare* de i gravi e de i leggieri, cioè se - Giorn. 2.640*

*48) \SALV.\... aria, a me basta che ella non impedisca il *moto circolare* de i mobili che per essa si dice che si - Giorn. 2.641*

*49) \SALV.\... è contrario al moto del fuoco, che sarà il *moto* deorsum: ma il *moto circolare*, che non è contrario né al sursum - Giorn. 2.651*

*50) \SALV.\... cadavero ed al vivente? E però, quando il *moto circolare* sia proprio degli elementi, dovrà esser comune de i - Giorn. 2.651*

*51) \SALV.\... quest'autore contro a chi dicesse, il principio del *moto circolare* de i gravi e de i leggieri esser un accidente - Giorn. 2.653*

*52) \SIMP.\... \SALV.\ Già mille volte si è detto che il *moto circolare* è naturale del tutto e delle parti, mentre sono - Giorn. 2.675*

*53) \SALV.\... non si muovon di moto retto, ma di un *moto* misto che anco potrebb'esser *circolare* schietto; - Giorn. 2.675*

*54) \SALV.\... credere che a quelli che dicon, tal *moto* non esser altrimenti retto, anzi più tosto *circolare*, paia di veder - Giorn. 2.728*
*55) \SALV.\... restando uno de' suoi estremi senza mutar luogo, il *moto* non può esser se non *circolare*: - Giorn. 2.735*
*56) \SALV.\... non lo separa dall'altro suo conterminale, adunque tal *moto* è *circolare* di necessità. - Giorn. 2.735*
*57) \SAGR.\... voi che quest'autore mettesse maggior diversità, tra il *moto* retto e 'l *circolare*, o tra il *moto* e la quiete? - Giorn. 2.751*
*58) \SIMP.\... quiete sicuramente. E quest'è manifesto; perché il *moto circolare* non è contrario al retto per Aristotile, anzi e' - Giorn. 2.752*
*59) \SAGR.\... un corpo naturale due principii interni, uno a 'l *moto* retto e l'altro al *circolare*, - Giorn. 2.753*
*60) \SAGR.\... principio interno stia immobile, e quella gli attribuisce il *moto circolare*: ma per la vostra concessione - Giorn. 2.753*
*61) \SALV.\... cielo. Quanto poi alla Luna, questa ha un *moto circolare* intorno alla Terra, dalla quale - Giorn. 3.155*
*62) \SALV.\...ed a Mercurio con Venere avete attribuito il *moto circolare* intorno al Sole senza abbracciar la Terra: - Giorn. 3.171*
*63) \SIMP.\... non se gli può negare, se gli attribuisce un *moto circolare* in un gran cerchio intorno al Sole in un anno - Giorn. 3.336*
*64) \SAGR.\... della Terra, come grave; il secondo è il *moto circolare*orizontale, per il quale restituisce e conserva il suo - Giorn. 3.370*
*65) \SAGR.\... con due moti retti semplici voi non comporrete mai un *moto circolare*, quali sono li due o i tre circolari diversi - Giorn. 3.378*
*66) \SAGR.\... se volete mantenere che 'l *moto* retto sia solo de gli elementi, e 'l *circolare* de' corpi celesti. - Giorn. 3.378*

4 Occorrenze ne *Il Saggiatore*:

*1) di più di 90 gradi, ei dà luogo al *moto* non *circolare*, ed ammette quello per linea ovata, anzi pur - 10.5*
*2) un cerchio massimo; e che in se stesso quel *moto* può esser fatto per linea*circolare*, ed anco per qual si voglia altra quanto si - 10.6*
*3) stima che degli altri due, presi l'uno dal *moto *circolare* e l'altro dalla piccolezza della paralasse, li quali - 14.3*
*4) all'orbe lunare, e quivi poi cangiare il suo *moto* retto in *circolare*? E come fa il Sarsi a sostenere per impossibil - 23.3*

<u>Espressione "moto retto"</u>

52 Occorrenze nel *Dialogo sopra i due massimi sistemi del mondo*:

*1) \SAGR.\... ma di già abitato da noi. Che se il *moto retto* è semplice per la semplicità della linea retta, e - Giorn. 1.28*
*2) \SAGR.\... come è detto, moti semplici e *moto* misto, chiamando semplici il circolare ed il *retto*, e misto il composto; Giorn. 1.28*
*3) \SAGR.\... semplici, ed a' composti il composto: ma per *moto* composto e' non intende più il misto di *retto* e circolare, - Giorn. 1.28*

*4) \SAGR.\... torniamo ad Aristotile, il qual mi definì, il *moto* misto esser quello che si compone del *retto* e del circolare; - Giorn. 1.32*

*5) \SALV.\... il quale sia ancora tanto più eccellente, quanto il *moto* circolare è più perfetto del *moto retto*: - Giorn. 1.33*

*6) \SALV.\... facile e manifesta. Imperocché quello che si muove di *moto* retto, muta luogo; e continuando di muoversi, si - Giorn. 1.35*

*7) \SALV.\... di mutar luogo, ed in conseguenza di muoversi di *moto retto*. In oltre, essendo il *moto retto* di sua natura infinito, perché infinita e indeterminata è la linea retta - Giorn. 1.35*

*8) \SALV.\... che in loro resti naturale inclinazione di più muoversi di *moto retto*, dal quale ora solo ne seguirebbe il - Giorn. 1.35*

*9) \SALV.\... cioè il disordinarsi. Possiamo dunque dire, il *moto retto*servire a condur le materie per fabbricar l'opera, - Giorn. 1.35*

*10) \SALV.\... furon per alcun tempo dal suo Fattore mossi di *moto retto*, ma che dopo l'esser pervenuti in certi e - Giorn. 1.35*

*11) \SALV.\... rivolti a uno a uno in giro, passando dal *moto *retto* al circolare, dove poi si son mantenuti e tuttavia - Giorn. 1.35*

*12) \SALV.\... per alcun tempo e per qualche spazio, di *moto retto*. Stante questo discorso, figuriamoci aver Iddio creato il - Giorn. 1.35*

*13) \SALV.\... Platone dire che gli desse di muoversi da principio di *moto retto* ed accelerato, e che poi, giunto a quel - Giorn. 1.35*

*14) \SALV.\... a quel tal grado di velocità, convertisse il suo *moto retto* in circolare, del quale poi la velocità naturalmente - Giorn. 1.35*

*15) \SALV.\... è moto circolare intorno al centro: adunque il *moto* circolare non s'acquisterà mai naturalmente senza il *moto retto* precedente, ma bene, acquistato che e' si sia - Giorn. 1.77*

*16) \SALV.\... ritrovare: e dico naturalmente, perché il *moto retto* che si ritarda, è il violento, che non - Giorn. 1.87*

*17) \SIMP.\... assegnato dall'istessa natura per fine e termine del *moto retto* deorsum; e non vegga parimente, muoversi - Giorn. 1.88*

*18) \SALV.\... ed in questo modo concludere che 'l *moto retto* competa egualmente a tutti i corpi mondani? - Giorn. 1.90*

*19) \SIMP.\... ora pronunziate (già che mettete in dubbio insino nel *moto* de i gravi se sia *retto* o no), come potete voi mai ragionevolmente negare che - Giorn. 1.91*

*20) \SIMP.\... che le materie gravissime, descendano verso il centro con *moto retto*, se, lasciate da una altissima torre, le - Giorn. 1.91*

*21) \SIMP.\... non è questo argomento più che evidente, cotal *moto* esser *retto* e verso il centro? Nel secondo luogo, voi - Giorn. 1.91*

*22) \SALV.\... lasciata per ora questa general considerazione, se il *moto retto* sia necessario in natura - Giorn. 1.121*

*23) \SALV.\... impossibile, perché naturale delle parti è il muoversi di *moto retto all'ingiù. - Giorn. 2.100*

*24) \SIMP.\... luogo, mentre dice che naturale delle parti è il *moto retto* al centro dell'universo, onde il circolare non gli - Giorn. 2.101*

*25) \SIMP.\... Il *moto retto* è naturalissimo delle parti della Terra e gli è eterno - Giorn. 2.107*

*26) \SIMP.\... e gli è eterno, né mai accaderà che di *moto retto* non si muovano, intendendo però sempre, rimossi gli - Giorn. 2.107*

*27) \SALV.\... esser impossibile che mobile alcuno si muova eternamente di *moto retto*, essendo che il *moto retto*, o vogliatelo in su o vogliatelo in giù, - Giorn. 2.118*

*28) \SALV.\... sia eterna, tuttavia, per non essere il *moto retto* di sua natura eterno, ma terminatissimo, non può - Giorn. 2.118*

*29) \SALV.\... stimato impossibile, che 'l sasso potesse muoversi di un *moto* misto di *retto* e di circolare; perché quando e' non avesse avuto - Giorn. 2.147*

*30) \SAGR.\... dall'albero, e sia in oltre vero che questo *moto* non arrechi impedimento o ritardamento al *moto retto* all'ingiù, naturale alla pietra, è forza che - Giorn. 2.240*

*31) \SALV.\.... Ma perché il *moto retto* del grave cadente è continuamente accelerato, è forza che - Giorn. 2.305*

*32) \SALV.\... descendente, partendosi dalla quiete, cioè dalla privazion del moto a basso, ed entrando nel *moto retto* in giù, è forza che passi per tutti i - Giorn. 2.305*

*33) \SALV.\... cadendo dalla sommità della torre C, venga movendosi del *moto* composto del comune circolare e del suo proprio *retto*. Imperocché, segnando nella circonferenza C D alcune parti - Giorn. 2.306*

*34) \SALV.\... della torre, che è quello che fa che il *moto retto* fatto lungo la torre ci si mostra sempre più e - Giorn. 2.306*

*35) \SAGR.\... e questa è, che stanti queste considerazioni, il *moto retto* vadia del tutto a monte e che la natura mai - Giorn. 2.311*

*36) \SIMP.\... Sarà un *moto retto* e perpendicolare, essendo la canna drizzata a perpendicolo. - Giorn. 2.360*

*37) \SALV.\... dirittura, non essendo nella linea circolare parte alcuna di *retto*. \SIMP.\ Io non intendo che 'l *moto* proietto sia a dirittura di tutto il circolare, ma - Giorn. 2.420*

*38) \SALV.\... che non fa angolo nessuno con la linea del *moto retto* fatto per la canna. - Giorn. 2.423*

*39) \SALV.\... mentre fu co 'l proiciente, e se da questo *moto* circolare deve passar al *moto retto*, qual dovrà esser questa linea retta? -Giorn. 2.425*

*40) \SALV.\... il mobile dopo la separazione, nel continuar il suo *moto retto*, si va sempre allontanando egualmente dal centro - Giorn. 2.437*

*41) \SALV.\... alcuni altri amici suoi, dimostra come l'accelerazione del *moto* retto de i gravi si fa secondo i numeri impari - Giorn. 2.571*

*42) \SALV.\... né ordinate né disordinate, non si muovon di *moto retto*, ma di un *moto* misto che anco potrebb'esser circolare schietto; - Giorn. 2.675*

*43) \SALV.\... noi resta visibile e osservabile una parte sola di questo *moto* misto, cioè la parte del *retto*, restandoci l'altra parte del circolare impercettibile, perché - Giorn. 2.675*

*44) \SIMP.\... vista in cosa tanto chiara s'inganna, e quel *moto* non è altrimenti *retto*, ma misto di *retto* e circolare. - Giorn. 2.697*

*45) \SALV.\... tanto in volerci far comprender co 'l senso, questo *moto* de i gravi descendenti esser semplice *retto* e non di altra sorte, - Giorn. 2.728*

*46) \SALV.\... indizio di credere che a quelli che dicon, tal *moto* non esser altrimenti *retto*, anzi più tosto circolare, paia di veder sensatamente - Giorn. 2.728*

*47) \SAGR.\... l'occhio semplice non si possa ingannare nel giudicar il *moto* retto de' gravi descendenti, e vuol che e' si inganni - Giorn. 2.751*

*48) \SAGR.\... voi che quest'autore mettesse maggior diversità, tra il *moto* retto e 'l circolare, o tra il *moto* e la quiete? - Giorn. 2.751*

*49) \SIMP.\... quiete sicuramente. E quest'è manifesto; perché il *moto* circolare non è contrario al *retto* per Aristotile, - Giorn. 2.752*

*50) \SAGR.\... un corpo naturale due principii interni, uno a 'l *moto retto* e l'altro al circolare, che due, pur - Giorn. 2.753*

*51) \SIMP.\... tra di loro molto differenti; poiché, oltre al *moto retto*, come grave, verso il centro, che non - Giorn. 3.336*

*52) \SAGR.\... elementari e di celesti, se volete mantenere che ‘l *moto retto* sia solo de gli elementi, e ‘l circolare de’ - Giorn. 3.378*

25 Occorrenze ne *Il saggiatore*:

*1) ... se bene era vero che il *moto* per cerchio massimo sempre appariva *retto*, non era però necessariamente vero il converso, cioè - 10.2*

*2) ...era però necessariamente vero il converso, cioè che il *moto* che apparisse *retto* fusse per cerchio massimo, - 10.2*

*3) come venivano ad aver supposto quegli che dall’apparente *moto retto* inferivano, la cometa muoversi per cerchio massimo: tra - 10.2*

*4) assai maggior mancamento è stato il lasciar senza considerazione il *moto retto*, poi che pur v’era il Kepplero che attribuito - 10.4*

*5) manchevole la scusa del Sarsi, perché non solamente il *moto* veramente *retto* apparisce per linea retta, ma qualunque altro, - 10.5*

*6) nominate stelle.” Concludiamo per tanto che dall’apparirci un *moto retto* altro non si può concludere salvo che l’esser fatto - 10.6*

*7) cioè il signor Mario, abbia attribuito alla cometa il *moto retto*, e poi, tre versi più a basso, - 10.8*

*8) basso, dice non esser bisogno alcuno d’escluder questo *moto retto*, il qual era certo e manifesto già mai non - 10.8*

*9) il quale non considerò il potersi ella muover di *moto retto*; e s’egli scusa il suo Maestro col dire - 11.2*

*10) trasgredì la legge: dico nell’inferir dall’apparenza del *moto retto* la circolazione per cerchio massimo, potendo esser del medesimo - 12.9*

*11) potendo esser del medesimo effetto causa il movimento realmente *retto* e qualunque altro *moto* fatto nell’istesso piano dove - 12.9*

*12) il Sarsi che noi con risolutezza abbiamo affermato, il *moto* della cometa dover necessariamente esser *retto* e perpendicolare alla superficie terrestre: - 19.4*

*13) non risolutezza, ma probabilità si è attribuito il *moto retto* in su alla medesima materia. E questo sia detto - 20.3*

*14) all’orbe lunare, e quivi poi cangiare il suo *moto retto* in circolare? E come fa il Sarsi a sostenere - 23.3*

*15) movendosi la cometa di semplice *moto retto*, fusse necessario ch’ella andasse sempre verso il vertice - 28.3*

*16) noi diciamo, che se la cometa si movesse di *moto retto*, ci apparirebbe muoversi verso il vertice e zenit, - 28.4*

*17) che dove quello dice, che o bisogna rimuovere il *moto retto* attribuito alla cometa, o vero, ritenendolo, aggiungere - 28.4*

*18) tuttavia la verità è, che segnati nel *moto retto* perpendicolarmente ascendente molti spazii eguali, i movimenti apparenti, - 31.3*

*19) signor Lottario, che può star benissimo in un istesso *moto retto* ed uniforme un’apparente diminuzione e grande e mezana e - 31.3*

*20) moto apparente, provare, il già più volte nominato *moto *retto* non competere in verun modo alla cometa (e dico - 31.4*

*21) fa’ tu ora concetto e tieni per sicuro che il *moto retto* del signor Mario in veruna maniera se gli assesta, - 31.4*

*22) il qual già apertamente ha scritto che un semplice *moto retto* non può bastare a soddisfare all’apparente mutazion della cometa - 31.4*

*23) mutazion di luogo che fece la cometa provar che ‘l *moto retto* del signor Mario non gli poteva competere, perché la - 31.5*

*24) piccola: e perché la verità è che a questo *moto retto* ne possono seguir mutazioni piccole, mediocri ed anco grandissime - 31.5*

*25) e mezo che poteva importar l'altra dependente dal proprio *moto retto*, tuttavia noi rimagniamo assai lontani da quel moto grandissimo che in lei si vide. - 32.2.*

# References

1. Bruno Giordano (1584). *Cena de le ceneri.* Anche online: https://www.liberliber.it/mediateca/libri/b/bruno/cena_de_le_ceneri/pdf/cena_d_p.pdf.
2. Caffarelli Vergara Roberto (2006). "Il principio d 'inerzia negli ultimi scrilli di Galilei". *«Cronos», 10*, 63-88.
3. Caffarelli Vergara Roberto (2009a). *Galileo Galilei and Motion. A Recontruction of 50 Years of Experiments and Discoveries*, Bologna 2009, (pubblicato insieme da SIFe SPRINGER).
4. Caffarelli Vergara Roberto (2009b). *Prove di Funzionamento della Macchina di Galilei*, in internet all'indirizzo: http://www.illaboratoriodigalileogalilei.it, nella pagina "*multimedia*", aprendo "*video1993-2009*".
5. Caffarelli Vergara Roberto (2010a), *La Macchina di Galileo Galilei*, https://www.vergaracaffarelli.it/styled-14/files/2005-la-macchina_di-galileo.pdf
6. Caffarelli Vergara Roberto (2010b). *La Macchina di Galilei. Gli esperimenti del 7 novembre 2005,*

*https://www.illaboratoriodigalileogalilei.it/ProveMacchinaGalilei.pdf*

7. Cohen I. B. (1967). "Newton's attribution of the first two laws of motion to Galileo". In: *Atti del Symposium Internazionale di Storia, Metodologia,Logica e Filosofia della Scienza "Galileo nella Storia e nella Filosofia della Scienza"*, Firenze, pp. XXV- XLIV.
8. Drago Antonino (1997). "La nascita del principio d'inerzia in Cavalieri e Torricelli secondo la matematica elementare di Weyl". In *Atti del XVII Congresso nazionale di storia della fisica e dell'astronomia*: Istituto di fisica generale applicata, Milano, Centro Volta, Villa Olmo, Como, 22-25 maggio 1997 / a cura di Pasquale Tucci.
9. Einstein Albert (1962). *Relativity,* in *The American Peoples Encyclopedia*, New York, Grolier Incorporated, p. 16-360.

10. Espinoza Fernando (2005). “An analysis of the historical development of ideas about motion and its implications for teaching”. In «*Physics Education*» 40 (2), p. 141.

11. Favaro Antonio (cur.) (1890-1909). *Le Opere di Galileo Galilei*, *Edizione Nazionale*, 1890-1909 , Voll. 20. Ristampe: 1929-1939 e 1964-1968. Firenze: Barbera.

12. Klein Etienne (2006). *Sette volte la rivoluzione*. Milano: Raffaello Cortina.

13. Koyré Alexandre (1939). *Ètudes galiléennes* (trad. it. *Studi galileiani*. Torino: Einaudi, 1976).

14. Newton Isaac (1687, 2001). *Philosophiæ Naturalis Principia Mathematica*. Londini: Jussu Societatis Regiæ ac Typis Josephi Streater. Prostat apud plures Bibliopolas. Anno MDCLXXXVII. In italiano *Principi di filosofia naturale*.Trad. italiana e riduzione a cura di Federigo Enriques e Umberto Forti. Milano: Fabbri editore, 2001.

15. Newton Isaac (1728). *Philosophiæ Naturalis Principia Mathematica*. Londini: translated by Andrew Motte.

16. Nicotra Luca (2022). “The authorship of the Principle of Inertia”. Parte I. «*Science & Philosophy*» Vol. 10 (1) pp. 81-110.

17. Panofsky Erwin (1956). “Galileo es a critic of the arts: aesthetic attitude and scientific thought”, in «*Isis*», XLVII, 1956, pp. 3-15.

18. Rossi Paolo (2006). “Galileo Galilei”. In *Storia della scienza* (a cura di P. Rossi). Roma: Gruppo Editoriale L’Espresso, , vol. 1°.

# The influence of social capital on health issues among transgender and gender diverse people: a rapid review

Davide Costa*

**Abstract**

This article aims to analyze the current literature on the social capital of transgender and gender diverse(TGD) people, given their fragility in social and health terms. The paper followed the guidelines developed by Tricco, Langlois, and Straus. The results of this paper reveal significant gaps in the literature relating to the social capital of TGD people and highlight how the various types of shared capital are for sexual health to be considered in future research on transgender health. This is the first article that analyzes in detail the relationship between social capital and TGD individuals. To date, there is no other scientific evidence in the literature in this regard. The paucity of scholarly evidence available for paper limits our ability to make conclusive statements about social capital of TGD people. Small sample sizes in the included studies warrant caution when deriving generalized conclusions about social capital.

**Keywords:** social capital, transgender and gender diverse people, health, review

* University "Magna Graecia" of Catanzaro, Italy; davide.costa@unicz.it.

Received on October 26, 2022. Accepted on December 20, 2022. Published on December 30, 2022. doi: 10.23756/sp.v10i1.834. ISSN 2282-7757; eISSN 2282-7765. 

# 1. Introduction

Research on the health of sexual and gender minorities (SGM), and in particular on transgender and gender diverse(TGD) individual, is very recently established (LeBlanc and Perry, 2021). This trend has established the following social, political, media, and legal activities on the rights and protection of LGBTQA+.

However, before continuing, is it good to ask who is TGD individual? To date, there are several definitions, in general, for this work, and I used the definition proposed by the World Professional Association of Transgender Health (WPATH): "transgender and gender diverse (TGD) to be as broad and comprehensive as possible in describing members of the many varied communities globally of people with gender identities or expressions that differ from the gender socially attributed to the sex assigned to them at birth. This includes people who have culturally specific and/or language-specific experiences, identities, or expressions, and/or that are not based on or encompassed by Western conceptualizations of gender, or the language used to describe it" (Coleman et al. 2022, p.3). The first research on the health of TGD people and other SGM began in the 1950s, in a cultural context that considered this gender variability as real crimes against nature and disease (Henry, 1948; Hirschfeld, 1948). Thus, SGM become the exclusive object of medicine.

In the 1970s, on the other hand, the process of demedicalization began thanks also to the activism that stood out in those years, although the research on health of TGD people still remained poor (LeBlanc and Perry, 2021). In the 1980s, on the other hand, the interest in research on the health of SGM health grew rapidly due to the spread of the HIV / AIDS epidemic. In those same years, a small number of researchers began to show interest in the relationship between TGD health and their social system of reference made up mainly of harassment and stigmatization (Hunter and Schaecher, 1987; Martin, 1982; Ross-Reynolds and Hardy, 1985). From these pioneering studies onwards, it has been documented, albeit with a residual number, that the TGD population is lacking in social relations, due to the stigma and nonacceptance of collectivities, and this is an incisive factor in their fragile health state. Regarding the fragility of health of TGD people, it is recalled that, compared to the cisgender population, they have a greater predisposition to different types of pathologies such as: suicidality (Cochran, Sullivan, and Mays, 2003; King et al., 2008), substance use disorders (King et al., 2008), obesity (Boehmer, Bowen, and Bauer, 2007), hypertension (Fredriksen-Goldsen, Kim, Barkan, Muraco, & Hoy-Ellis, 2013), cardiovascular disease (Fredriksen-Goldsen et al., 2013), Type II diabetes (Beach, Elasy, and Gonzales, 2018), chronic pain (Fredriksen Goldsen et al., 2017), cancer (Stinchcombe, Wilson, Kortes-Miller, Chambers, & Weaver, 2018), acquired hypothyroidism, COPD, PTSD, schizophrenia, diabetes, asthma, obesity, personality disorders, rheumatoid arthritis/ osteoarthritis, psychotic disorders, fibromyalgia, chronic pain and fatigue, anemia, tobacco use disorders, hyperlipidemia, anxiety disorders, major depressive affective disorders(Hardacker et al., 2018). Therefore, all of these diseases are also caused by the "loss of support networks, isolating the individual from their chosen family, the family of origin, colleagues, friends, and health system supports" (Hardacker et al., 2018,p.100). Thus, if “Social relationships matter for health and well-being” (LeBlanc and Perry, 2021, p. 136), experiences of stigma,discrimination, and isolation negatively affect the quality of life of TGD people (Petruzzella, Feinstein and Lavner, 2019).

Although much research has shown how large and toothless social networks are able to improve access to different types of social resources and mediate better psychophysical health (Kim, Fredriksen- Goldsen, Bryan and Muraco, 2017), in the case of TGD people there is very little literature on this subject. It is important to remember that there is another way to define social relations and their relevance to health, namely, the so-called concept of social capital (Coleman 1988, Putman, Leonardi and Nanetti, 1993).

The notion of social capital arises in the sociological, political, and economic fields, to provide explanations for the cobehavior of cooperating citizens in society (Kawachi e Kennedy, 1997), but for some time "Public health researchers have turned to the concept of social capital to explain the heterogeneity of the population state of health in all geographical areas" (Lochner et al.,1999, p.259). In general, social capital could be defined as the set of some characteristics of social life such as: social networks, norms, and trust that allow us to achieve common goals (Putnam 2000). According to the review of the literature on social capital proposed by Aldrige et al. (2002), it is possible to distinguish a real taxonomy of social capital:bonding social capital is the type of social capital that refers to parental and ethnic ties, i.e., strong ties (Banfield 1958, Gittel and Vidal, 1998); bridging social capital, these are the weakest social ties, and therefore attributable to relationships with acquaintances and friends (Granovetter, 1973, 1985) and linking social capital: concerning the connections between individuals belonging to different social levels (Woolcock, 2001).

Moreover, another theorist of social capital, Pierre Bourdieu (1977,1986), argues that it is a set of resources of a societal type, and therefore reflects the set of social inequalities that actually structure it (Bourdieu 1986). Thus, according to this approach, central is the power behind social capital, which varies according to the stability and location of social relations. Social capital turns out to be important for health in general, with a sociological approach.

Already many centuries ago, Emile Durkheim had shown how integration and social cohesion had relevance in pathologies in psychiatric disorders such as suicide (Durkheim, 1897).

Therefore, based on what has been reported so far, this review was guided by the following question research: how has the social capital of TGD people been studied in the literature? More specifically, this article aims to analyze the current literature on the social capital of TGD people, given their fragility in social and health terms.

# 2. Methods

A rapid review was conducted to produce a timely, contextualized, and in-depth synthesis of the current literature on the social capital of TGD. The review followed the guidelines developed by Tricco, Langlois, and Straus (2017).

The search was conducted in April-May 2022 and was performed in three databases: MEDLINE, Scopus and Web of Science. This study focuses on the keyword "social capital and transgender and gender divers people" and the research was realized with the use of two rows of this filter: article title, abstract, keywords.

Qualitative and quantitative studies, published without time limits, were included. To include the various studies, they must have studied directly, both with quantitative, qualitative, or mixed methods and in a direct way on a semantic and empirical level, the concept of social capital on TGD people regardless of gender or age.

To represent the population of interest, inclusion criteria were to all types of subjects that fit the definition of TGD (including nonbinary) and of all ages. Studies were excluded, however, whether the concept of social capital was or was only briefly hinted at or explored but without technical terminology; and those studies whose results were presented in an unclear and dispersive way with respect to the initial objectives were also excluded; besides excluded articles not dealing with TGD people.

I started with 199 studies, 4 duplicate studies were removed, and the remaining 195 studies were screened for relevance by me at the title and abstract level. Upon reviewing titles and abstracts, 103 studies were excluded because they did not meet the inclusion criteria (e.g., wrong population, setting, methods, or research question). 92 studies were subsequently screened at the full-text level by three reviewers and a final sample of 9 articles was included (Figure 1) and charted for data extraction.

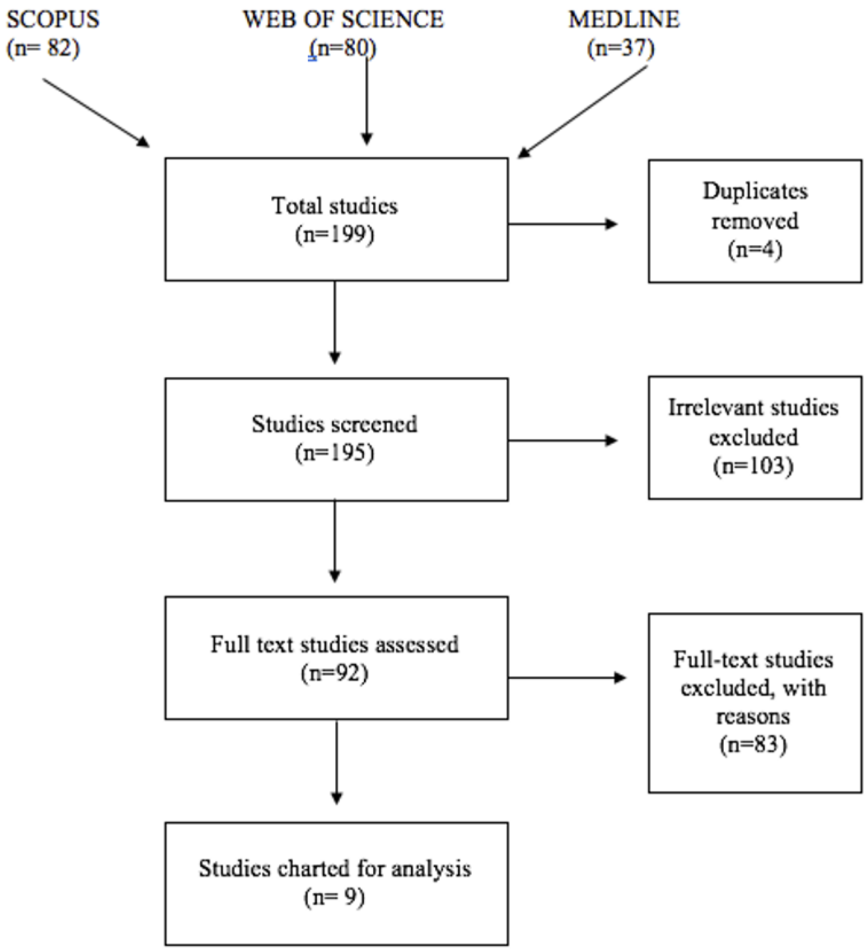



Figure 1. Flowchart outlining stager of the rapid review process and number of sources retrived and selected at each stage.

The data extraction chart contained fields for reference information, region where the study occurred, purpose, methods, samples, and results relating to the social capital of TGD people.

# 3. Findings

The nine included studies were published between 2008 and 2022. Included studies are summarised in Table 1.

| **Author(s)** | **Years** | **Main Purpose** | **Age** | **Study design** | **Methods** | **Contest of research** | **Country/City** |
|---|---|---|---|---|---|---|---|
| 1. Rogério M. Pinto, Rita M. Melendez & Anya Y. Spector | 2008 | To describe how a sample of urban racial and ethnic minority MTFs use their gendered social networks to develop social capital and engage in political action. | From 18 to 53 | Qualitative | Interview | Community -based health care clinic | USA-New York |
| 2. Sara Green-Hamann & John C. Sherblom. | 2014 | To compare and analyze differences in the social capital and support communication provided by Alcoholics Anonymous,cancer caregivers, and transgender identity support groups. | Not reported | Quantitatie | Comparative analysis | Second Life (virtual world) | USA |
| 3. Elena A. Erosheva, Hyun-Jun Kim. Charles Emlet, and Karen I. Fredriksen-Goldsen. | 2015 | This study examines global social networks—including friendship, support, and acquaintance networks—of lesbian, gay, bisexual, and transgender (LGBT) older adults. | From 50 to 80+ | Quantitatie | Survey | Community agencies | USA |
| 4. Andrew King and Ann Cronin | 2015 | The purpose of this paper is to contribute to debates about lesbian, gay, bisexual and transgender (LGBT) housing later in life by placing these in a theoretical context: social capital theory. | Older people | Review | Narrative review | Scientific literature | USA |
| 5. Amaya G. Perez-Brumer, Sari L. Reisner, Sarah A. McLean, | 2017 | To understand social capital as a social determinant of health and examines its relationship to HIV vulnerabilities | From 18 to 44 | Qualitative | Focus group | Community Task Force | USA-Lima |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Alfonso Silva-Santisteban, Leyla Huerta, Kenneth H. Mayer, Jorge Sanchez, Jesse L. Clark Matthew J. Mimiaga, and Javier R. Lama. | | among TW in Peru. | | | | | |
| 6. Sel J. Hwahnga,Bennett Allenc , Cathy Zadoretzkyd, Hannah Barber Doucete , Courtney McKnightf and Don Des Jarlaisf | 2021 | To analyzed social capital and health outcomes among transwomen of color | From 23 to 50 | Qualitative | Focus group | AIDS Institute of New York State Department | USA-New York |
| 7. Rayner Kay Jin Tan, Caitlin Alsandria O'Hara, Wee Ling Koh, Daniel Le, Avin Tan, Adrian Tyler Calvin Tan, Chronos Kwok, Sumita Banerjee and Mee Lian Wong. | 2021 | This paper attempts to explore the association between measures of social capital and patterns of sexualized substance use among a sample of YMSM in Singapore | form 18 to 25 | Quantitative | Prospective cohort study | Community organizations | China-Singapore |
| 8. Meagan Zarwell, Jennifer L. Walsh, Katherine G. Quinn, Andréa Kaniuka, Alexandra Patton, William T. Robinson and Robert J. Cramer. | 2021 | The aim of this paper is to create a new tool to measure social capital within social networks of sexual minority men and gender minority individuals. | From 18 to older | Quantitative | Survey | Pride Festival in Milwaukee | USA-Milwaukee |
| 9. Yuekang Li, Vanessa D. Fabbre & Eleni Gaveras | 2022 | This study uses the autobiographical life narratives of trans older adults | from 50 to over 80 | Qualitative | Biographical interview | Personal and professional networks of the co- | USA |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | to develop the concept of authenticated social capital, which can be used to diversify theorising about both trans older adults and the lives of other stigmatised and marginalised groups in an increasingly diverse 21st-century global society. | | | | creators. | |

Manual full-text review revealed that studies occurred across the USA(n=8) and only one in China(Singapore). Five studies focused on SGM, including and not just TGD people(Hamann and Sherblom, 2014; King and Cronin, 2015; Erosheva, Emlet, and Fredriksen-Goldsen., 2015; Zarwell et al. 2021; Tan et al. 2021). Four studies focused only transgender people (Pinto, Melendez and Anya, 2008; Perez-Brumer et al.,2017; Hwahnga et al., 2021; Li, Fabbre and Gaveras, 2022).

Only one study is a narrative review(King and Cronin, 2015). Regarding the age of the subjects included in the analyzed papers, 4 studies focused on an elderly population (Erosheva, Emlet, and Fredriksen-Goldsen., 2015; King and Cronin, 2015; Zarwell et al. 2021; Fabbre and Gaveras, 2022); on the other hand, in four studies the population was made up of young people and adults (Pinto, Melendez and Anya, 2008; Perez-Brumer et al.,2017; Hwahnga et al., 2021; Tan et al. 2021).

Regarding the research contexts, and therefore the area of recruitment, as can be seen in the table, most of the studies were conducted through the involvement of associations and / or healthcare contexts.

### Definitions of social capital used in the articles

First of all, in the nine papers included, different types of definitions of social capital were identified and reported in the following Table 2.

| **Author(s)** | **Definition of social capital used** |
|---|---|
| Rogério M. Pinto, Rita M. Melendez & Anya Y. Spector | "Social capital refers to sharing information and social values in networks that promote survival and access to resources (Bourdieu, 1986; Coleman, 1988)". |
| Sara Green-Hamann & John C. Sherblom. | "The concepts of social capital are used to describe the social support provided in groups, but little research connects that communication to the stressor for which the group is formed". |
| Elena A. Erosheva, Hyun-Jun Kim. Charles Emlet, and Karen I. Fredriksen-Goldsen. | "The concept of social capital involves a notion of social relations as an available resource. Social capital can be defined as a function of social structure—a system of social relations—producing advantage for individuals |

| | |
|---|---|
| | who are within that structure (Coleman, 1988). Social ties—with kin, partners, adult children, friends, neighbors, or with fellow members of organizations—constitute social capital of older adults that can give them access to social, emotional, and practical support (Gray, 2009)". |
| Andrew King and Ann Cronin | The authors used the classical definitions by Putnam and Bourdieu. |
| Amaya G. Perez-Brumer, Sari L. Reisner, Sarah A. McLean, Alfonso Silva-Santisteban, Leyla Huerta, Kenneth H. Mayer, Jorge Sanchez, Jesse L. Clark Matthew J. Mimiaga, and Javier R. Lama. | "Bonding social capital (, bridging social capital (inter-group relations) and linking social capital (relations with institutions of power) are needed to inform public health strategies". |
| Sel J. Hwahnga,Bennett Allenc , Cathy Zadoretzkyd, Hannah Barber Doucete , Courtney McKnightf and Don Des Jarlaisf | "Social capital as a group-level construct, versus an individual-level resource, in which social support is an important component. Social capital can further be defined as the social resources that evolve through social networks or structures characterized by mutual trust". |
| Rayner Kay Jin Tan, Caitlin Alsandria O'Hara, Wee Ling Koh, Daniel Le, Avin Tan, Adrian Tyler Calvin Tan, Chronos Kwok, Sumita Banerjee and Mee Lian Wong. | Social capital is a concept "that reflect access to social networks, psychological feelings of affiliation, or even explicit forms of affiliation that presuppose participation in interest groups or activities". |
| Meagan Zarwell, Jennifer L. Walsh, Katherine G. Quinn, Andréa Kaniuka, Alexandra Patton, William T. Robinson and Robert J. Cramer. | "The theory of social capital argues that social connections and networks are valuable to members of groups who may derive resources from within their networks". |
| Yuekang Li, Vanessa D. Fabbre & Eleni Gaveras | The concept of social capital is used "to explore the nature, role and value of social networks, connections and forms of community in the lives of older adults, and its importance for wellbeing in later life has received increased attention in social science and policy domains". |

It is possible to note that in most of the studies, the definitions and conceptualizations of Coleman(Pinto, Melendez and Anya, 2008; King and Cronin, 2015;), Bourdieu (Pinto, Melendez and Anya, 2008; King and Cronin, 2015; Zarwell et al. 2021; Li, Fabbre and Gaveras, 2022) and Putnam(King and Cronin, 2015; Zarwell et al. 2021; Tan et al. 2021; Hwahnga et al., 2021; Li, Fabbre and Gaveras, 2022) have been used, even if in many cases reinterpreted by the authors; being the highest theoretical references available in the literature to date.

Definitions and authors other than those mentioned can be found, for example, in the work by Hamann and Sherblom(2014)that they used two different authors for the definition of social capital by Lin(1999) that "Social capital is embedded in the social network of a group.

By participating in the group's social network, an individual gains access to certain informational and emotional resources that are helpful in responding to life's stresses. These resources provide a type of social capital.Support groups offer individuals social capital resources to help respond to their life stressors" (Hamann and Sherblom, 2014, p.1132); the second definition used by Hamann and Sherblom(2014) is that of Granovetter that "Strong-tie networks of close friends offer a bonding social capital"(Hamann and Sherblom, 2014, p.1133). Perez-Brumer et al. (2017) used the notion proposed by Woolcock (2004), which in fact originated from the re-elaboration of the above-mentioned "fathers" of the concept of social capital.

Regarding definitions, the most recent article by Li, Fabbre and Gaveras(2022), included in this article, contains two different definitions, one little known and the other entirely created by the authors. In the first case, for the authors, the best-known theories on social capital do not take into account the cultural and social systems inherent to gender issues, especially with respect to the distribution of power in society; this is why these researchers used Anselm Strauss's (1978)approach according to which trans genders experience "non-authentication" experiences.

In the second case, Li, Fabbre and Gaveras(2022) proposed a new type of social capital specific for the TGD social capital analysis, the “authenticated social capital”. This type of social capital is that blends the main theoretical elements of social capital with the aim, however, of overcoming the classic vision of this concept which “often relies on cisheteronormative norms and expectations, while authenticated social capital aims to resist these” (p.13).

**The main methods used to study the social capital of TGD people**

In the eight included articles, except the narrative review by King and Cronin (2015), it was possible to detect a certain variability of research methods and techniques used for the study of social capital of TGD people.

In table 3 provides a general overview of the search tools used in the articles included.

| Author(s) | Study design | Method and instrument used |
|---|---|---|
| Rogério M. Pinto, Rita M. Melendez & Anya Y. Spector | Qualitative | Community-based participatory research. The instrument use was semistructured interview according the approach by "Lincoln and Guba (1985)"based on grounded theory |
| Sara Green-Hamann & John C. Sherblom. | Quantitative | Comparative analysis of optimal matching and social capital influences. The theoretical approach used was optimal matching model. |
| Elena A. Erosheva, Hyun-Jun Kim, Charles Emlet, and Karen I. Fredriksen-Goldsen. | Quantitative | Survey. The social network measurement are typically used in social network studies (Burt, 1984); also it was used and modified the index to estimate the network size of LGBT individuals; the diversity index was used to study Network diversity; while to measure Family relations, identity disclosure, religious activity, and service utilization, scales built by the authors were used. |
| Amaya G. Perez-Brumer, Sari L. Reisner, Sarah A. McLean, Alfonso Silva-Santisteban, Leyla Huerta, Kenneth H. Mayer, Jorge Sanchez, Jesse L. Clark Matthew J. Mimiaga, and Javier R. Lama. | Qualitative | Focus group events. The authors used the community-level perspective. For qualitative analysis was conducted using Dedoose. |

| | | |
|---|---|---|
| Sel J Hwahnga,Bennett Allenc,Cathy Zadoretzkyd, Hannah Barber Doucete, Courtney McKnightf and Don Des Jarlaisf | Qualitative | Focus groups events. The authors used the thick and thin trust approach. For the analysis was used thematic approach to analyzing the data (Guest et al., 2012). |
| Rayner Kay Jin Tan, Caitlin Alsandria O'Hara, Wee Ling Koh, Daniel Le, Avin Tan, Adrian Tyler Calvin Tan, Chronos Kwok, Sumita Banerjee and Mee Lian Wong. | Quantitative | Prospective cohort study. To study was used Personal social capital(PSCS-16); to describe the perceived quantity and quality of their relationships with individuals and organizations, it is used a Likert scale. Connectedness to the LGBT community was an eightitem scale adapted from Frost and Meyer. Outness was measured t with the scale developed by Mohr and Fassinger |
| Meagan Zarwell, Jennifer L. Walsh, Katherine G. Quinn, Andréa Kaniuka, Alexandra Patton, William T. Robinson and Robert J. Cramer. | Quantitative | Crosssectional survey. For the study, the authors have modified an instrument that measured social capital within constructed families of GBMSM, the Constructed Family Social Capital Scale by Zarwell and Robinson. |
| Yuekang Li, Vanessa D. Fabbre & Eleni Gaveras | Qualitative | Interpretive content analysis. The instrument used was in-depth interview. The authors used: interpretive content analysis (Drisko and Maschi 2016),; twocircle coding process. Compared to content analysis techniques that rely solely ondeductive coding and quantitative measures (Baxter 1991), interpretive content analysis facilitates additional attention to meaning-making and subjectivity in qualitative data (Drisko and Maschi 2016). |

First of all, it is possible to note that there is a parity of qualitative and quantitative study designs. About this: four studies used a qualitative design study (Pinto, Melendez and Anya, 2008; Perez-Brumer et al.,2017; Hwahnga et al., 2021; Li, Fabbre and Gaveras, 2022); four studies used a quantitative study design(Hamann and Sherblom, 2014; Erosheva, Emlet, and Fredriksen-Goldsen., 2015; Tan et al. 2021; Zarwell et al. 2021).

Qualitative approaches were based on the use of instruments such as: focus group (Perez-Brumer et al.,2017; Hwahnga et al., 2021); interview (Pinto, Melendez and Anya, 2008; Li, Fabbre and Gaveras, 2022).

Quantitative approaches were based on questionnaires. From the analysis of these articles with quantitative approaches, two studies (Erosheva, Emlet, and Fredriksen-Goldsen. 2015; Tan et al., 2021; Zarwell et al. 2021) interesting data have been revealed, united by the fact that some instruments, or parts of them, have been modified to specifically study the social capital of SGM. Erosheva, Emlet, and Fredriksen-Goldsen (2015) for the analysis of an aspect of social capital, that is, the social network size, they have modificated the index to evaluate the dimensions of LGBTQA+ people.

In fact, the index they have modified is similar to that used to analyze the size of social networks, i.e., according to the sum method (McCarty, Killworth, Bernard, Johnsen and Shelley, 2001). It is an indicator that is based on the amount of social contact (with friends, relatives, etc.).

"The modified summation index uses groups defined by sexual identity, gender identity, and age because these groups reflect the basic composition of the population of interest better than typical relational categories. It has been shown that the summation method yields a valid and reliable proxy for the actual network size" (Erosheva, Emlet, and Fredriksen-Goldsen, 2015, p.8).

Zarwell et al., 2021, have modificated the Constructed Family Social Capital Scale by Zarwell and Robinson (2018), an instrument that merges social cohesion and the indicators used to study social networks, to measure the social capital of SGM, which belongs to built families.

Zarwell et al. (2021), with their modifications, they created the Network Social Capital Stairs, that which contains the same questions as the previous scale, with some modifications on questions relating to individuals who are part of the social networks of the study subjects, instead of concentrating alone on the family as in the starting scale.

Thus "this additional measure of compositional quality was added because SGMS face unique barriers to health care and discrimination in health care settings due to their marginalized identity and therefore may benefit from social network members who they can talk to about LGBTQA+ related healthcare"(Zarwell et al., 2021,p. 4).

### Types of social capital studied

Regarding the types of social capital of the TGD people that have been analyzed in the articles included are synthesized in Table 4.

| Author(s) | Bonding capital | Bridging capital | Linking capital | Trust |
|---|---|---|---|---|
| Rogério M. Pinto, Rita M. Melendez & Anya Y. Spector | X | | | |
| Sara Green-Hamann & John C. Sherblom. | X | X | | X |
| Elena A. Erosheva, Hyun-Jun Kim. Charles Emlet, and Karen I. Fredriksen-Goldsen. | | X | | |
| Andrew King and Ann Cronin | X | X | | X |
| Amaya G. Perez-Brumer, Sari L. Reisner, Sarah A. McLean, Alfonso | X | X | X | X |

| | | | | |
|---|---|---|---|---|
| Silva-Santisteban, Leyla Huerta, Kenneth H. Mayer, Jorge Sanchez, Jesse L. Clark Matthew J. Mimiaga, and Javier R. Lama. | | | | |
| Sel J. Hwahnga,Bennett Allenc , Cathy Zadoretzkyd, Hannah Barber Doucete , Courtney McKnightf and Don Des Jarlaisf | **X** | **X** | **X** | **X** |
| Rayner Kay Jin Tan, Caitlin Alsandria O'Hara, Wee Ling Koh, Daniel Le, Avin Tan, Adrian Tyler Calvin Tan, Chronos Kwok, Sumita Banerjee and Mee Lian Wong. | **X** | **X** | | **X** |
| Meagan Zarwell, Jennifer L. Walsh, Katherine G. Quinn, Andréa Kaniuka, Alexandra Patton, William T. Robinson and Robert J. Cramer. | **X** | | | **X** |
| Yuekang Li, Vanessa D. Fabbre & Eleni Gaveras | **X** | **X** | | **X** |

It is possible to see how bonding capital has been studied in all articles included (Pinto, Melendez and Anya, 2008; Hamann and Sherblom, 2014; Erosheva, Emlet, and Fredriksen-Goldsen, 2015; King and Cronin 2015, Perez-Brumer et al.,2017; Hwahnga et al., 2021; Tan et al. 2021; Zarwell et al., 2021; Li,Fabbre and Gaveras, 2022).

Bridging capital was studied in 8 articles on 9(Hamann and Sherblom, 2014; Erosheva, Emlet, and Fredriksen-Goldsen, 2015; Perez-Brumer et al.,2017; Hwahnga et al., 2021; Tan et al. 2021; Zarwell et al., 2021; ; Li,Fabbre and Gaveras, 2022).

Linking capital was studied in only two studies (Perez-Brumer et al.,2017; Hwahnga et al., 2021).

Trust was studied in 7 studies (Hamann and Sherblom, 2014; Perez-Brumer et al.,2017; Hwahnga et al., 2021; Tan et al. 2021; Zarwell et al., 2021; Li,Fabbre and Gaveras, 2022).

Three articles (Erosheva, Emlet, and Fredriksen-Goldsen, 2015; Tan et al., 2021; Hwahnga et al., 2021)found some subcategories of scapital. Erosheva, Emlet, and Fredriksen-Goldsen(2015) and Tan et al. (2021) proposed two subcategories of social capital: sociocentric network and egocentric networks.

The sociocentric (Erosheva, Emlet, and Fredriksen-Goldsen,2015) or group-centred(Tan et al. 2021) dimension concerns all relationships between people within a welldefined group such as a village and egocentric(Erosheva, Emlet, and Fredriksen-Goldsen, 2015) or individual-center(Tan et al. 2021) dimension regards the personal network.

Hwahnga et al. (2021), instead, proposed subcategories that relate to social trust(Putnam, 2000)"that were implicit within reciprocity, social norms, and civic engagement. Thick trust, as a component of bonding social capital, was thus "bolstered by dense networks of social exchange".

This trust, as a component of bridging or linking social capital, was directed at a "generalized " and depended on some level of social capital networks and expectations of reciprocity"(p.3).

**Main findings about social capital of TGD**

As regards the results achieved in all nine articles analyzed, first of all, they are fully in line with the general literature which considers social capital as a real health promoter (Pinto, Melendez and Anya, 2008; Hamann and Sherblom, 2014; Erosheva, Emlet, and Fredriksen-Goldsen, 2015; King and Cronin 2015;Perez-Brumer et al.,2017; Hwahnga et al., 2021; Tan et al. 2021; Zarwell et al., 2021; Li,Fabbre and Gaveras, 2022). King and Cronin (2015) regarding the importance of social capital, arguing that "Social capital counts" (p.22) for the health of SGM.

In detail, some results are particularly interesting. Pinto, Melendez and Anya (2008)found that the TGD people involved considered the clinic dedicated to their health treatment, as a real place for training and for the implementation of their social capital thanks to the various services provided by the clinic. Furthermore, the clinic also provided a further positive contribution to the social capital of the subjects studied, also thanks to the interaction with other TGD people, sharing advice, experiences, etc. on their path of building their identity (Pinto, Melendez and Anya, 2008).

In Hamann and Sherblom's (2014) study, however, it is clear that through the study of social capital, it is possible to understand how support groups, inherent to the problems of social inclusion and not only, of TGD people are promoters of social support in the growth of both bonding and bridging capital. Furthermore, and this is a very interesting data, according to Green, the TGD group is the most inclined to increase social capital because "The TI group communication builds group inclusion and social identity through the use of "we" statements, specialized language, and acronyms. Only in their expressions of emotional empathy and sympathy do the three groups look qualitatively similar in their support communication". A similar result, regarding this aspect related to Erosheva, Emlet, and Fredriksen-Goldsen's (2015) study, it is shown that TGD people have the largest composition of social capital, greater predisposition towards network diversity, greater ties with other SGMS as: lesbian, gay men and bisexual. Perez-Brumer et al.(2017) paper, on the other hand, in the results, underlines the need to make social capital a real measure, and tool, to improve and prevent the different forms of social exclusion that trangenders experience, not only at the level relational, but above all at the level of health services, with the aim of implementing the degree of acceptance and resilience of these, and not only, SGM.

In this regard, Zarwell et al. (2021), with their new scale, the “Network Social Capital Scale”, argued that not only does the new scale allow to improve the quality of measurement of social capital, but also to be able to demonstrate the unequivocally the salutary effect that social capital has not only on perceived health, but also on programming and management of health systems in the field of SGM.

The study of social capital by Tan et al. (2021) has allowed us to demonstrate how the low social capital of the SGM, dictated by nonacceptance, stigma, etc., causes, at an early age, the use of drugs, alcohol and above all early use of sex and various forms of prostitution.

The study by Hwahng et al. (2021) comes to register as the social capital, especially in the area of trust, in the dimension of "thick trust" (bond capital) and "thin trust" (bridge / connection capital) highly relevant to TGD people in their relationship with health and the health system.

Li,Fabbre and Gaveras (2022), with “their” new form of social capital, the “authenticated social capital”, argued that this type of social capital is composed of different dimensions and levels, so much so that it poses itself as a new paradigm for analyzing the social capital of SGM.

In particular, these authors argued that with this new form of social capital it can favor empowerment, and thus simplify self-affirmation and promote the transmission of knowledge, which ultimately challenges the marginalized identities, thus giving greater importance to the alternative social networks created by transoriented communities.

Thus, the concept of social capital and community is defined on the basis of a sense of shared identity rather than on the basis of coordinates and / or physical and spatial proximity.

# 4. Discussion

This review identified the major issues related to social capital in health issues among TGD people. From the analysis of the included papers, four macro areas of analysis emerged from the point of view of the data collected: the types of definitions used, the methods and techniques of measurement, the size of the social capital analyzed and the main findings achieved.

As for the definitions used, as already reported, it is interesting to note that only in 2022, with the study of Li, Fabbre and Gaveras (2022) a new way of defining and studying the social capital was developed, with their authenticated social capital. It is a reinterpretation and unification between the approach of Bourdieu and Putnam, by means of an author little known as Strauss (1978).

The results achieved, as reported in the previous paragraph, however, concern not only the TGD elderly, therefore it would be necessary to apply the concept of authenticated social capital also to other age groups, to effectively validate this new approach.

The relevant aspect, however, relates to the fact that, although there is a certain univocity and accessibility in the use of the notion of social capital in TGD people, there are very few studies in this regard.

This aspect requires significant reflection since social capital is being made up of social relations, these have a direct and indirect impact on health and inequality related to the status of SGM (LeBlanc and Perry, 2021).

This paper confirms the fact that research on SGM rarely focuses on the concept of social capital. Yet social exclusion and low social capital have a particularly negative impact on the health of individuals such as TGD people (Petruzzella, Feinstein and Lavner, 2019).

As far as the methodological approaches used are concerned, the fact that with the 8 original articles, there is a parity between qualitative (Pinto, Melendez and Anya, 2008; Perez-Brumer et al.,2017; Hwahnga et al., 2021; Li, Fabbre and Gaveras, 2022); and quantitative (Hamann and Sherblom, 2014; Erosheva, Emlet, and Fredriksen-Goldsen., 2015; Tan et al. 2021; Zarwell et al. 2021) approaches, is a significant step forward compared to what was the trend until some time ago(Schilt and Lagos, 2017) because before 2010, there were no quantitative data on TGD people, because all questionnaires used up to then only provided a binary view in the choice of gender (male and female) (Westbrook & Saperstein 2015).

Clearly, due to the scarcity of studies, it is certainly not possible to generalize. This aspect is important because, especially from a sociological point of view, the data on TGD people derives mainly from qualitative studies, due to the fact that qualitative approaches allow to grasp the phenomenological and lived aspects in a more in-depth way (Compton 2015, Rubin 2003).

It would be interesting to improve this aspect as well, to resort to approaches based on mixed methods(Pearce 2012) in such a way as to be able to integrate the qualitative aspects of the social capital of TGD people and the quantitative ones, which have not yet been applied in this field of research. Regarding the types of social capital studied, the studies in this article allow us to argue that bonding capital has been more studied.

Subsequently, bridging capital was analyzed in many of the articles (Hamann and Sherblom, 2014; Erosheva, Emlet, and Fredriksen-Goldsen, 2015; Perez-Brumer et al.,2017; Hwahnga et al., 2021; Tan et al. 2021; Zarwell et al., 2021; Li, Fabbre and Gaveras, 2022) included in this review.

In neither of the two types analyzed, a concept was taken into consideration that would be in an intermediate position between bonding and bridging capital, which for TGD people and in general for SGM, constitutes a fundamental component of their social capital (LeBlanc and Perry, 2021), i.e. the families of choice (Weston, 1997).

Families of choice for SGM, and in particular for TGD people, constitute a set of networks of broad and intimate relationships (Fish & Russell, 2018).

This type of "family", thus, is the result of the deconstruction of the concept of family from a biological and juridical point of view, which is the foundation of the social capital of the SGM, but which has not yet been analyzed with the profile of social capital (Thomeer, Donnelly, Reczek, and Umberson, 2017).

While the least studied type of social capital is linking; perhaps one or more in-depth studies would be necessary, also because one of the main problems inherent to the social capital of TGD people and their state of health, i.e. the barriers they encounter in the health context, and which involve a further reduction of capital social status and their state of health (Petruzzella, Feinstein, and Lavner, 2019).

Finally, this review demonstrates a striking absence of research about the social capital of TGD and in general for SGM.

# 5 Conclusions

This paper confirms that there is a marked absence of academic research about the social capital of TGD people. This paper confirms that there is a marked absence of academic research about the social capital of TGD people.

This paper presents strengths and limitations. Strengths of this review include using the concept of social capital in TGD people, and consideration for the multitude of factors that converge and interact with the influence this concept in their health status. In its design, this article followed the guidelines outlined by Tricco, Langlois, and Straus (2017).

As a result, this review is transparent, synthesises current knowledge about interest, and can be used to inform future sociological and not only interventions. The paucity of scholarly evidence available for review limits our ability to make conclusive statements about social capital of TGD people. Small sample sizes in the included studies warrant caution when deriving generalised conclusions about social capital. I invite readers to consider that the themes presented here have been drawn from much larger, more complex research on the subject.

Thus, the social capital of TGD people is little studied, despite being a protective factor for their health, which already appears particularly fragile and affected by many chronic diseases, drug addictions, psychiatric disorders, etc. It was possible to note how recently new scales and new ways of approaching social capital have been proposed, created specifically for this type of SGM, also in view of the increasing visibility of TGD people in the social, political and above all healthcare space for what concerns their right to health, acceptance and elimination of health barriers.

# References

Aldridge, S., D.Halpern, and S. Fitzpatrick. 2002. Social Capital: A Discussion Paper London: Performance and Innovation Unit. Available at www. social-capital. net/docs/PIU, 20.

Beach, L. B., T.A. Elasy and G. Gonzales. 2018. Prevalence of Self-Reported Diabetes by Sexual Orientation: Results from the 2014 Behavioral Risk Factor Surveillance System. LGBT health, 5(2), 121-130. https://doi.org/10.1089/lgbt.2017.0091.

Boehmer, U., D.J. Bowen and G.R. Bauer. 2007. Overweight and obesity in sexual-minority women: evidence from population-based data. Am J Public Health. 97(6):1134-40. doi: 10.2105/AJPH.2006.088419.

Bourdieu, P. 1977. The economics of linguistic exchanges. Social science information, 16(6),645-668.https://doi.org/10.1177/053901847701600601

Cochran, S.D., V.M. Mays and J.G. Sullivan. 2003. Prevalence of mental disorders, psychological distress, and mental health services use among lesbian, gay, and bisexual adults in the United States. J Consult Clin Psychol.71(1):53-61. doi: 10.1037//0022-006x.71.1.53.

Coleman, E., Radix, A. E., Bouman, W. P., Brown, G. R., De Vries, A. L. C., Deutsch, M. B., ... & Arcelus, J. (2022). Standards of care for the health of transgender and gender diverse people, version 8. International Journal of Transgender Health.

Coleman, J. S. (1988). Social capital in the creation of human capital. American journal of sociology, 94, S95-S120."https://doi.org/10.1086/228943.

Compton, D. L. R. 2015. LG (BT) families and counting. Sociology Compass, 9(7), 597-608.https://doi.org/10.1111/soc4.12273.

Durkheim, E., 1897. In: Simpson, G. (Ed.), Suicide. The Free Press, New York, p. 1951 (Spaulding, J.A. and Simpson, G. (transl)).

Erosheva, E. A., H. J, Kim, C. Emlet and K. I. Fredriksen-Goldsen. (2016). Social Networks of Lesbian, Gay, Bisexual, and Transgender Older Adults. Research on aging, 38(1), 98-123. https://doi.org/10.1177/0164027515581859.

Fish, J. N. and S.T. 2018. Queering methodologies to understand queer families. Family Relations,67(1),12-25. https://doi.org/10.1111/fare.12297

Fredriksen-Goldsen, K. I., and H. J. 2017. The Science of Conducting Research With LGBT Older Adults- An Introduction to Aging with Pride: National Health, Aging, and Sexuality/Gender Study (NHAS). The Gerontologist, 57(suppl 1), S1-S14. https://doi.org/10.1093/geront/gnw212

Fredriksen-Goldsen, K.I., H.J. Kim, S.E. Barkan, A.Muraco and C.P. Hoy-Ellis. 2013. Health disparities among lesbian, gay, and bisexual older adults: results from a population-based study. Am J Public Health.103(10):1802-9. doi: 10.2105/AJPH.2012.301110.

Gittell, R., & Vidal, A. 1998. Community organizing: Building social capital as a development strategy.Sage.https://doi.org/10.4135/9781452220567

Granovetter, M. S. 1973. The strength of weak ties. American journal of sociology, 78(6), 1360-1380.https://doi.org/10.1086/225469

Green-Hamann, S. and J.C. Sherblom. 2014. The influences of optimal matching and social capital on communicating support. Journal of health communication,19(10),1130-1144. https://doi.org/10.1080/10810730.2013.864734

Hardacker, C., Ducheny, K., & Houlberg, M. (Eds.). (2018). Transgender and gender nonconforming health and aging. Springer.https://doi.org/10.1007/978-3-319-95031-0

Henry, G. W. (1948). Sex variants: A study of homosexual patterns. New York, NY: PB Hoeber.https://doi.org/10.1097/00007611-194812000-00025

Hirschfeld,M. (1948). "Sexual anomalies: The origins, nature and treatment of sexual disorders: A summary of the works of Magnus Hirschfeld MD". New York, NY: Emerson Books.

Hunter, J. and R. Schaecher. 1987. "Stresses on lesbian and gay adolescents in schools. Children and Schools, 9(3), 180-190. doi: 10.1093/cs/9.3.180.

Hwahng, S. J., B. Allen, C. Zadoretzky, H. Barber Doucet, C. McKnight and D. Des Jarlais. 2021. Thick trust, thin trust, social capital, and health outcomes among trans women of color in New York City. International journal of transgender health, 23(1-2), 214-231. https://doi.org/10.1080/26895269.2021.1889427.

Kawachi, I., B.P. Kennedy. 1997. Health and social cohesion: why care about income inequality? British Medical Journal 314, 1037±1040." https://doi.org/10.1136/bmj.314.7086.1037

Kim, H. J., K.I. Fredriksen-Goldsen, A.E Bryan, and A. Muraco. 2017. Social Network Types and Mental Health Among LGBT Older Adults. The Gerontologist, 57(suppl 1), S84-S94. https://doi.org/10.1093/geront/gnw169

King, A. and A. Cronin. 2016. Bonds, bridges and ties: applying social capital theory to LGBT people's housing concerns later in life. Qual Ageing; 17 (1): 16-25. doi: 10.1108. QAOA-05-2015-0023.

King, M., J. Semlyen, S.S.Tai, H. Killaspy, D. Osborn, D. Popelyuk, and I. Nazareth . 2008. A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. BMC Psychiatry. 18;8:70. doi: 10.1186/1471-244X-8-70

Lamaison, P., & Bourdieu, P. 1986. From rules to strategies: An interview with Pierre Bourdieu. Cultural anthropology, 1(1), 110-120.https://doi.org/10.1525/can.1986.1.1.02a00060

LeBlanc A. J. B. L. Perry. 2021. "Sexual and Gender Minority Health Vol: 21", Emerald Publishing Limited, England. https://doi.org/10.1108/S1057-6290202121

Li, Y., V. D. Fabbre, and E. Gaveras. 2022. Authenticated social capital: conceptualising power, resistance and well-being in the lives of transgender older adults. Culture, health & sexuality, 1-16. Advance online publication. https://doi.org/10.1080/13691058.2022.2044519

Lochner, K., I. Kawachi AND B.P. Kennedy. 1999. Social capital: a guide to its measurement. Health & place, 5(4), 259-270. https://doi.org/10.1016/s1353-8292(99)00016-7

Martin, A. 1982. "Some issues in the treatment of gay and lesbian patients". Psychotherapy Theory Research and Practice, 19(3), 341.https://doi.org/10.1037/h0088445

Pearce, L. D. (2012). Mixed Methods Inquiry in Sociology. American Behavioral Scientist, 56(6), 829-848. https://doi.org/10.1177/0002764211433798

Perez-Brumer, A. G., S. L. Reisner, S. L., S. A. McLean, A. Silva-Santisteban, L. Huerta, K. H. Mayer, J. Sanchez, J. L. Clark, M. J. Mimiaga, and J.R. Lama. 2017.

Leveraging social capital: multilevel stigma, associated HIV vulnerabilities, and social resilience strategies among transgender women in Lima, Peru. Journal of the International AIDS Society, 20(1), 21462. https://doi.org/10.7448/IAS.20.1.21462

Petruzzella, A., B.A. Feinstein, and J.A. Lavner. 2019. Sexual orientation-related stigma and relationship functioning among female same-sex couples. Journal of lesbian studies, 23(4), 439-450.https://doi.org/10.1080/10894160.2019.1614861

Pinto, R. M., R.M. Melendez, and A. Y. Spector. 2008. Male-to-Female Transgender Individuals Building Social Support and Capital From Within a Gender-Focused Network. Journal of gay & lesbian social services, 20(3), 203-220. https://doi.org/10.1080/10538720802235179

Putnam, R. D. 2000. Bowling alone: America's declining social capital. In Culture and politics (pp. 223-234). Palgrave Macmillan, New York. https://doi.org/10.1007/978-1-349-62397-6_12

Putnam, R. D., R. Leonardi, R. Y. Nanetti. 1993. Making democracy work: Civic traditions in modern Italy.NJ, USA: Princeton University Press. https://doi.org/10.1515/9781400820740

Ross-Reynolds, G. and B. S. Hardy. 1985. ""Crisis counseling for disparate adolescent sexual dilemmas: https://doi.org/10.1080/02796015.1985.12085176 Pregnancy and homosexuality"". School Psychology Review, 14(3), 300-312

Rubin H. 2003. Self-Made Men: Identity and Embodiment Among Transsexual Men. Nashville, TN: Vanderbilt Univ. Press" https://doi.org/10.2307/j.ctv17vf6c0

Ryan, B. 1959. Banfield: The Moral Basis of a Backward Society. Edited by Walter C. McKain, Jr.(Book Review). Rural Sociology, 24(1), 60.

Schilt, K. And D. Lagos. 2017. The development of transgender studies in sociology. Annual Review of Sociology, 43, 425-443. https://doi.org/10.1146/annurev-soc-060116-053348

Stinchcombe, A., K. Wilson, K. Kortes-Miller, L. Chambers and B. Weaver 2018. Physical and mental health inequalities among aging lesbian, gay, and bisexual Canadians: cross-sectional results from the Canadian Longitudinal Study on Aging (CLSA). Canadian journal of public health = Revue canadienne de sante publique, 109(5-6), 833-844. Advance online publication. https://doi.org/10.17269/s41997-018-0100-3

Strauss, A. 1978. A Social World Perspective. Studies in Symbolic Interaction 1: 119-128.Tan, R., C. A. O'Hara, W. L. Koh, D. Le, A. Tan, A.Tyler, C. Tan, C. Kwok, S. Banerjee and M. L. Wong. 2021. Social capital and chemsex initiation in young gay, bisexual, and other men who have sex with men: the pink carpet Y cohort study. Substance abuse treatment, prevention, and policy, 16(1), 18. https://doi.org/10.1186/s13011-021-00353-2

Thomeer, M. B., R. Donnelly, C. Reczek, and D. Umberson. 2017. Planning for future care and the end of life: A qualitative analysis of gay, lesbian, and heterosexual couples. Journal of Health https://doi.org/10.31219/osf.io/2qpa7 and Social Behavior, 58(4), 473-487."

Tricco, A. C., E. V. Langlois, and S. E. Straus. 2017. Rapid Reviews to Strengthen Health Policy and Systems: A Practical Guide. Gen eve: World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/258698/9789241512763eng.pdf;jsessionid=CA8E7498F8842638636E550E74F1C22B? sequence=1

Westbrook, L. and A. Saperstein. 2015. New categories are not enough: Rethinking the measurement of sex and gender in social surveys. Gender & Society, 29(4), 534-560. https://doi.org/10.1177/0891243215584758

Weston, K. 1997. Families we choose: Lesbians, gays, kinship. New York, NY: Columbia University Press.

Woolcock, M. 2001. The place of social capital in understanding social and economic outcomes. Canadian journal of policy research, 2(1), 11-17.

Zarwell, M., J. L. Walsh, K. G.Quinn, A. Kaniuka, A. Patton, W.T.Robinson and R.J. Cramer 2021. A psychometric assessment of a network social capital scale among sexual minority men and gender minority individuals. BMC public health, 21(1), 1918. https://doi.org/10.1186/s12889-021-11970-8

# Order and disorder in the cities

Umberto Pagano*

**Abstract**

In recent years a paradigm has emerged for which urban liveability coincides with the existence of conditions of order, rationality, predictability and safety. If we combine this with the enormous technological progress applied to the management of urban ecosystems and the strongly transitional nature of our age (digital transition, climate change, ecological transition ...), we understand why in the last twenty years the concept of "Smart City" has been one of the most successful. But exactly what are we talking about when we talk about Smart Cities? Actually, the process of smartification does not only concern the urban dimension but, in some way, seems to apply to so many aspects of life. What kind of rationality is hidden in the dynamics of smartification? Are there dark sides of the Smart Cities? Are there alternatives to the Order based on standardization, digital surveillance, massive use of increasingly invasive technologies? These are categories whose application is generally argued with the need to generate "sustainable" ways of life but to what extent are these categories sustainable themselves? Martin Heidegger warned that the fact that "everything works" is exactly the problem and not the solution. Is humanity generating an increasingly *irrational rationality*?
The provocation launched by some Authors (above all Richard Sennett) is that there is the possibility of an antagonism to this process, designing cities as something open, never concluded, *dis-organized*. But what exactly does this *disorder* consist of? Is it a mere utopia or is it really possible to develop concrete categories and urban planning practices consistent with it?

**Keywords**: Sociology, Urban Planning, Architecture, Smart City, Transition.[1]

* University "*Magna Græcia*" of Catanzaro, Italy; umberto.pagano@unicz.it
[1] Received: 2022-10-29; Accepted: 2021-12-28; Published: 2022-12-31; doi: 10.23756/sp.v10i1.943.

# 1. Introduction

The "escape from the city" is a myth that, in macro terms, has a negligible impact on the process of progressive – and rapid – urbanization of the world. It is estimated that by 2050 70% of the world's population will live in a city or a megacity (UN, 2018), a threshold that Europe has already crossed for some time[2]. If we add to this the projections on population growth and the effects of climate change, with the associated migratory dynamics, it is well understood how a considerable portion of the near future of humanity depends, more or less directly, on the way in which the challenges posed by pressure on urban centres and the need for complex urban regeneration processes will be met.

For twenty years now, an enormous interest has developed around Smart Cities, as part of a wider process of *smartification* of the world, in which the term "smart", weighed down by abuse, now appears substantially transfigured and often reduced to a mere marketing etiquette. Basically, nowadays it refers to a product or process based on an "intelligent" optimization of resources and results, thanks to the advanced and integrated use of ICT and digitalization. But exactly what kind of intelligence are we talking about? Who is the subject that generates it? And more: what exactly is a Smart City? The optimization of resources takes place according to which perspective? For whose benefit? Of course, "smart" is everything but neutral term, with an aura of positivity that already in the beginning makes explicit the determination to deny the critical aspects, which instead are by no means marginal. On closer inspection, the rationality behind *smartification* is mostly the capitalistic rationality, the interest of capital in creating forms of functional "order", predictability and control that do not necessarily coincide with the interest of citizens and community as a whole. In the *smartification* take place predictability, "positivity" and pornographic over-exposure – as intended by Jean Baudrillard (1995)[3] – that lead to a sort of "closure", to a scientific (or *scientistic*) and *dataistic* organization of life, which also has precise implications on spaces.

The result of *urban smartification* – and other forms of *smartification* – generates a situation that is only apparently paradoxical: increasingly "intelligent" cities – and/or systems of objects –, increasingly "rational" spaces and temporalities and, at the same time, increasingly dumb citizens.

Perhaps today the role of Urban Planning and Sociology is also unfolding in all its implications and consequences – and more or less symbolic violence

[2] Already in 2018 the European urbanization rate was 74% (UN, 2018).

[3] For a discussion of the subject, see: Pagano, 2007. The theme of the *annihilation of negativity*, as a progressive elimination of the dialectic with the *Other* and of all that is indeterminacy and mystery, has also been developed, with noteworthy results, by Han (2013; 2015), who frames it in a Hegelian perspective.

(Cfr. Bourdieu, 1997) and violations – the dialectic "*Order*" vs. "*Disorder*", for the conceptualization of nowadays' and tomorrow's cities

# 2. Order

Although it has been a topic of discussion for twenty years, there is no uniformity on the meaning of the term "Smart City". This is partly inevitable, since it is an "object" that can be conceptualized and treated from multiple perspectives, in which the wide and intense use of information and communication technologies remains however essential. The schematization of Obringer and Nateghi (2021, 2) provides a useful framework to grasp heterogeneities, convergences and specificities of the different approaches proposed over the last 15 years, starting from the well-known conceptualization by Giffinger *et al.* (2007) (Table 1), based on 6 fundamental dimensions, of which Murgante and Borruso (2013) made a useful synthesis, framing the principal variables involved[4].

| *Authors* | *Definition/Conceptualization* |
|---|---|
| Giffinger *et al.* (2007) | A Smart City has several characteristics: *smart economy* (competitiveness), *smart people* (social and human capital), *smart governance* (participation), *smart mobility* (transportation and ICT), *smart environment* (natural resources) and *smart living* (quality of life). [See note n.4 for details]. |
| Caragliu *et al.* (2011) | A city which invests in "human and social capital and |

[4] *Smart Economy*: Employment rate; presence of innovative enterprises, presence and quality of universities and research institutes; infrastructures (roads, railways, airports, electronic infrastructures, etc.). *Smart Environment*: Air quality, percentage of separate collection of municipal waste (also electrical and electronic equipment waste), presence of green spaces in the city, efficiency and quality of water supply (water leakage and water treatment). *Smart Governance*: Not only related to e-government, percentage of ecological cars, use of recycled paper, energy saving, adoption of ecological policies for city planning and development, ability to network with other municipalities. *Smart Living*: Investments in culture and welfare providing several services, from childcare facilities to community libraries, from counselling structures for old people to cinemas, number of people below poverty level, hospital emigration rate, immigrants social integration, criminality rate. *Smart Mobility*: Extensive and efficient public transportation network, park and ride, great diffusion of ecological cars, limited traffic areas, cycle paths, bike and car sharing. *Smart People*: Education and early school leaving level, number of women working and holds positions within the administration, presence of foreign students, political participation, involvement in voluntary associations, newspapers diffusion and level of participation to cultural events (Murgante e Borruso, 2013, 635).

| | |
|---|---|
| | traditional (transport) and modern (ICT) communication infrastructure fuel sustainable economic growth and a high quality of life, with a wise management of natural resources, through participatory governance”. |
| Deakin and Al Waer (2011) | A city that works with the community to implement ICT that ultimately improves the quality of life for the community. |
| Nam and Pardo (2011) | A city that “infuses information into its physical infrastructure to improve conveniences, facilitate mobility, add efficiencies, conserve energy, improve the quality of air and water, identify problems and fix them quickly, recover rapidly from disasters, collect data to make better decisions, deploy resources effectively, and share data to enable collaboration across entities and domains”. |
| Batty *et al.* (2012) | A city in which “ICT is merged with traditional infrastructures, coordinated and integrated using new digital technologies”. |
| Kitchin (2014) | A city which has an extensive network of sensors and is capable of harnessing big data analytics to improve the function of the city. |
| Neirotti *et al.* (2014) | A Smart City should “optimise the use and exploitation of both tangible (e.g. transport infrastructures, energy distribution networks, natural resources) and intangible assets (e.g. human capital, intellectual capital of companies, and organisational capital in public administration bodies)”. |
| Angelidou (2015) | A city that takes a “humane” approach to integrate technology throughout the city, with a goal to advance human and social capital. |
| Marsal-Llacuna *et al.* (2015) | A city which improves “urban performance by using data, information and information technologies (IT) to provide more efficient services to citizens, to monitor and optimize existing infrastructure, to increase collaboration amongst different economic actors and to encourage innovative business models in both the private and public sectors”. |
| Ahvenniemi *et al.* (2017) | Smart Cities use technology to enable sustainable development. |

Table 1 – Some significant definitions of “Smart City” in literature (Obringer e Nateghi, 2021, 2).

However, despite the heterogeneity, a fairly precise line of development can be identified in the evolutionary process of the conceptualization of the theme. A line that goes from the mere automation/digitalization of processes and services (with a focus mainly on physical and infrastructural aspects) – at the beginning of the 2000s – to a more mature idea of Smart City as a socially inclusive context (in which generation of human capital and citizens' participation in the processes gain an increasing importance), and then moves – around 2010 – to a greater centrality of the “quality of life”, up to nowadays, to a major attention (undoubtedly favoured also by the pandemic scenario) for the aspects of social interaction in respect of health and environment (Borruso and Balletto, 2022, 94).

But, for the purposes of this essay, I do not consider particularly interesting to dwell on these aspects (that was anyway appropriate to recall), which have been widely analysed in the literature, because they are essentially reasonings on the ends, mostly agreed in their generic nature, where instead the crucial discourse, in my perspective, is the relationship between means and ends or, even more, the typical tendency of means, especially in hyper-technical societies, to become ends.

The current concern is that society is moving towards dystopian scenarios in which the application of algorithmic mechanisms and mass surveillance, through a very dense and very extensive network of sensors, is consolidating the power of a superior and homologating Intelligence[5], which proceeds by progressive weakening of any other intelligence and rationality. Paraphrasing Marx, we could say that the dominant smartness is the smartness of the ruling class. Although today the concept of class has a lower explanatory efficacy than in the past, what we mean here is that the rationality implicit in *the smartification process* is that of advanced Capitalism. And, in its essence, it is the kind of *irrational rationality* which – as Adorno and Horkheimer (1947) taught us – connotes the "short circuit" of the Enlightenment.

The *Order* of smartness is an order in which takes place the progressive separation between technical rationality and reasonableness, thus becoming less and less human. In the smartification scenario based on Big Data and algorithms, man becomes more and more a dataistic, more and more reduced to a string, set of measures, mere quantity (infinitesimal quantity, compared to the dataistic bigness that overwhelm us).

These might seem like abstract speculations but, on the contrary, they have extremely concrete implications. Think of digital bureaucracy: if in some fields digitization has triggered de-bureaucratization mechanisms, in other cases it is generating forms of *absolute bureaucracy*, in which the rationalization process has gone so far as citizens do not interface (albeit through standardized procedures) with a human bureaucrat but with automatic computer-driven systems (automatic vocal responders, automatic emails, etc.), with which they can interact only using choices and language settings provided, thus undergoing a total subordination to algorithmically determined times and methods, preventing any form of protest. It is about squaring the circle of the process that Max Weber described. It is a further and long step

[5] It is increasingly common to combine Artificial Intelligence and Machine Learning with other emerging technologies (such as IoT, autonomous vehicles, cloud computing, big data, cobots, cyber-physical systems…) to generate advanced urban solutions. Examples include: the use of deep learning and high-performance computing (HPC) for traffic predictions using sensor data, incident prediction, disaster management, logistics and urban planning, event detection for urban governance, disease detection. For an effective overview of these topics, see: Yigitcanlar *et al.*, 2020, which also provides a wide bibliography on the subject.

toward a more and more de-humanized form of “rational” bureaucracy that introduces us to a more and more inhuman and unreasonable order.

The fundamental trait of Smart Cities is an *Order* based on predictability. The *System* fears all that is not predictable with acceptable precision, and tends to establish, through advanced technology, a scopic regime aiming to the total control of minds and actions, to increasingly sophisticated measurement of every performance, in which man transubstantiates himself into data. It is a system where «Es funktioniert alles. Das ist gerade das Unheimliche, daß es funktioniert und daß das Funktionieren immer weiter treibt zu einem weiteren Funktionieren und daß die Technik den Menschen immer mehr von der Erde losreißt und entwurzelt»[6] (Heidegger, 1966/1976, 208).

As Rem Koolhaas (2014, 58) has well pointed out: «This regime has had a very big impact on cities and the way we understand cities. With safety and security as selling points, the city has become vastly less adventurous and more predictable». The ultimate goal of this process is the total predictability of man, his submission to the system in which he is completely deprived of autonomy. In their each and every activity, the citizens of the Smart City will be increasingly dependent, like children to be controlled, from an Order aiming at the progressive elimination of all that is ambiguous, opaque, unpredictable, mysterious, unclear, different, *other*.

> When we look at the visual language through which the smart city is represented, it is typically with simplistic, child-like rounded edges and bright colours. The citizens the smart city claims to serve are treated like infants. We are fed cute icons of urban life, integrated with harmless devices, cohering into pleasant diagrams in which citizens and business are surrounded by more and more circles of service that create bubbles of control (…). Where is the possibility of transgression? And rather than discarding urban intelligence accumulated over centuries, we must explore how what is today considered “smart” [compares] with previous eras of knowledge (Koolhaas, 2014, 58).

Virtually, the Order of the Smart City is an Order that is based on the High Definition of the data, but – as Baudrillard would likely say – to the highest definition of the data corresponds the lowest definition of meaning (*Cfr*. Baudrillard 1995; Pagano, 2007, 37-38). This Order works incessantly in the minds, incessantly smoothes the sphere of values.

> Maybe it is no coincidence that “liveable”, flat, cities like Vancouver, Melbourne and even Perth are replacing traditional metropolises in our imaginary (…). Because the smart city movement has been apolitical in its declarations, we also have to ask about the politics behind the improvements on offer. A new trinity is at

[6] «Everything is functioning. That is precisely what is awesome, that everything functions, that the functioning propels everything more and more toward further functioning, and that technicity increasingly dislodges man and uproots him from the earth». [Trans. Sheeha, ed., 1981].

> work: traditional European values of liberty, equality, and fraternity have been replaced in the 21st century by comfort, security and sustainability. They are now the dominant values of our culture, a revolution that has barely been registered (Koolhaas, 2014, 58-59).

An Order whose ultimate goal is that “everything works” orderly, predictably, smoothly, with no harshness, no contradiction, without dialectics, *hygienically*. The Smart City is *smooth*, in the sense suggested by Byung-Chul Han:

> From the perspective of hygienic reason, any ambivalence and any secret are also perceived to be dirty. Pure is transparency, and things become transparent when they fit into the smooth streams of information and data. Data have something pornographic and obscene about them. They have no inside, no flip sides; they are not ambiguous. In this they differ from language which does not permit things to come into perfectly clear focus. Data and information deliver themselves total visibility and they make everything visible. Dataism introduces the second Enlightenment. Acts, which presuppose a free will, belong to the dogmas of first Enlightenment. The second Enlightenment smoothens such acts into operations, into a data-driven process which takes place without any autonomy or dramatic orchestration of the subject. Acts become transparent when they are operationalized, when they submit themselves to computable and controllable process. Information is pornographic form of knowledge. Knowledge also contains negativity in the sense that is often gained against a resistance. Knowledge is altogether different temporal structure from that of information. In stretches between past and future. Information, by contrast, dwells in a smoothened-out time that is made up of indifferent point-like presences. This is a time without events [*Ereignis*] and destiny. The smooth is something one just likes. It lacks the negativity of opposition [*Gegen*]. (…). Smooth communication is free from any negativity of the other or alien. (…). The resistance coming from the other disturbs the smooth communication of the same. The positivity of smoothness accelerates the circulation of information, communication and capital. (Han, 2015/2018, 9-10).

# 3. Disorder

The basic idea of this essay came to me when a few months ago I read the essay by Richard Sennet and Pablo Sendra “*Designing Disorder. Experiment and Disruption in the City*” (2020), which was born with the interesting idea of a sort of connection with the famous text by Sennet himself “*The uses of disorder: personal identity and city life*” (1970), published exactly 50 years earlier, where he argued that the idea of an orderly, functional and perfectly functioning city, designed in a completely rational and efficient way, expresses a profoundly undesirable paradigm. Half a century later this idea not only maintains its own logic but, somehow, appears even more current and convincing.

At the basis of the reasoning there is an idea in some ways provocative but stimulating. After decades of hyper-rational urban planning, the elimination of amorphous, imprecise, incomplete spaces and areas, in favour of a perfect rationalization and separation of spaces, times, functions, the streets of many cities – also and above all those considered most liveable – are more and more "orderly" but increasingly lifeless, similar-morgues as sterile as oppressive. Liveability has become synonymous with "order", an order based on predictability, rationality (often increasingly irrational), certain boundaries, separation, in the illusion that controlling bodies and behaviours increases safety. For the sake of argument, let's assume for this to happen, what is the price?

Throughout the essay by Sennett and Sendra the term "Smart City" is never used, yet the Authors' discourse is profoundly close to the considerations developed in the first part of this contribution. Smartification is a manifestation of the Order as Sennet understands it.

> (…) something has gone wrong – radically wrong – in our conception of what a city itself should be. Imagining the good city became ever more difficult as planning become legalistic and bureaucratic after World War II. This presents a paradox (…) [which] can be traced to one big fault: the *overdetermination* of both the city's visual forms and its social functions. *The technologies which make possible experimentation have been subordinated to a regime of power which wants order and control* (Sennett in Sennett and Sendra, 2020, 27).

This paradox affects the very spirit of the city. Hyper-determination and predictability can only be the result of "closed", "finite", predetermined, standardized functions and processes, of containment of vitality and imagination, informality and improvisation. Ultimately, one of the final results of this process is people less and less autonomous in handling complex and unexpected situations.

Smooth and waterproof cities, without harshness, as beautiful as they are non-sensual, like Jeff Koons' sculptures, with their "sacralization of smoothness" (Han, 2015, 6); cities where borders, limits, edges are impenetrable and generate compartments functionally connoted and watertight. The opposite of openness so important to Sennett, continuously generated and re-elaborated inside the city, through negotiable limits and borders, non-absolute separations working like osmotic membranes, which absorb and expel, breathe, gasp, even cough. In short: *porous cities*. An open city is not smooth but it is rough, contradictory, unfinished. The open city, in this sense, is a profoundly human city, which presents fundamental flaws, fractures in the mechanisms of overdetermination, hyperdetermination and predictability; a city that "knows" how to surprise, destabilize, even disappoint. Disorder is, in short, the ability to break the order of technologically assisted dataistic predetermination. It is not a chaotic city, although this may be a weak point of the argument: the subtlety of the distinction between disorder and chaos. The difference is clear in logical terms, perhaps more nuanced in concrete reality. Naples, that Sennett proposes as an example of open city (p. 30), is undoubtedly a very porous city, partly dis-ordered (in the positive sense) but in the meantime chaotic and full of pockets of profound decay.

The way of designing cities with the logic of smartification produces a functional rationality that claims to provide spaces with pre-established endowments of meaning, thus failing in «providing communities the time and space to evolve, which is needed for growth» (p. 29). The outcome is a city, contrary to appearances, fragile. «The '*Brittle City*' is a symptom of society operating on a large scale as a closed system repressing anything that doesn't fit in, ensuring that nothing sticks out, offends or challenges» (*Ibidem*).

For Sennett and Sendra, then, the role of the urban planner – at least a urban planner not subservient to the system – should be close to Jane Jacobs' vision:

> In her view, big capitalism and powerful developers tend to favour homogeneity: determinate, predictable and balanced in form; the role of the radical planner is therefore to champion dissonance. In her famous declaration, «if density and diversity give life, the life they breed is disorderly» (Jacobs, 1961). The drivers of fast time – developers, investors, national actors – want their cities to be closed in form; that is, to be quantifiable, determinate, balanced and well integrated. The investor knows what he or she is getting» (Sennett and Sendra, 2020, 30-31).

It is necessary to unhinge the hyperdeterministic, "second-illuminist" smartification, through the logic of incompleteness: creating gaps, openings, meanings, even though contradictions, asperities, dissonances.

For Sennett and Sendra, in this process it is fundamental that the interventions on the places should not be top-down but the result of a participatory planning by the population, in which the urban planner from time to time, also according to the contexts and the mandate received, takes the commitment of facilitator, mediator, even activist, but without ever replacing the population involved.

I believe this is something that can work in some places, maybe in many places, but it's not a recipe for all places. In the population the different opinions are unlikely to have the same weight and social dignity, the same chance to determine, for the existence of minority but preponderant subjects, for the pressures of lobbies and capitals from outside the context, also for dynamics of overwhelming and criminal behaviour, and so on. Furthermore, there are contexts where there is a deep distrust of many people towards politicians and decision makers, so they are very sceptical about being involved in co-planning initiatives.

No doubt that an effective and far-sighted urban project can rarely be the result of simple standardized or slightly re-modelled replicas. At the very basis of a demanding intervention on the places, especially if it intends to be "vivifying", opening, un-closed, *non-smart*, there must be the ability "to read" the community: «(...) and you need to come up late with a proposal once you've listened» (p. 128).

This somehow reminds me of the first pioneering experiences of a new kind of urban design carried out in Italy after the World War II[7], based on collaboration of Italian scholars with foreign colleagues (mainly American): multi-disciplinary research groups in which architects and urban planners worked in the field together with sociologists, anthropologists, social psychologists, economists, historians, even philosophers, in an attempt to listen the populations and understand their *Weltanschauung* and their real needs.

The vivid stories of those experiences, told by protagonists as extremely stimulating and fruitful, convey a lesson that is perhaps worth recovering by contemporary urban planners and sociologists.

One of the most interesting projects was the one on Matera, a city that at the time represented a *unicum* of peasant culture, since in a very ancient settlement, dug into the tuff, in about 3,000 caves (called "*Sassi*"[8]), over 16,000 peasants lived in precarious and demeaning conditions, forced to walk for hours every day to reach the lands (that they did not own) where they worked. Adriano Olivetti, at the time president of the National Institute of Urban Planning, convinced the U.N.R.R.A.[9] to promote an intervention through the construction of a "new town" ("*La Martella*"), a "model" of village innovative for architectural and urban planning solutions. However, there was the need to preserve the balance of an ancient peasant culture, to understand the needs and desires of the people, often reluctant to leave the settlement, despite the awful living conditions, as well as to select which families should be relocated, since the new village could not accommodate everyone; in short, to avoid or at least "manage" a "*cultural apocalypse*" (Cfr. De Martino, 1977). It was then decided to create a "*Study Group on the Sassi of Matera*"[10], first experience in Italy of a deeply interdisciplinary approach to urban planning issues, which would then be the reference for several

---

[7] Among the most significant experiences were those animated by the "*Portici Group*", directed by Manlio Rossi-Doria, and the Adriano Olivetti's "*Community Movement*".

[8] "*Stones*" in Italian Language.

[9] "*United Nations Relief and Rehabilitation Administration*", an international humanitarian organization founded in 1943 to provide aid and assistance to the countries most affected by the war. It began operating in Europe as soon as the allied forces began the liberation of the Mediterranean and Balkan countries. Its action was mainly concentrated in some European countries, including Italy, through programs aimed at supporting the weakest sections of the population but also at resuming both agricultural and industrial production.

[10] For further information see: Musatti *et al.*, 1956; Bilò e Vadini, 2016.

subsequent projects. Also part of it was the American philosopher, of German origin, Friedrich G. Friedmann, who wrote this in a letter to Ludovico Quaroni, the urban planner in charge of the project:

> The community we are studying is a human community, it is more than an environment (physical and human), it is the activity of suffering and creating. It interprets what it undergoes and tries to transform by interpreting. It is culture (that is, the way of feeling and solving problems) in crisis. It is a society with a type of consciousness, which changes (which enters the historical consciousness, as they say) (…). In certain areas of life a void is created – old forms of life fall and the new ones are not yet ready – and therefore pseudo-solutions, abstractionisms come out.
> (...) It is evident: where a whole culture, a way of life, changes (or even collapses), technical aids are not enough; they are, at best, aids towards new forms of culture. What right do we have to intervene? What is or should be our philosophy of intervening? Of course, we must distinguish between problems and problems. There are the ultimate problems, those poetic-metaphysical tensions, a certain human sensitivity, which are expressed in various aspects of life. By studying them, we help ourselves, expanding our human experience: we cannot think to provide solutions for problems of this kind (it would be like wanting to abolish humanity itself). But then, there are other problems (it would be useful to trace the exact limits between the two groups of problems), I would say external tensions, which claim to be resolved in one way or another. What is our motivation for trying to help others to solve them? Why do we want to take the peasant and transfer him to *La Martella*? Why in the historical configuration does this represent a peaceful rational development (as opposed to irrational and violent developments)? It is important to clarify the philosophy of intervention, of our intervention, also for the efficiency of our attempts. It is important to make a list of problems (which means understanding the community in its true vitality) and of relationships between these problems (intangible and explicit ones); it will be necessary to distinguish those that must be solved by means of "external" intervention and those that, on the other hand, must find their natural course (...) I believed, and still believe today, that in order to understand human reality we do not need a detailed description (I would say: from outside), but we do need intimate penetration, dictated not by sentimentality, but by a deep sense of social responsibility. In other words, not a coldly positivist study: I am convinced that there is objectivity beyond the narrow field of today's science, moral objectivity, if you like… (Friedmann, 1951)[11].

I quoted a long passage, but I think it is worth reporting it because, despite it concerns an experience distant in time, it returns an ethical dimension of research and urban planning that I believe is still valid today, regardless of the specific context. Whether it is about a Cambodian countryside or a New York neighbourhood, I believe that this type of approach may represent one of the greatest safeguards against smartification of the world.

[11] This is an extract from the letter sent by F.G. Friedmann to Ludovico Quaroni on November 18th, 1951, reported in: Marselli, 1990, 222-223. [My translation from Italian].

# References

[1] Adorno T., Horkheimer M. (1947), *Dialektik der Aufklärung*, Amsterdam, Querido Verlag.

[2] Ahvenniemi H., Huovila A. (2020), “How do cities promote urban sustainability and smartness? An evaluation of the city strategies of six largest Finnish cities”, in Environment, Development and Sustainability, 23, pp. 4174-4200.

[3] Angelidou M. (2015), “Smart Cities: A conjuncture of four forces”, in *Cities*, 47, pp. 95-106.

[4] Batty M., Axhausen K.W., Giannotti F., Pozdnoukhov A., Bazzani A., Wachowicz M., Ouzounis G., Portugali Y. (2012), “Smart cities of the future”, in *The European Physical Journal Special Topics*, 214(1), pp. 481-518.

[5] Baudrillard J. (1995). *Le crime parfait*, Paris, Éditions Galilée.

[6] Bilò F., Vadini E. (2016), *Matera e Adriano Olivetti*, Roma/Ivrea, Comunità Editrice.

[7] Borruso G., Balletto G. (2022), “Smart Cities: nuove sfide”, in Morri R., Pasquinelli d’Allegra D., Pesaresi C. (a cura di), *Il cammino di un geografo, un geografo in cammino. Scritti in onore di Gino De Vecchis*. Milano, FrancoAngeli.

[8] Bourdieu P. (1997). *Méditations pascaliennes*, Paris, Le Seuil.

[9] Caragliu A., Bo C.D., Nijkamp P. (2011), “Smart cities in Europe”, in *Journal of Urban Technology*, 18(2), pp. 65-82.

[10] de Martino E. (1977), *La fine del mondo. Contributo all’analisi delle apocalissi culturali*, Torino, Einaudi.

[11] M., Al Waer H. (2011), “From intelligent to smart cities”, in *Intelligent Buildings International*, 3(3), pp. 140-152.

[12] Giffinger R., Fertner C., Kramar H., Kalasek R., Pichler-Milanovic N., Meijers E. (2007), *Smart cities – Ranking of European medium-sized cities. Technical Report*, Wien, Centre of Regional Science.

[13] Han B.-C. (2013). *Im Schwarm. Ansichten des Digitalen*, Berlin, MSB Matthes & Seitz.

[14] Han B.-C. (2015). *Die Errettung des Schönen*, Frankfurt am Main, S. Fisher Verlag. Eng. Trans. *Saving Beauty,* Cambridge, Polity Press, 2018.

[15] Heidegger M. (1966), *Nur noch ein Gott kann uns retten*, interview, September 23rd, 1966, published in *Der Spiegel* on May 30th 1976, pp.193-219. Eng. Trans. "Only a God Can Save Us", in *Heidegger: The Man and the Thinker,* ed. Sheehan T., pp. 45-67, Chicago, Precedent Publishing Inc., 1981.

[16] Jacobs J. (1961), *The death and life of great American cities*, New York, Random House.

[17] Kitchin R. (2014), "The real-time city? Big data and smart urbanism", in *GeoJournal*, 79(1), pp. 1-14.

[18] Koolhaas R. (2014), "Cities that are truly smart", in AA.VV. *Digital Minds for a New Europe*, Brussels, The Lisbon Council for Economic Competitiveness and Social Renewal, pp. 58-59 (the essay is an adaption from a talk given at the High-Level Group Meeting on Smart Cities in Brussels on September 24th 2014).

[19] Marsal-Llacuna M.-L., Colomer-Llinas J., Melendez-Frigola J. (2015). "Lessons in urban monitoring taken from sustainable and livable cities to better address the smart cities initiative", in *Technological Forecasting and Social Change*, 90, pp. 611-622.

[20] Marselli G.A. (1990), "Sociologia del vecchio e nuovo Mezzogiorno", in *Storia del Mezzogiorno* Vol. XIII, pp. 173-233, Napoli, Edizioni del Sole.

[21] Murgante B., Borruso G. (2013), "Cities and Smartness: A Critical Analysis of Opportunities and Risks", in *Lecture Notes in Computer Science*, June 2013, 7973, pp. 630-642.

[22] Musatti R., Friedmann F.G., Isnardi G. (1956), *Lo studio di Matera: saggi introdut*tivi, (ed.) Commissione per lo studio della città e dell'agro di Matera, UNRRA-CASAS I giunta, Roma.

[23] Nam T., Pardo T.A. (2011). "Conceptualizing smart city with dimensions of technology, people, and institutions, in *Dg.o '11 Proceedings of the 12th annual international digital government research conference: Digital government innovation in challenging times*, pp. 282-291, College Park, Maryland, Association for Computing Machinery.

[24] Neirotti P., De Marco A., Cagliano A.C., Mangano G., Scorrano F. (2014). “Current trends in smart city initiatives: Some stylised facts”, in *Cities*, 38, pp. 25-36.

[25] Obringer R., Nateghi R. (2021), “What makes a city ‘smart’ in the Anthropocene? A critical review of smart cities under climate change”, in *Sustainable Cities and Society*, 75 (2021) 103278

[26] Pagano U. (2007). *L’uomo senz’ombra. Elementi di sociologia dell’inautentico*, Milan, FrancoAngeli.

[27] Sennett R. (1970), *The uses of disorder: personal identity and city life*, New York, Knopf.

[28] Sennett R., Sendra P. (2020), *Designing disorder: experiments and disruptions in the city*, New York, Verso.

[29] United Nations, Department of Economic and Social Affairs, Population Division (2018). *The World’s Cities in 2018 - Data Booklet* (ST/ESA/ SER.A/417).

[30] Yigitcanlar T., Butler L., Windle E., Desouza K.C., Mehmood R., Corchado J.M. (2020), “Can Building ‘Artificially Intelligent Cities’ Safeguard Humanity from Natural Disasters, Pandemics, and Other Catastrophes? An Urban Scholar’s Perspective”, in *Sensors* 2020, 20(10), 2988.

# It is Impossible to Teach Special Relativity Without Deceiving the Student

Denis Thomas[*†]

**Abstract**

As the title asserts, it is impossible to teach the theory of special relativity without deceiving the student, which means that everyone who already accepts the theory as truth has been deceived. The resulting problem from this deception is, not only is science being held back as people not being told truth, these people are passing their deception onto others, even using time dilation as an answer to the distant starlight problem which many use to attack the account of Biblical creation instead of focusing on the error which yields such exaggerated stellar measurements. The focus of this paper is to expose many of the deceptions within physics texts used to deceive the student, along with several lies which have been told in support of the theory, such as GPS requiring relativity, the Hafele-Keating experiment, muons, etc., while also revealing Einstein's confusion concerning light.

**Keywords:** Hafele-Keating, GPS, light clock, length contraction, Lorentz transformation [‡§]

[†] 160 Indian Ridge Drive, North Augusta, South Carolina 29860, USA. Drillstop0@gmail.com; 803-845-0936

[§] Received on December 1, 2022. Accepted on December 29, 2022. Published on December 30, 2022. DOI: 10.23756/sp.v10i1.1036. ISSN 2282-7757; eISSN 2282-7765. 

# 1. Introduction

Special Relativity is a false theory which many people have accepted as truth. This theory has negatively impacted many of the areas of science, including astronomy, where some astronomers have tried to invoke time dilation as an answer to how we can observe distant starlight if the Earth was only created about 6,000 years ago. But if these men would have instead applied God's advice and examined the flaw of stellar magnitude, which is invalid because of wave interference, they would have reached the same conclusion which astronomer Alan Hirshfield wrote: "a star's brilliance reveals nothing about its remoteness,"[1] and thereby removed one of the obstacles which cause some to reject the truth of the Bible. And words such as spacetime would not be heard from the pulpit. Special Relativity rejects the fact that all truth is absolute as words get redefined, with claims such as simultaneous does not mean simultaneous for everyone.

Special Relativity is a comparative relationship between two or more vantage points, with the claim of a difference from classical relativity as the movement of one vantage point, or frame of reference, approaches the speed of light. The confusion and math of Hendrik Lorentz is what Einstein accepted and based this theory upon, along with his own confusion as he pondered different scenarios referred to as a gedankenexperiment, or a thought experiment. Because Einstein did this and many elevate him to god status, several texts equate these examples of mental reasoning to scientific experiments.

Several of these thought experiments from physics texts will be presented here. Most of the examples and illustrations requires at most some simple math to understand the error being presented to the student. Look for the error in the examples. The error will be explained following the example. Although the velocity of light is very fast, approximately 300,000,000 meters per second, in most cases the mathematical variable $c$ is used to denote it.

Many people have been confused into thinking Einstein's theory of mass equivalence, $E=MC^2$, is related to special relativity, but it is not. However, because of the deception of special relativity, the deception of relativity has expanded to included relativistic mass, and relativistic energy. Einstein's theory of mass equivalence was published on November 2, 1905, while the theory of special relativity was published on September 26, 1905. Einstein also had two other papers published earlier that same year.

Special Relativity is taught like a magician's trick, as the magician tries to confuse the audience and divert attention from the truth of what he is doing, which is done with a wide variety of examples. A good magician can easily fool you, requiring extra thought to understand the deception. If you can imagine the number of text pages a student must cover from a near 1,000-page text during each physics class, along with that of his/her other classes and reports to write, it is easy to understand how the unsuspecting student does not take the time to verify what is being taught is true, especially after having for years been indoctrinated to believe that science is an unbiased pursuit of truth.

Within most physics texts teaching special relativity will be a statement such as "contrary to common sense", which should be a red flag to the reader, but the authors will try to persuade the student to ignore common sense and accept "the reality of special relativity". Common sense is something God has given man, and he should use it. "According to the principle of relativity, no inertial frame of reference is more correct than any other in the formation of physical laws. Each observer is correct in his or her own frame of reference."[2]

## 2.1 Example #1: time dilation

The following example[3] occupies six pages of the students' text, with enough quoted and summarized here that the deception should be apparent: "A dramatic illustration of time dilation is provided by identical twins, one an astronaut who takes a high-speed round-trip journey in the galaxy while the other stays home on Earth. When the traveling twin returns, he is younger than the stay-at-home twin. How much younger depends on the relative speeds involved. If the traveling twin maintains a speed of 50% the speed of light for one year (according to clocks aboard the spaceship), 1.15 years will have elapsed on Earth…. One question often arises: Since motion is relative, why doesn't the effect work equally well the other way around? Why wouldn't the traveling twin return to find his stay-at-home twin younger than himself? We will show that, from the frames of reference of both the earthbound twin and the traveling twin, it is the earthbound twin who ages more."

A spaceship has a flashing light on it which blinks once every six minutes. If the spaceship is at rest relative to Earth, once the initial flash is received on Earth, another flash will be observed every six minutes. Nothing special about that. "When motion is involved, the situation is quite different. It is important to note that the speed of the flashes will still be $c$, no matter how the ship or receiver may move. How frequently the flashes are seen, however, very much depends on the relative motion involved. When the ship travels toward the receiver, the receiver sees the flashes more frequently. This happens not only because time is altered due to motion, but mainly because each succeeding flash has less distance to travel as the ship gets closer to the receiver. If the spaceship emits a flash every six minutes, the flashes will be seen at intervals of less than six minutes. Suppose that the ship is traveling fast enough for the flashes to be seen twice as frequently. Then they are seen at intervals of 3 minutes.

If the ship recedes from the receiver at the same speed and still emits flashes at 6-minute intervals, these flashes will be seen half as frequently by the receiver – that is, at 12-minute intervals. This is mainly because each succeeding flash has a longer distance to travel as the ship gets farther away from the receiver. The effect of moving away is just the opposite of moving closer to the receiver. So, if the flashes are received twice as frequently when the spaceship is approaching (6-minute intervals are seen every 3 minutes), they are received half as frequently when it is receding (6 minute intervals are seen every 12 minutes)." Examples are given with the conclusion the astronaut twin aged less.

## 2.2 Comments on Example #1

Contrary to what the student was told in the text, time is not altered due to motion. Such a claim reveals a lack of understanding of what time is. Simply defined: "time is a system of information exchange, how that God and men communicate events with respect to the rotation of earth. What you plan to do tomorrow, what you did last year, how long Jesus was in the tomb, how long it took God to create the heaven and the Earth, how fast something travels,"[4] including the speed of light, cars, etc. and time itself all relate to the rotation of Earth. A clock does not determine time any more than a barometer determines atmospheric pressure. Both are merely instruments attempting to measure quantities. Yet it is a fundamental flaw with special and general relativity, the concept that "time is different for different observers."[5]

The formula to solve the problem correctly is $t = f \pm (f \times v_c)$, where t equals the time between the flashes observed on Earth; f equals the actual time between flashes on the spaceship; ± equals the travel direction of the spaceship with + distance getting farther from Earth, and – as distance gets less to Earth; $v_c$ equals the velocity of the spaceship expressed as a fraction of the speed of light. The author had stated the actual time between flashes of light on the spaceship was 6 minutes and the velocity of the spaceship was such that the flashes are received every 3 minutes while traveling toward Earth. Inserting this information into the formula we have $3 = 6 - (6 \times v_c)$; $0 = 3 – 6v_c$; $v_c = 0.5$. He claimed that if going away from Earth at the same velocity the flashes would be observed on Earth every 12 minutes. We get a different answer when we input the information into the formula: $t = 6 + (6 \times 0.5)$; $t = 9$ minutes. Special relativity does not factor in direction of travel. The example was that of the magician trying to confuse the student. Why did the author lie?

In another physics text[6], the authors provide an example of the twin paradox with Helen taking a starship flight to a star 9.5 ly from Earth and instantly returning at the same velocity, 0.95c, where she left her brother George. The calculations are given, showing George has aged 20 years and Helen has aged only 6.25 years. The authors had stated: "only one inertial reference frame measures proper time", and Helen's clock is "the clock that measures proper time. "George is expecting Helen to be younger than he is. Helen is expecting George to be younger than she is. Here's the paradox! It's logically impossible for each to be younger that the other at the time when they are reunited." The authors try to explain the paradox by stating: "George spends the entire time in an inertial reference frame, but Helen does not. The situation is not symmetrical. The principle of relativity applies only to inertial reference frames… Helen's analysis and calculations are not correct because she was trying to apply an inertial reference frame result to a non-inertial reference frame." This contradicts what the authors stated about Helen's clock. While the authors refer to the acceleration Helen felt, (the Lorentz calculations which George used are based upon Helen moving at constant velocity, which also contradicts what the author claims about Helen's motion being non-inertial) relative motion teaches us that, if physical sensations are ignored, she would be viewing George as the one experiencing the acceleration, thus they should reach the same conclusion about each other. So why did these authors lie?

## 3.1 Example #2: simultaneity

"Two events are simultaneous if they occur at the same time. Our everyday experiences and intuition suggest that the notion of simultaneity is "absolute"; that is, two events are either simultaneous or not for all observers. However, to determine if two events are simultaneous (or not), involves the measurement of time, and our studies of time dilation show that different observers do not always agree on measurements involving clocks and time intervals." An illustration is provided showing two lightning bolts striking both ends of a moving boxcar at the same instant an observer on the ground is located equal distance from the front and rear of the boxcar. Ted is the observer centered on the boxcar while Alice, the ground observer is equal distance from the front and rear of the railcar as the lightning bolts strike while the railcar passes by. "We now ask, Did the two lightning bolts strike simultaneously?" The author then tells that Alice viewed the lightning bolts as simultaneous, while explaining that because of Ted's motion he viewed the forward lightning strike first, and then states: "two observers must always agree on the order of two events that occur at the same point in space….In Ted's reference frame, the two lightning bolts are not simultaneous….Time dilation and the relative nature of simultaneity mean that special relativity conflicts with many of our intuitive notions about time…That is very different from Newton's picture, in which time is an absolute, objective quantity, the same for all observers."[7]

## 3.2 Comments on Example #2

With both the stationary and the moving observer, the event of the lightning strikes is simultaneous. But, because of motion the moving observer's perception of the event differs from that of the stationary observer. Had the moving observer realized that he was moving, he could have calculated which flash occurred first, or if they were simultaneous, if he also knew his velocity. Indeed, the moving observer could claim that the forward pulse of light arrived at him first. But the stationary observer would also make the same claim, that the forward pulse arrived at the moving observer prior to the rear light pulse. One of the tricks of the magician is to confuse you, which is what these authors try to do when claiming simultaneous does not mean simultaneous for everyone. But truth is true for everyone. There is no such thing as a relative truth. So why did the author lie?

Imagine you are in an electric golf cart and your friend is on the 50-yard line, which is located equal distance from speakers at each end of the field. At the instant you pass your friend, an announcement comes from the speakers. Your friend hears the sound from both speakers simultaneously, but because of your motion you hear the announcement from the closest speaker, prior to the echo sound from the other speaker. You are intelligent enough to know that the sound from both speakers was simultaneous, and it was because of your motion the sound had farther to travel from the distant speaker, you heard it last. Apply your intelligence to the lightning bolt thought experiment.

With special relativity, it is essential that you believe that simultaneous does not mean simultaneous for each observer. As one author, who used two firecrackers exploding instead of the two lightning bolts, wrote: "The paradox of Peggy and Ryan contains the essence of relativity, and it's worth careful thought. First, review the logic until you're certain there is a paradox, a logical impossibility. Then convince yourself that the only way to resolve the paradox is to abandon the assumption that the explosions are simultaneous in Peggy's reference frame. If you understand the paradox and its resolution, you've made a big step toward understanding what relativity is all about."[8] (Peggy was the moving observer.) Just say it to yourself: special relativity is true until you convince yourself.

"To talk about time, about simultaneity at a distance, you have to synchronize your clocks."[9] The issue of simultaneity is one of the errors Einstein made within the theory of special relativity. He describes two clocks, A & B and expands it to include clock C, in space and proposes an imaginary but incorrect way to synchronize the clocks:

"If at the point A of space there is a clock, an observer at A can determine the time values of events in the immediate proximity of A by finding the positions of the hands which are simultaneous with these events. If there is at the point B of space another clock in all respects resembling the one at A, it is possible for an observer at B to determine the time values of events in the immediate neighbourhood of B. But it is not possible without further assumption to compare, in respect of time, an event at A with an event at B. We have so far defined only an "A time" and a "B time." We have not defined a common "time" for A and B, for the latter cannot be defined at all unless we establish by definition that the "time' required by light to travel from A to B equals the "time" it requires to travel from B to A. Let a ray of light start at the "A time" $t_A$ from A towards B, let it at the "B time" to be reflected at B in the direction of A, and arrive again at A at the "A time" t0B. In accordance with definition the two clocks synchronize if tB − tA = $t0_A$ − tB. We assume that this definition of synchronism is free from contradictions, and possible for any number of points; and that the following relations are universally valid: -

1. If the clock at B synchronizes with the clock at A, the clock at A synchronizes with the clock at B.
2. If the clock at A synchronizes with the clock at B and also with the clock at C, the clocks at B and C also synchronize with each other.

Thus with the help of certain imaginary physical experiments we have settled what is to be understood by synchronous stationary clocks located at different places, and have evidently obtained a definition of "simultaneous," or "synchronous," and of "time." The "time" of an event is that which is given simultaneously with the event by a stationary clock located at the place of the event, this clock being synchronous, and indeed synchronous for all time determinations, with a specified stationary clock. In agreement with experience we further assume the quantity $\frac{2\mathrm{AB}}{t'_A - t_A} = c$ to be a universal constant the velocity of light in empty space."[10]

It is impossible to synchronize three remotely isolated stationary clocks, A, B, & C using only a ray of light. While Einstein's example required the ray of light to be emitted, reflected, received back with the first clock calculating the flight time of the light ray and thus adjust the clock (although he does not mention adjusting the clock), that at best could only work for two stationary clocks. Imagine clocks A & B are remotely separated by exactly 5 light-seconds. Clock A starts at time 0 and sends a light pulse to B. As B receives the pulse and reflects it back, it now starts and is lagging clock A by 5 seconds. Clock A receives the reflected pulse at A = 10, just as B = 5. Clock A must now calculate the round-trip time of the light pulse, divide that by 2, and subtract that from its own time in order to synchronize with clock B. That was with both clocks stationary. To also synchronize clock C as Einstein proposed would be impossible. While Einstein claimed that his imaginary physical experiments "settled what is to be understood by synchronous stationary clocks located at different places", he then applied his definition of clock synchronization in the next section of his theory to moving clocks. He stated: "*let the time* $\tau$ *of the* *moving system be determined for all points of the moving system at which there are clocks at rest relatively to that system by applying the method, given in §*1”. The two lightning bolt thought experiment shows the impossibility of the clocks of the moving system synchronizing, as the motion would cause the time for the light to go from A to B to not equal the time from B to A, just as the motion of the boxcar caused the flash from the forward lightning bolt to strike the rear of the boxcar prior to the rear flash arriving at the front. It also provides support for the claim that the one-way velocity of light has never been measured, only the two-way velocity such as with the Michelson-Morley experiment.

Imagine today’s atomic watches which are adjusted via a radio signal broadcasted by the Naval Observatory. On Earth, they would each be accurate to within microseconds. It would be impossible for astronauts on the moon and on Mars to have their watches synchronized using this radio signal, which travels at the speed of light. Einstein referred to his imaginary method of synchronization numerous times within his theory.

## 4.1 Example #3: Newtonian mechanics

“A spaceship (S’) moves with speed $v_{S'/E}$ = 1000 m/s relative to the earth (E). It fires a missile (M) with speed $v_{M/S}$ = 2000 m/s relative to the earth. (a) Newtonian mechanics tells us that the missile moves at a speed of 3000 m/s relative to the earth. (b) Newtonian mechanics tells us that the light beam emitted by the spaceship moves at a speed greater than c relative to earth; this contradicts Einstein’s second postulate.”[11]

## 4.2 Comments on Example #3

All types of waves travel at a constant velocity until conditions change. Contrary to the claim of the textbook author, Newtonian mechanics does not teach that the velocity of waves is added to the velocity of the craft. “Water waves produced by a slow-moving tugboat have the same speed as those produced by a high-powered speedboat.”[12] The waves of a train horn do not travel faster when the train is moving. However, when the medium upon which the wave is traveling is also moving, only then is the wave traveling faster with respect to something which is stationary. Boat waves in a river move faster than on a lake with respect to shore; audible voices inside a moving vehicle travel faster with respect to the ground; light waves in an expanding section of space would move faster if space actually expands. While galaxies expand, there is zero real evidence that space expands. (The example of inflating a balloon with dots on it provides zero evidence of space expanding, but is a visual aid for those teaching the Big Bang.) Why did the author lie?

# 5.1 The light clock

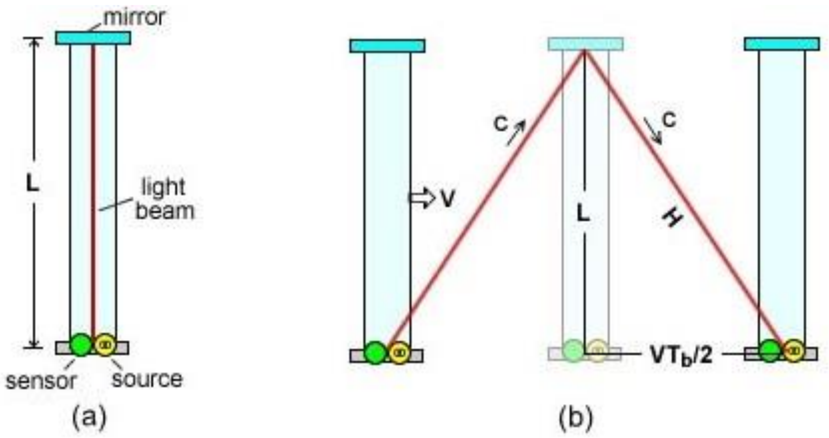



Fig. 1 the imaginary light clock

Figure 1, illustrated above[13], is an imaginary clock, which is the most frequently used example in the teaching of special relativity. In most physics texts, it is comprised of two horizontal mirrors with a pulse of light bouncing between the mirrors, while in the illustration above, and in some other physics texts, there is only one mirror, with the light source and sensor at the bottom. The diagonal path of figure (1b) is what is taught as the "path of light as seen from a position of rest." [14]

# 5.2 Comments on the light clock

There are numerous errors with the light-clock illustration. Can you spot them? The light clock errs in that light is portrayed as behaving as would an object with mass, acceleration is ignored, the stationary observer's observation is never shown from a point of inertial motion, the observational delay is ignored, and the thought that it is the clock which determines the time is completely wrong. A clock is merely an instrument attempting to measure a quantity. It can no more determine time than a barometer determines atmospheric pressure. Every instrument is subject to instrument error.

For the motion of one object to be considered inertial, it must be compared to another object and its movement must be perpendicular along a straight line to the other object. For simplicity, it is easiest to reference this line as the *x* axis, along with placing the "stationary" reference at *x*0.

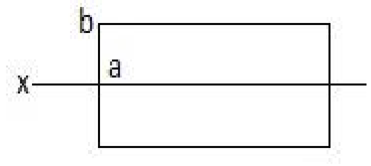



Fig. 2

Consider the box above, Fig. 2, moving along the x-axis. Only those points touching the x-axis can be considered inertial with respect to x, while the distance from x to b would be constantly changing at a nonlinear rate. Suppose the distance from a to b is 5’, which would remain unchanged as the box moves. If the distance from x to “a” is 10’, then by Pythagorean’s theorem the distance from “x” to “b” would be determined to be 11.18’. As the box moves to the right to xa = 15’, the distance from x to b becomes 15.81’. The stationary observer watching a moving lightclock would only view the vertical motion of the light along the y-axis, and not from the off-to-the-side vantage point illustrated in physics texts.

Space has only three dimensions. These can be plotted on an *x*, *y*, *z* graph. Adding t for time to a graph shows the position of these co-ordinates at a particular instant. But time is not a fourth dimension. Time is not a dimension, but a measurement of quantity which can be plotted on a straight line, with “The Beginning” at t = 0, the past to the right of that, counting up to the present, with the future to the right of that.

Often within physics texts, the time for the light to go from the bottom mirror to the top mirror and back to the bottom mirror is defined as one tick of the clock, while some define it as two ticks. Consider the following example: Bob is our moving observer, while Alice is stationary. Each time the pulse of light of the light clock hits the bottom mirror it causes a flash of light (one tick of the clock) to be emitted in the direction of Alice. Our clock could then be compared with the spaceship from example #1 which now flashes a light every second. We can then use our formula from “comments on example #1” to determine the time Alice would observe between flashes: $t = f \pm (f \times v_c)$. Bob is moving away at 0.50c. The flashes from his light clock, which are at one second intervals, are then observed every 1.5 seconds by Alice, and if Bob instantly reverses directions, returning at the same velocity, Alice will observe the flashes every 0.5 second. However, if we apply the Lorentz equation to the problem the answer is quite different, so let’s solve the problem as the students of physics are taught to do and look for the error. (Fig. 3 is another illustration of the light clock.)

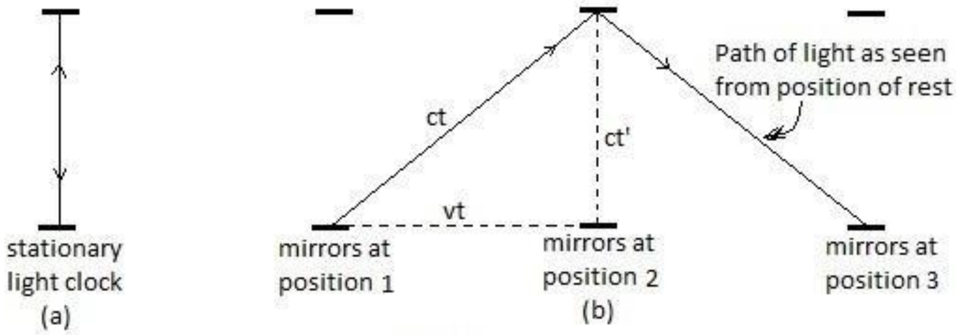



Fig. 3

We know that we can find the length of the hypotenuse of a right triangle using Pythagorean's theorem, where the square of the hypotenuse (c) is equal to the sum of the square of the adjacent side (a) of the triangle and the square of the opposite side (b) of the triangle. $a^2 + b^2 = c^2$. This equation works, whether the dimensions are expressed as length, velocity, or time. Since distance equals velocity multiplied by time, we can substitute that into our Pythagorean equation. Thus, we have $(ct)^2 = (vt)^2 + (ct')^2$, (with *c* being the velocity of light, and *v* being the velocity of the moving observer), which can be reduced to that of Fig.4:

$$t = \frac{t'}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Fig. 4

Inputting Bob's velocity, 0.5*c*, and the time of *t'* (1 second), we come up with t = 1.15 seconds in both directions, considerably different than the 1.5 seconds while receding and 0.5 seconds while returning we concluded above. So, which one is wrong? Actually, they both are. If Bob's clock merely consisted of a timer which caused a light to flash, then the 1.5 seconds between observed flashes while moving away from Alice, and the 0.5 seconds while approaching would be correct. But, since we used the imaginary light clock, we need to understand Bob's clock stopped functioning once he went into motion, a result of one of the characteristics of light. But first, let's expose another lie told to the students.

Relative motion means that a person's motion is relative to another, and that each can consider themselves as the one who is stationary. This is true whether or not the motion considered is inertial, accelerating, or decelerating, as each can, if the feelings of acceleration and deceleration are ignored, view the other as the one experiencing these types of motion, which is probably why early astronomers viewed the celestial objects as rotating around Earth. They did not feel their motion. In most physics texts which I have reviewed, the authors claim the reason why the moving observer did not reach the same conclusion as the Earth-bound stationary observer is because the moving observer was not in an inertial frame of reference. This would mean that only during the acceleration and deceleration would Bob's motion be non-inertial, and once at inertial velocity he would be considering Alice's clock as moving slower at the same rate she views his clock. But no physics text will factor the non-inertial motion into their math, but only make the statement this is why there is a difference. So why did the authors lie? The Lorentz transformation formula ONLY applies when both frames of reference are inertial with respect to each other.

Imagine the following thought experiment: an apparatus is located on the floor of an airplane and shoots paintballs vertically at the ceiling at regular intervals. We know that once the plane is traveling at a constant velocity, the paintballs will follow a vertical path to the ceiling. This is because the paintballs located inside the apparatus gained potential energy upon acceleration, which was changed into kinetic energy upon launch, giving the paintball the momentum of the forward velocity of the plane. For something to be moving, even at an inertial rate, it had to at some time in the past experience acceleration. The paintballs launched during acceleration will follow a diagonal path toward the rear of the plane. Now consider the light clock with its ball of light bouncing between two mirrors. Upon acceleration, the mirrors would move out of the path of the light pulse and the clock would stop working. Consider the light clock illustrated in Fig. 1, with a light flashing at regular intervals, bouncing off a mirror and returning to a sensor, the clock will still not function once the plane is in motion. This is because the photon of light did not exist prior to being emitted, and does not contain physical mass, wherewith to gain the potential energy to give it the forward momentum once emitted. As it travels up toward the upward mirror, the movement of the plane will move the mirror out of the light's path. One author used a laser pulse in the light clock (Fig. 1) and wrote: In other words, the stationary observer "concludes that because of the motion of the vehicle, if the light is to hit the mirror, it must leave the laser at an angle with respect to the vertical direction."[15] This meant the comparison of the light shot vertical in the case of the stationary light clock to when the laser was fired at an angle to hit the moving mirror. It is not comparing apples to apples, yet the author still guides the student into comparing the angled to the vertical light path to conclude time dilation. So why do these authors lie?

From Alice's perspective, our stationary observer, she could only view the vertical up and down motion of Bob's imaginary light clock, and not the diagonal path presented in each physics text. If, for example, the upper limit of the flight path of light pulse of the light clock is coordinate y10 on a graph, by drawing in the vanishing point of a perspective drawing, and considering the flight time of the light back to Alice, you should recognize the pulse of light is not rising and falling at a linear rate. However, if Bob's clock just emitted a flash of light every second (or at the same intervals as Alice's flashing clock), Alice could determine both the direction and velocity of Bob using our formula from comments on example #1: $t = f \pm (f \times v_c)$.

Because the illogical light clock is used in the majority of texts, physics and astronomy, which cover special relativity, another thought experiment will be provided here. Alice and Bob each have two radio receivers and one transmitter beside them. One receiver is tuned to receive the transmitter beside it, while the other is tuned to receive the other observer's transmitted signal. Both transmitters broadcast a pulse every second, similar to that of the U.S. Naval Observatory's WWV radio. Alice's transmitter broadcasts on 10 MHz, while Bob's on 20 MHz. Each receiver is equipped with a counter, counting each pulse received. Each counter reads zero just as Bob travels away from Alice. The Doppler effect causes a shift in the frequency (which can be calculated with the wavelength formula) each of them are receiving from the other's transmitter, but our imaginary receivers have auto tune ability and continues to receive the signal which travels at the same speed as light. In this imaginary scenario, regardless of acceleration, velocity, deceleration, or direction of movement, both observers conclude the exact same about the other and the four counters will not again synchronize until Bob returns back to his original position.

The following example, in addition to the comments on example #5, will further illustrate this flaw in Einstein's thinking. We live on a rotating Earth. "The moon's average distance from Earth is 238,855 miles. Since light travels at 186,000 miles per second, it would take 1.28 seconds for a pulse of light to go to the moon. If we make an imaginary circle of average lunar orbit based upon $2\pi r$, we have a circle of 1,500,770 miles. Divide that by 86,400, the number of seconds in a day. Our zenith moves along this circle at 17.37 miles per second. Multiply that by the light travel time of 1.28 seconds for 22.2 miles. If, as claimed by Lorentz, Einstein, and those claiming special relativity is true, light carries the forward momentum of motion, it would be necessary to aim the laser for lunar laser ranging at a spot 22.2 miles away from the retroflectors left by the Apollo astronauts. But they don't do that!"(Thomas, 2022)[16] (Lunar laser-ranging, contrary to the true science of laser ranging, is another of the deceptions taught to students, but is not the focus of this paper.)

Hendrik Lorentz recognized the speed of light is constant, and wrote the somewhat simple equation which is now referred to as the Lorentz factor, which seems to be based upon the math of Woldemar Voigt.[17] Since distance = velocity multiplied by time, the formula became $c^2t^2 = v^2t^2 + c^2t'^2$ with the conclusion that t was a longer duration of time that t' (t prime). The formula was further reduced to $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ , which is the relationship of the velocity of the moving object to the speed of light. (The Greek letter gamma, γ, is often used to designate the Lorentz factor.) But Lorentz erred in thinking light would behave similar to objects with physical mass. Einstein accepted Lorentz's confusion and wrote his theory based upon it. The Lorentz transformation formula is used not only in time dilation, but also in length contraction, relativistic momentum, and several others, such as relativistic mass. Since the formula is based on the thought that light would have the forward momentum as objects with physical mass, which are illustrated with the light clock and disproven with our examples, it should be recognized that in whatever application the Lorentz transformation formula is used, that it is a part of false science.

# 6.1 The moving elevator

Although Einstein's elevator thought experiment is considered foundational for general relativity, and not thought of until 1907, it offers further proof of Einstein's confusion along with revealing one of the flaws with the light clock thought experiment, again disproving special relativity. This "experiment" is also referred to as the moving lab, where a moving lab is referenced instead of an elevator.

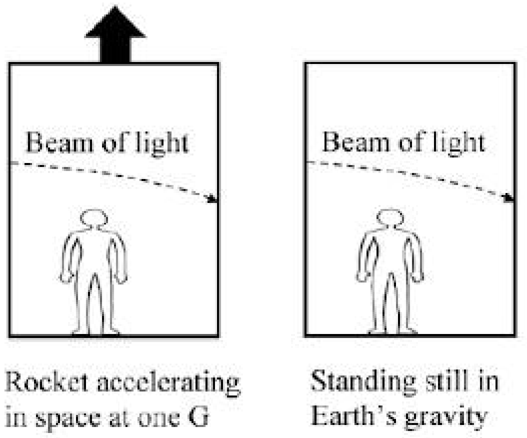



Fig. 5

The clipart image, Fig. 5 above[18] is similar to what is in numerous physics texts.

## 6.2 Comments on the moving elevator

The illustration on the left of Fig. 5 is correct because as the light waves leave the source on the left the motion of elevator, or rocket, is causing the walls to rise, giving the illusion that the light beam is bending with respect to the walls. That illusion would exist whether the elevator was accelerating, decelerating, or in inertial motion. Simply by rotating that illustration 90°, such that the elevator is now going horizontal and the light vertical, you invalidate the claims of the light clock, as it would be obvious that the light would not bounce between the mirrors as taught in physics texts. The illustration on the right shows Einstein's illogical thinking upon which he theorized that gravity would bend light in a similar manner, which has led to the misconception of curved space. He had thought that since acceleration of the elevator or rocket would "bend" the light that the acceleration of gravity would also bend light. However, the accelerating elevator only gave the illusion of light bending. Light is not deflected by gravity. While Newton may have theorized gravity bending light, gravity is similar to a permanent magnet in that its attraction is steady and on objects with mass, while the electromagnetic waves of radio and light alternate and are not affected by stationary magnets. Because you are possibly thinking of the 1919 eclipse, that will be covered next.

## 7.1 The eclipse of 1919

"Einstein predicted that starlight passing close to the Sun would be deflected by an angle of 1.75 seconds of arc – large enough to be measured….(Measuring this deflection has become a standard practice at every total eclipse since the first measurements were made during the total eclipse of 1919.)….In every instance, the deflection of starlight has supported Einstein's prediction."[19]

## 7.2 Comments on the 1919 eclipse

Frequently, lies are supported by additional lies. That is the case with special and general relativity, as many scientists have claimed their experiment or test proves the theory. Arthur Eddington's claim of the 1919 solar eclipse is such an example, which many cite as proof of relativity. His bias certainly added to his lack of objectivity, having the year prior written a book on general relativity, and counted himself as one of three men who understood the theory. "It is worthwhile mentioning at this point that none of later solar eclipse missions in 1922, 1929, 1936, 1947 and 1952 yielded conclusive results about the amount of light deflection (Newtonian or Einsteinian, cf. [10, p. 68]).”[20] So why did the author lie?

Amateur astronomer Donald Bruns attempted to repeat Eddington’s exercise with the eclipse observation of August 21, 2017, with the claim that his test also confirms general relativity, and some hail this as further proof of general relativity. But Domingos Soares (2019) of the Physics Dept. of Federal University of Minas Gerais, when comparing Eddington’s claim with that of Bruns, wrote “the impossibility of a conclusive result therein will clearly emerge.”[21] We know that light bends when going through mediums of different densities on angle. The hot, interplanetary gases surrounding the sun are responsible for the observed deviation, similar to the fact that the starlight of our sun traveling through Earth’s atmosphere makes the upper limb of the sun visible at sunrise when it is actually more than 34 minutes of a degree below the horizon.[22]

# 8.1 Length contraction

“In relativistic mechanics, there is no such thing as absolute length, or absolute time.”[23] “Length contraction suggests that objects in motion are measured to have a shorter length than they do at rest. No actual shrinkage is implied, merely a difference in measured results, just as two observers in relative motion measure a different frequency for the same source of sound (the Doppler effect).”[24]

# 8.2 Comments on length contraction

“No actual shrinkage is implied, merely a difference in measured results.” Oh! What nonsense that students are being indoctrinated to believe. Many, if not most, other physics texts do not suggest that length contraction is not actual.

Since distance, time, and velocity are mathematically related, and the velocity of light is constant, a change in the value of time in time dilation requires a change in the value of distance, length contraction. The comments on sections 5.2 and 6.2 should be clear enough for you to recognize the impossibility of time dilation, and that would invalidate claims of length contraction, relativistic mass, relativistic energy, relativistic momentum, etc.

# 9.1 The Hafele-Keating experiment

“Time dilation has been confirmed also for not-so-fast motion.”[25] Four cesium clocks were, in 1971, flown around the world both in an eastward, and again in a westward direction, in what is referred to as the Hafele-Keating experiment. Compared to the U.S. Naval Observatory clock, the clocks are reported to have gained 273 ± 7 nanoseconds on the westward flight, and to have lost 59 ± 10 nanoseconds on the eastward flight. "These results provide an unambiguous empirical resolution of the famous clock "paradox" with macroscopic clocks."[26]

## 9.2 Comments on the Hafele-Keating experiment

It is impossible for an experiment to prove a deceptive theory true. Dr. Louis Essen, the man who invented the cesium atomic clock, and who was one of two men to determine the number of oscillations of cesium to equal the ephemeris second of 1900 (which became the 1966 definition of the atomic second)(determining the definition to have an error tolerance of ±20 oscillations, although the error tolerance was not included when the definition of the SI second was changed), and who more accurately measured the speed of light, proving it to be 16 km/s higher than what the scientific community believed, stated that special relativity "would retard the rational development of science."[27] He was right, as so much of the student's time has been wasted on learning a false theory, supported with lies. This man who invented the atomic clock also wrote the following concerning the Hafele-Keating experiment: "Four atomic clocks were flown round the world and the times recorded by them were compared with the times recorded by similar clocks in Washington. The results obtained from the individual clocks differed by as much as 300 nanoseconds. This absurdly optimistic conclusion was accepted and given wide publicity in the scientific literature and by the media as a confirmation of the clock paradox. All the experiment showed was that the clocks were not sufficiently accurate to detect the small effect predicted."[28]

Within years of adopting the atomic second, scientists observed that elevated atomic clocks were not staying in sync with those of lower elevations. Some attributed this to special and general relativity, yet the cause is gravity affects the resonant frequency of every isotope. The IAU in 1976 revised the definition of the System International, SI, second, effective 1977, such that it is the atomic second at sea level.[29] The atomic clocks used in the Hafele-Keating experiment were not counting what is now SI seconds. Therefore, it would be incorrect to compare them to a clock that was counting accurately. That statement is also true for those atomic clocks in GPS satellites. Unless the elevated or moving atomic clock, such as is in a GPS satellite, uses a different definition of the number of oscillations to equal a second, it will not be counting SI seconds and would accrue instrument error which would require regular corrections. The student was told the Hafele-Keating experiment confirmed time dilation. Why did the author lie?

## 10.1 Muons

"Very detailed studies have been done on unstable particles called muons that are created at the top of the atmosphere, at a height of about 60 km, when high-energy cosmic rays collide with air molecules….. We wouldn't detect muons at the ground at all if not for time dilation."[30]

## 10.2 Comments muons

Hailed as evidence of time dilation and special relativity, this claim concerning muons is a prime example of dumbed-down education, where the claim is that muons could not reach Earth's surface before decaying if it were not for the distance getting shorter from the muon's perspective, length contraction, and the time of flight for the muon getting shorter from the ground's perspective, time dilation. But what they are actually saying with their dumbed down philosophy is that if were not for length contraction, where a mile is no longer a mile, muons would be traveling at about fifteen times the speed of light. Meteorites fall through Earth's atmosphere at a much slower velocity and most burn up, but the miraculous muons can do it because of special relativity. The claimants ignore the fact that if cosmic rays can cause muons to be formed at an elevation of 60 km (some physics texts say 100 km, while another states 4.8 km), these same rays can cause them to be formed at ground level. Ah! But the UV solar radiation is more intense at the upper atmosphere, causing more muons to form there.

## 11.1 The precession of Mercury

"From the special theory of relativity, we know that measurements of space and time undergo transformations when motion is involved. Likewise with the general theory: Measurements of space differ in different gravitational fields – for example, close to and far away from the Sun….Careful measurements showed that Mercury's orbit precesses about 574 seconds of arc per century. Perturbations by other planets were found to account for all but 43 seconds of arc per century….And then came the explanation of Einstein, whose general relativity field equations applied to Mercury's orbit predict an extra 43 seconds of arc per century!"[31] "Careful observations of Mercury's orbit during the 1800's showed that it does indeed precess, with each precession cycle taking more than 20,000 years."[32]

# 11.2 Comments on the precession of Mercury

"The Sun will often move outside of its average location by over a full radius. For some reason this never appears to be considered when modelling Mercury's perihelion motion."[33] We now know that there are some minor planets which cross Mercury's orbit and could be partly responsible for the deviation. In fact, the Jet Propulsion Lab Small-Body database lists 362 Mercury "crossers" and 561 Mercury "grazers"[34], with the distinction between the two on how they interact with Mercury's orbit. At least two of these minor planets which cross Mercury's orbit, 1998 $RO_1$ and 1999 $KW_4$, have their own moon. Thus, general relativity does not add anything to the science of Mercury's orbit. It is not enough to calculate the effect of just one or two of these minor planets with the conclusion that since their gravitational influence is not enough to be responsible for Mercury's precession, general relativity must still be true. The effects of all 923 minor planets must be considered along with the sun's movement from its average location.

# 12.1 Problems resulting from the deception of special relativity

Most people critical of Einstein's theory are ridiculed and several in the academic community have had their career threatened if they speak against this god of science. That is a human behavior problem. The deception of special relativity is foundational to many other deceptions the student is taught: general relativity; the Schwarzschild radius from which black holes are theorized; curvature of space; fabric of space; time as a fourth dimension; gravity bending light; spacetime, gravity affecting time; relativistic momentum; relativistic mass; wormholes; gravitational lensing; event horizon, the Lorentz transformation, the Minkowski calculations; etc. Time wasted learning these deceptions, along with learning evolution, the Big Bang, false claims as to the number of stars and stellar distances, is time in which the student is not learning truth.

Nothing has meaning, unless it is relative. And that especially includes words. Without reference to sci-fi movies, words such as Klingon and cyborg, have no meaning to the listener. Teaching the student into thinking simultaneous does not mean simultaneous for everyone, that one observer can be inertial while the other is not and yet use the Lorentz formula, which only applies when both observers are inertial, that the length of your measuring stick is not the same, or that people can have their own time is deceiving the student into not believing in absolute truth, truth is relative, you can have your truth while I have mine. Unfortunately, the deceived student then parrots their deception (thoughtless repeating) to others as adults, instead of applying God's instruction: "prove all things, hold fast that which is good." (If you cannot find an error in the comments on sections 5.2 and 6.2, you should do as the child who had been deceived into believing in Santa does when he learns the truth: he quits believing the lie.) (Just as parents who deceive their children should be ashamed for lying, so also should anyone who teaches special or general relativity to be true. The Almighty Creator God hates lying.[35])

# 13 Conclusion

*"A man may imagine things that are false, but he can only understand things that are true, for if the things be false, the apprehension of them is not understanding."* Isaac Newton

Although "there have been hundreds of papers and dozens of books written on the refutation of special relativity over the last 100 years,"[36], I know of none pointing out some of the textbook deceptions used to teach the theory, or that it is based upon the misconception that light could have the same forward momentum as objects with physical mass. The examples presented represents the more common of the deceptions used to teach special relativity, followed by explanations revealing the deception. If the theory were true, the deceptions would not be necessary to explain it. Some people accuse Einstein of plagiarism, as he did not give credit to his sources. If he was guilty of plagiarism with the theory of relativity, it is likely he is guilty of it with his other papers. Others point out the mathematical errors Einstein made within the theory. As Dr. Louis Essen stated: "Einstein's use of a thought experiment, together with his ignorance of experimental techniques, gave a result which fooled himself and generations of scientists." Every area of science which uses the false equation of Lorentz should be viewed as producing deceptively false conclusions.

While Einstein had said that "the distinction between the past, present, and future is only an illusion",[37] I am confident my future with Jesus is not an illusion, nor what He has done for me in the present and in the past. You also can have confidence in the eternal promises of God, if you repent of your sins and believe that Jesus paid for your sins, or you can have the curses if you chose the path to hell.

# Acknowledgements

I would like to thank my wife, Karen, who has endured my years of studying this subject for how to expose special relativity in a fashion most honest people would understand this theory as deceptively false.


# References


[1] Parallax, Alan Hirshfield, c.2001, pg xii
[2] University Physics, Hugh Young & Riger Freedman, 9th edition, ©1996, page 1198.
[3] Conceptual Physics, tenth edition by Paul Hewitt, ©2009, pages 696-702.
[4] Science From A Different Perspective, Denis Thomas, ©2016, page 68.
[5] The Cosmic Perspective, 2nd edition, Jeffrey Bennet, Megan Donahue, Nicholas Schneider, Mark Voit, ©2002, page 410
[6] College Physics, Randy Knight, Brian Jones, & Stuart Field, ©2017, pages 888-891
[7] College Physics Reasoning and Relationships, Nicholas Giordano,©2010, pages 927-928.
[8] College Physics, Randy Knight, Brian Jones, & Stuart Field, ©2017, pages 886.

[9] Einstein's Clocks, Poincare's Maps, ©2003, page 13
[10] https://rb.gy/finf2n Accessed 17 Dec 2022
[11] University Physics, Hugh Young & Riger Freedman, 9th edition, ©1996, page 1194.
[12] Physics, James Walker, ©2002, page 948.
[13] https://rb.gy/vwxf1x Accessed 28 Nov 2022
[14] Conceptual Physics, tenth edition by Paul Hewitt, ©2009, page 694.
[15] College Physics, 9th edition, volume 2, Raymond Serway & Chris Vuille, ©2012, page 891.
[16] https://philpapers.org/rec/THOTSO-52 Accessed 17 Dec 2022
[17] https://www.liquisearch.com/woldemar_voigt/the_voigt_transformation Accessed 28 Nov 2022
[18] https://rb.gy/xqqx3y Accessed 17 Dec 2022
[19] Conceptual Physics, tenth edition by Paul Hewitt, ©2009, page 723.
[20] Soares, Domingos. "The 1919 Eddington eclipse." Page 5. doi:10.13140/RG.2.2.33288.88321
[21] Ibid.
[22] Navigation and Nautical Astronomy, Benjamin Dutton, 1943, section 1252, page 324.

[23] College Physics, 9th edition, volume 2, Raymond Serway & Chris Vuille, ©2012,
page 889.
[24] Physics, Volume 1, fourth edition, Robert Resnick, David Halliday, Kenneth Krane, ©1992, page 472
[25] Conceptual Physics, tenth edition by Paul Hewitt, ©2009, page 695.
[26] Principles of Physics, Frank Blatt, ©1983, page 667.
[27] https://rb.gy/xmaujm Accessed 28 Nov 2022
[28] Ibid.
[29] XVIth General Assembly Grenoble, France 1976. page 11, Recommendation 5b https://www.iau.org/static/resolutions/IAU1976_French.pdf Accessed 28 Nov 2022
[30] College Physics, Randy Knight, Brian Jones, & Stuart Field, ©2017, pages 890.
[31] Conceptual Physics, tenth edition by Paul Hewitt, ©2009, page 727.
[32] The Cosmic Perspective, Second edition, Jeffrey Bennett, Negan Donathue, Nicholas Schneider, & Mark Voit, ©2002, page 419.
[33] "Mercury's Perihelion Advance." https://rb.gy/2d7nr8. Accessed 28 Nov 2022
[34] "List of Mercury-crossing minor planets." https://rb.gy/bxhual.
Accessed 4 Nov 2022.
[35] Proverbs 6:16-19; Revelation 21:8
[36] "Special Relativity." https://rb.gy/6qfeqx. Accessed 28 Nov 2022.
[37] https://rb.gy/gcsizt Accessed 17 Dec 2022

# A logical-metaontological approach to the problem of (meta)data veracity in systems for automatic extraction of metadata from scientific-legal articles

(Un approccio logico-metaontologico al problema della *(meta)data veracity* nei sistemi di estrazione automatica di metadati da articoli scientifico-giuridici)

Simone Cuconato*

**Abstract**

In an increasingly data-driven world, the question of data – or metadata – veracity is now a central issue not only in the world of information but also in the legal one. Data veracity describes a closeness to truth on a higher level than a measure such as accuracy does. High veracity data is data that can be relied upon when making decisions, thus reducing the risk of basing choices on untrue information. The article uses epistemic logic **T** to model structured metadata automatically extracted from legal papers, and the tools of metaontology to propose a definition of veracity as truthmaker.

**Keywords**: (meta)data veracity; truthmaker; applied epistemic logic; metadata modelling§

* Department of Computer Engineering, Modelling, Electronics and Systems Engineering, University of Calabria and Istituto di Informatica e Telematica (IIT) - CNR, Via P.Bucci, 87036, Rende (CS), Italy; simone.cuconato@unical.it.

§ Received on May 9th, 2022. Accepted on December 28th, 2022. Published on December 30th, 2022. DOI: 10.23756/sp.v10i1.784. ISSN 2282-7757; eISSN 2282-7765. 

*Simone Cuconato*

**Sunto**

In un mondo sempre più guidato dai dati, la questione della veridicità dei dati – o metadati – è una questione ormai centrale non solo nel mondo dell'informazione, ma anche in quello giuridico. La veridicità dei dati descrive una vicinanza alla verità ad un livello più alto di una misura come l'accuratezza. I dati ad alta veridicità sono dati su cui si può fare affidamento quando si prendono decisioni, riducendo, in questo modo, il rischio di fondare le proprie scelte su informazioni non veritiere. L'articolo usa la logica epistemica **T** per modellare metadati strutturati estratti automaticamente da articoli scientifico-giuridici, e gli strumenti della metaontologia per proporre una definizione di *veracity* come *truthmaker*.

**Parole chiave:** *(meta)data veracity*; *truthmaker*; logica epistemica applicata; modellizzazione dei metadati

# 1. Introduzione

Sono passati più di dieci anni da quando Chris Anderson, allora caporedattore dell'influente rivista tecnologica *Wired*, pubblicò un articolo intitolato "*The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*"[1]. L'articolo di Anderson è diventato rapidamente il manifesto ideologico dell'entusiasmo "datacentrico" ed è articolato lungo due punti chiave.

Primo: "fidatevi, è conveniente". Google ci ha insegnato che non è importante capire perché una pagina web è "migliore" di un'altra, ma è sufficiente fidarsi dell'ordinamento prodotto dall'algoritmo PageRank. La comodità di ricevere una risposta molto semplice a una domanda potenzialmente molto complicata, senza dover necessariamente sviluppare alcuna analisi semantica o causale, è diventata presto la chiave del successo di Google.

Secondo: "i modelli scientifici sono obsoleti". La disponibilità senza precedenti di dati prodotti più o meno consapevolmente da tutti noi ci permette di ripensare radicalmente la relazione tra i dati e i meccanismi che li generano.

---

[1] [2].

Secondo Anderson possiamo smettere di cercare modelli: invece di procedere per "congetture” e “confutazioni" nello spiegare le osservazioni, il diluvio di dati ci permette di rinunciare al laborioso compito di costruire modelli per i fenomeni di interesse, in favore del compito molto più facile di analizzare le correlazioni individuate da sofisticati algoritmi statistici.

In questo lavoro, ci muoveremo in una direzione opposta a quella tracciata da Anderson.

Primo: “non fidarsi”. Oggi più che mai è necessario porre l’accento sulla qualità delle informazioni e sulla veridicità dei dati: un’analisi semantica è necessaria.

Secondo: “i modelli scientifici sono fondamentali”. È molto difficile pensare ai dati senza che essi rispondano a un'ipotesi di modellizzazione. L’idea semplicistica che *petabyte* di dati possano essere autosufficienti e che i dati possano essere visti come un sostituto della modellazione scientifica non è sostenibile.

In particolare, in questo contributo: *i*) proporremo un modello basato sulla logica epistemica per formalizzare i metadati estratti da articoli scientifico-giuridici tramite sistemi di estrazione automatica fondati sull’intelligenza artificiale; *ii*) porremo l’accento sul tema della qualità dei dati grazie alla definizione di un principio metaontologico di veridicità come *truthmaker*. Se fino a pochi anni fa il costo dell'informazione era l'aspetto più rilevante, al contrario, oggi la qualità delle informazioni è diventata più importante che mai. Per questo motivo, la veridicità dei dati – nel nostro caso metadati – è stata proposta come la quarta "V" – accanto a Volume, Varietà e Velocità – dei *big data*[2].

Ma cosa sono i metadati? Nel mondo dell’informazione, i metadati rappresentano la base informativa di “secondo livello”, che descrive, struttura e gestisce i dati primari o le informazioni su cui vengono appoggiate le risorse

[2] Si veda [11].

informative[3]. Attualmente i metadati sono necessari non solo per gestire, conservare e reperire gli oggetti informativi, ma rappresentano le pedine fondamentali nel *Semantic Web* avendo un ruolo chiave nell'indicizzazione e nell'identificazione, nella classificazione e nella catalogazione, nella conservazione, nella verifica dell'integrità e dell'affidabilità e nella gestione dei diritti, nonché nella distribuzione, nella ricerca e nel recupero delle risorse digitali. I metadati, siano essi descrittivi, strutturali, amministrativi o per la *long term digital preservation*, alla fine, sono accomunati da un unico obiettivo multifunzionale: quello di contribuire a una gestione e conservazione più chiara e modulare degli oggetti digitali. La metadatazione automatica permette di estrarre direttamente i metadati dalle fonti documentali. L'estrazione dei metadati come tecnologia fondamentale per il processo automatico dei documenti ha avuto un grande successo in numerose applicazioni e domini. Molte delle soluzioni proposte da tali sistemi sono basate tecniche sub-simboliche di intelligenza artificiale, come il *machine learning* (ML).

In generale, le modalità con cui il ML permette agli algoritmi di fare apprendimento con i dati sono classificate in cinque categorie: *i*) apprendimento supervisionato, nel quale vengono presentati al modello scelto gli esempi formati dagli input e relativi output desiderati con lo scopo di far apprendere una regola generale in grado di mappare gli input negli output; *ii*) apprendimento non supervisionato, nel quale vengono forniti al modello scelto solo gli esempi formati dagli input, senza alcun output atteso, con lo scopo di fargli apprendere in autonomia una qualche struttura nei dati d'ingresso; *iii*) apprendimento semi-supervisionato, nel quale vengono combinati i due approcci precedenti con una prima fase supervisionata sui dati aventi input e output associato, e una successiva fase non supervisionata su dati di cui non si conosce l'output associato; *iv*) apprendimento con rinforzo, con il quale si interagisce con un ambiente dinamico in cui raggiungere un certo obiettivo e a mano a mano che si esplora il dominio del problema vengono forniti dei *feedback* in termini di ricompense o punizioni secondo il comportamento eseguito; *v*) apprendimento con trasferimento, nel quale il modello scelto impara ad affrontare un certo

[3] Una delle migliori introduzioni ai metadati è [13].

problema generico e, successivamente, si prende la conoscenza creata per usarla nell'affrontare un altro problema simile o più specifico. La figura 1 mostra una tassonomia dei metodi di apprendimento automatico.

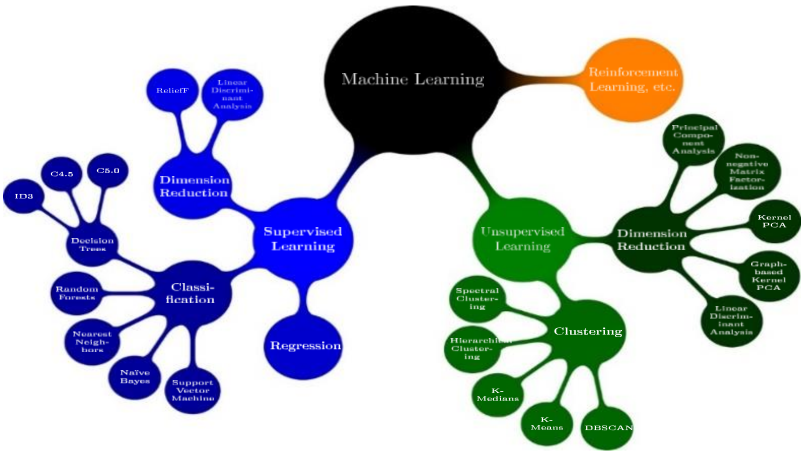



**Figura 1** Una illustrazione schematica della tassonomia dei metodi di ML[4]

Sempre più sistemi di estrazione automatica di metadati basati su tecniche di apprendimento automatico sono diventati strumenti centrali nel mondo dell'informazione. In questo lavoro, useremo CERMINE per estrarre metadati da articoli scientifico-giuridici. CERMINE è un *framework open-source* per estrarre metadati strutturati da articoli scientifici nativi digitali. Il *framework* è basato su un *workflow* modulare e le implementazioni della maggior parte dei passi sono basate su tecniche di apprendimento automatico supervisionato e non supervisionato. Il *workflow* modulare, rappresentato in figura 2, consiste in tre percorsi (*ii* e *iii* eseguiti in parallelo): *i*) il percorso di estrazione della struttura di base richiede un file pdf come input e produce una struttura geometrica gerarchica in formato TrueViz. TrueViz è uno strumento in grado di classificare le entità di ogni pagina della struttura in quattro categorie: zone, linee, parole e caratteri. A sua volta, ogni zona è etichettata secondo altre quattro categorie: metadati, riferimenti, corpo e altro; *ii*) il percorso di estrazione dei metadati analizza le parti di metadati della struttura struttura geometrico-gerarchica, il risultato è un insieme di metadati del documento in formato XML; *iii*)

[4] Figura tratta da [1].

l'estrazione dei riferimenti estrae una lista di riferimenti bibliografici dal documento.

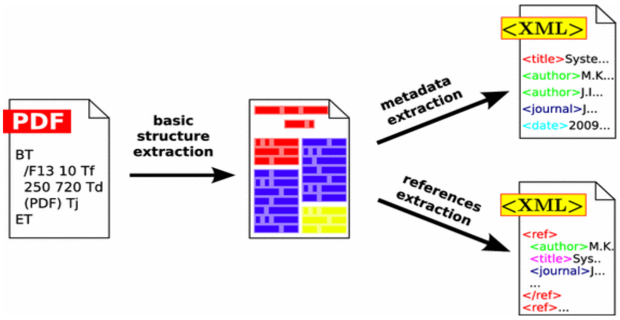



**Figura 2** L’architettura del *workflow* di CERMINE[5]

Nel contesto degli articoli di ricerca, i metadati sono di solito di natura descrittiva e detengono una grande importanza poiché forniscono una breve panoramica su un articolo scientifico mostrando informazioni come il titolo, i suoi autori, la rivista, la bibliografia, ecc. Spesso, i ricercatori tendono a decidere la pertinenza dell'articolo con il loro dominio di interesse basandosi sulle informazioni dei metadati. Lo scopo di questo articolo è quello di applicare la logica epistemica ai sistemi di estrazione automatica di metadati da articoli scientifico-giuridici e di proporre una definizione di veridicità come *truthmaker*.

## 2. Logica epistemica standard

La logica epistemica è un’estensione della logica classica che ha come oggetto di studio gli enunciati di credenza e di sapere[6]. Nell’epistemologia contemporanea è ampiamente condivisa l’idea secondo cui la verità è una condizione necessaria della conoscenza. Per tale motivo: *i*) si dice che la conoscenza è *fattiva*, ossia si presuppone la verità della proposizione conosciuta; *ii*) perché si abbia conoscenza è necessario intrattenere una credenza; *iii*) la credenza deve essere giustificata. Per lungo tempo la verità, la credenza e la

[5] Figura tratta da [16].

[6] Per un’introduzione italiana alle logiche intensionali e modali si veda[8], [9], [17], mentre per uno studio più mirato alla logica epistemica si veda [18].

giustificazione sono state considerate condizioni congiuntamente sufficienti perché si abbia conoscenza. Dagli anni ’60 in poi, grazie ai lavori di Gettier[7], gli epistemologi contemporanei hanno sostenuto che, oltre alle tre suddette condizioni, ne occorrano altre[8]. Tuttavia, per quanto i logici siano particolarmente interessati al complesso dibattito che si è sviluppato tra gli epistemologi riguardo alla strategia da adottare per caratterizzare esaustivamente la conoscenza, nelle logiche epistemiche la conoscenza è generalmente caratterizzata come semplice credenza vera. In questo modo, i logici trattano le attribuzioni di conoscenza e credenza come formule contenenti operatori modali. Semanticamente questo significa che, nel valutare il valore di verità di una formula associata a un operatore epistemico, si prenda in esame un insieme di circostanze alternative a quelle attuali. Tali circostanze alternative prendono in letteratura il nome di *mondi possibili*. Poniamo per esempio che un soggetto creda che Mario Draghi sia il presidente del Consiglio italiano e che Barack Obama sia il Presidente degli Stati Uniti. I mondi compatibili con le sue credenze saranno tutti e soli i mondi in cui è vero che Mario Draghi è il presidente del Consiglio italiano e che Barack Obama è il Presidente degli Stati Uniti. Ma in base alla semantica dei mondi possibili il possesso o meno della conoscenza dipende da come stanno le cose nel mondo attuale: il nostro soggetto non può sapere che Barack Obama è il Presidente degli Stati Uniti, dato che è falso.

Sintatticamente, il linguaggio della logica epistemica proposizionale è il linguaggio della logica proposizionale classica con l’aggiunta di uno specifico operatore epistemico unario tale che

$K_a\varphi$ si legge “l’agente $a$ sa che $\varphi$”

In generale, un agente può essere una persona reale, un giocatore in un gioco, un robot, una macchina, un “processo” o, nel nostro caso, un *framework* di estrazione automatica di metadati. Hintikka ha fornito una prima pioneristica formalizzazione delle attribuzioni di credenza in un linguaggio modale

[7] [10].

[8] Approfondire tale tema ci porterebbe lontani dagli scopi di questo lavoro. Rimandiamo a chi fosse interessato a un recente approccio formale al problema di Gettier a [20].

sfruttando delle strutture semantiche dette *model set*. Tuttavia, sarà solo dopo le pubblicazioni dei lavori di Kripke che Hittinka elaborerà un'interpretazione semantica degli operatori epistemici che possiamo presentare in termini di semantica standard dei mondi possibili secondo le seguenti linee:

> $K_a\varphi$: è vera in un mondo possibile $w$ a condizione che $\varphi$ sia vera in tutti i mondi compatibili con le credenze intrattenute dal soggetto epistemico in $w$.

Pertanto l'idea intuitiva delle logiche modali epistemiche è associare a un dato soggetto epistemico un insieme di mondi, che corrispondono a tutte le situazioni compatibili con le credenze del soggetto stesso. Vediamo ora come catturare queste intuizioni in termini formali.

**Definizione 1** [Sintassi di $\mathcal{L}_K$] Dato un insieme $\mathcal{P}$ di variabili proposizionali ed un insieme finito di agenti $\mathcal{A}$, definiamo il linguaggio epistemico $\mathcal{L}_K$ come segue:

$$\varphi := p|\neg\varphi|\varphi \wedge \varphi|K_a\varphi$$

Dove $p \in \mathcal{P}$ e $a \in \mathcal{A}$.

**Definizione 2** [Modello epistemico] Dato $\mathcal{P}$ ed $\mathcal{A}$ un modello epistemico $M{:}\langle W, R^{\mathcal{A}}, V^{\mathcal{P}}\rangle$ è una tripla dove

- $W \neq \emptyset$ è un insieme di mondi possibili $w_i$, a volte chiamato il dominio di $M$, e denotato $\mathcal{D}(M)$;
- $R^{\mathcal{A}}$ è una funzione, che produce una relazione di accessibilità $R_a \subseteq W \times W$ per ogni agente $a \in \mathcal{A}$;
- $V^{\mathcal{P}}{:}\, W \rightarrow (\mathcal{P} \rightarrow \{vero, falso\})$ è una funzione tale che per ogni $p \in \mathcal{P}$ e per ogni $w_i \in W$, determina quale sia il valore di verità $V^{\mathcal{P}}(w_i)(p)$ di $p$ nel mondo possibile $w_i$.

**Definizione 3** [Verità nel modello] Dato un modello $M{:}\langle W, R^{\mathcal{A}}, V^{\mathcal{P}}\rangle$ definiamo la verità di una formula $\varphi$ $M, w_i \vDash \phi$ come segue:

| | | |
|---|---|---|
| $M, w_1 \vDash p$ | sse | $V(w_1)(p) = vero$ con $p \in \mathcal{P}$ |
| $M, w_1 \vDash \varphi \wedge \psi$ | sse | $M, w_1 \vDash \varphi$ e $M, w_1 \vDash \psi$ |

| | | |
|---|---|---|
| $M, w_1 \vDash \neg\varphi$ | sse | non $M, w_1 \vDash \varphi$ (spesso scritto $M, w_1 \nvDash \varphi$) |
| $M, w_1 \vDash K_a\varphi$ | sse | $M, w_2 \vDash \varphi$ per ogni $w_2$ tale che $w_1 R_a w_2$ |

**Definizione 4** [Assiomi e regole di inferenza] Il sistema di prova della logica epistemica che useremo è assiomatizzato utilizzando gli assiomi di **T** e la regola del *modus ponens* e della necessitazione come riportato in tabella 1:

| Sistema | Regole | Assiomi | Proprietà di R |
|---|---|---|---|
| **T** | **MP** e **Nec** | $K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$<br>$K_a\varphi \to \varphi$ | $R$ riflessiva |

**Tabella 1** Logica epistemica **T**

La riflessività di $R$ garantisce che il principio

**T** $K_a\varphi \to \varphi$

sia valido.

# 3. Logica e metaontologia dei metadati

Vediamo ora come adattare la logica epistemica standard alla modellizzazione dei metadati[9]. A livello sintattico nel nostro modello avremo solo una particolare tipologia di proposizioni $p_{\mathcal{E}}$

$$p_{\mathcal{E}} =_{def} \mathcal{E}_{m_i}^{d_i}$$

[9] Per un'applicazione della logica epistemica ai sistemi di estrazione automatica di metadati da articoli scientifici sul Covid-19 si veda [6].

dove $\mathcal{E}_{m_i}^{d_i}$ si legge "estrae il metadato $m_i$ dal documento $d_i$".

**Definizione 5** [Sintassi di $\mathcal{L}_{K_\mathcal{E}}$] Dato un insieme $\mathcal{P}_\mathcal{E}$ di variabili proposizionali ed un insieme finito di *framework* $\mathcal{F}$, definiamo il linguaggio epistemico $\mathcal{L}_{K_\mathcal{E}}$ come segue

$$\varphi := p_\mathcal{E} | \neg\varphi | \varphi \wedge \varphi | K_a\varphi$$

Dove $p_\mathcal{E} \in \mathcal{P}_\mathcal{E}$ e $a \in \mathcal{F}$.

A livello semantico, invece, sostituiremo il concetto di mondo possibile con quello di *estrazione possibile*. L'idea intuitiva alla base dell'applicazione della logica epistemica alla modellizzazione dei metadati è associare a un dato *framework* un insieme di possibili estrazioni, che corrispondono a tutte le situazioni compatibili con le credenze del *framework* stesso.

**Definizione 6** [Modello epistemico per metadati] Dato $\mathcal{P}_\mathcal{E}$ e $\mathcal{F}$ un modello epistemico per metadati $M: \langle E, R^\mathcal{F}, V^{\mathcal{P}_\mathcal{E}} \rangle$ è una tripla dove

- $E \neq \emptyset$ è un insieme di estrazioni possibili $e_i$;
- $R^\mathcal{F}$ è una funzione, che produce una relazione di accessibilità $R_a \subseteq E \times E$ per ogni agente $a \in \mathcal{F}$;
- $V^{\mathcal{P}_\mathcal{E}}: E \rightarrow (\mathcal{P}_\mathcal{E} \rightarrow \{vero, falso\})$ è una funzione tale che per ogni $p_\mathcal{E} \in \mathcal{P}_\mathcal{E}$ e per ogni $e_i \in E$, determina quale sia il valore di verità $V^{\mathcal{P}_\mathcal{E}}(e_i)(p_\mathcal{E})$ di $p_\mathcal{E}$ nell'estrazione possibile $e_i$.

**Definizione 7** [Verità del modello epistemico per metadati] Dato un modello epistemico per metadati $M: \langle E, R^\mathcal{F}, V^{\mathcal{P}_\mathcal{E}} \rangle$ definiamo la verità di una formula $\varphi$ $M, e_i \vDash \varphi$ come segue:

| | | |
|---|---|---|
| $M, e_1 \vDash p_\mathcal{E}$ | sse | $V(e_1)(p_\mathcal{E}) = vero$ con $p_\mathcal{E} \in \mathcal{P}_\mathcal{E}$ |
| $M, e_1 \vDash \varphi \wedge \psi$ | sse | $M, e_1 \vDash \varphi$ e $M, e_1 \vDash \psi$ |
| $M, e_1 \vDash \neg\varphi$ | sse | non $M, e_1 \vDash \varphi$ |
| $M, e_1 \vDash K_a\varphi$ | sse | $M, e_2 \vDash \varphi$ per ogni $e_2$ tale che $e_1 R_a e_2$ |

**Definizione 8** [Assiomi e regole di inferenza] Il sistema di prova della logica epistemica che useremo è assiomatizzato utilizzando gli assiomi di **T** e la regola del *modus ponens* e della necessitazione come riportato in tabella 2:

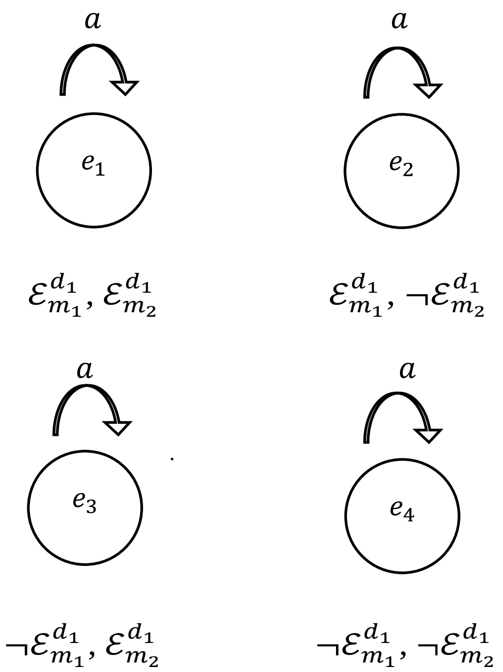
| **Sistema** | **Regole** | **Assiomi** | **Proprietà di R** |
|---|---|---|---|
| **T** | **MP** e **Nec** | $K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$<br><br>$K_a\varphi \to \varphi$ | $R$ riflessiva |

**Tabella 2** Logica epistemica **T**

**Definizione 9** [Struttura $\mathcal{S}$] Una struttura $\mathcal{S}$ è della forma $\mathcal{S} = \langle \mathcal{F}, E, \mathcal{P}_{\mathcal{E}}, M, D \rangle$, dove

$\mathcal{F} = \{a, b, c, \dots\}$ è un insieme non vuoto di *framework* di estrazione automatica di metadati,
$E = \{e_1, \dots, e_m\}$ è un insieme non vuoto di possibili estrazioni ($|E| = m \in \mathbb{N}$),
$\mathcal{P}_{\mathcal{E}} = \{p_{\mathcal{E}_1}, \dots, p_{\mathcal{E}_m}\}$ è un insieme non vuoto di proposizioni ($|\mathcal{P}_{\mathcal{E}}| = m \in \mathbb{N}$),
$M = \{m_1, \dots, m_m\}$ è un insieme non vuoto di metadati ($|M| = m \in \mathbb{N}$),
$D = \{d_1, \dots, d_m\}$ è un insieme non vuoto di documenti ($|D| = m \in \mathbb{N}$).

$\mathcal{S}$ è una struttura nella quale occorrono possibili estrazioni $E$. $\mathcal{F}$ è l'insieme dei *framework* di estrazione automatica di metadati, mentre $\mathcal{P}_{\mathcal{E}}$ è l'insieme delle proposizioni. Infine, $M$ è l'insieme dei metadati e $D$ l'insieme dei documenti (nel nostro caso di articoli scientifico-giuridici). All'interno della struttura possiamo rappresentare un modello relazionale usando un grafo in cui le possibili estrazioni sono nodi e la relazione epistemica è indicata tramite frecce come illustrato in figura:




In questo grafo abbiamo una situazione nella quale dato un documento in entrata e due metadati, un agente di estrazione sa che si possono verificare quattro possibili estrazioni: l'estrazione in cui entrambi i metadati vengono estratti correttamente, l'estrazione in cui il metadato uno viene estratto correttamente mentre il due no, l'estrazione in cui il metadato due è estratto correttamente mentre l'uno no, ed infine l'estrazione in cui entrambi i metadati non sono riportati correttamente. Analizziamo, ora, più in dettaglio cosa vuol dire che in una estrazione una proposizione è vera o falsa. Come già sappiamo, la verità di una formula proposizionale dipende "dalla situazione del mondo", o nel caso di una proposizione epistemica "è vera in $w$ a condizione che sia vera in tutti i mondi accessibili da $w$". Le situazioni sono formalizzate usando valutazioni e in $\mathcal{S}$ sappiamo che una proposizione $p_{\mathcal{E}}$ "è vera in $e$ a condizione che sia vera in tutte le possibili estrazioni accessibili da $e$"

$$V^{\mathcal{P}_{\mathcal{E}}}: E \to (\mathcal{P}_{\mathcal{E}} \to \{vero, falso\})$$

Inoltre, poiché sappiamo che $p_{\mathcal{E}}$ ha la forma $\mathcal{E}_{m_i}^{d_i}$ scriveremo che è vero (V) o falso (F) che "nell'estrazione $e_i$ un framework estrae il metadato $m_i$ dal documento $d_i$" nel seguente modo

$$\underbrace{\mathcal{E}_{m_i}^{d_i}}_{e_i} = \mathrm{V/F}$$

Ma cosa vuol dire che in una estrazione un metadato estratto è vero? Detto altrimenti, che cosa vuol dire che un *framework* estrae correttamente un metadato da un articolo scientifico-giuridico? Per rispondere a queste domande occorre presentare la teoria dei *truthmakers* e definire la veridicità come *truthmaker*. La teoria dei *truthmakers* è una interessante teoria metaontologica proveniente dal mondo della filosofia analitica che esplora la relazione tra ciò che è vero e ciò che esiste[10]. La teoria ha radici profonde nel pensiero occidentale e, da un lato, veicola una nostra intuizione emergente: se, ad esempio, è vero che il cane è sullo zerbino è perché il cane è "di fatto" sullo zerbino; dall'altro, rappresenta l'idea alla base di una celebre teoria della verità, ossia il corrispondentismo:

($\mathcal{C}$) Dire la verità è dire come "stanno le cose nel mondo"

La teoria dei *truthmakers* a cui ci rifaremo in questo lavoro è quella sviluppata dal filosofo australiano David Malet Armstrong in *Truth and Truthmakers*[11]:

($\mathcal{T}$) Per ogni verità, $p$, esiste un ente, $T$, tale che $T$ rende vero $p$ se e solo se non è possibile che $T$ esista e $p$ sia falso

[10] Il termine metaontologia – usato per la prima volta in [19]– indica l'indagine che mira a determinare quale sia il modo di caratterizzare la nozione di ontologia.

[11] [3]. Per una traduzione italiana delle principali opere di Armstrong si veda [7], mentre per un'introduzione al pensiero del filosofo australiano si veda[4]. Per un confronto tra la metaontologia del *Tractatus logico-philosophicus* di Wittgenstein e la teoria dei *truthmakers* di Armstrong rimandiamo a [5]. Infine, è importante specificare che: *i*) la relazione di *truthmaking* non è una relazione univoca: una proposizione può avere molti *truthmakers* e un oggetto può rendere vere molte proposizioni; e *ii*) ai fini di questo articolo non è necessario impegnarsi in una particolare ontologia dei *truthmakers* (come fatti, sostanze, proposizioni vere, ecc.).

Nel nostro dominio possiamo riformulare il principio $\mathcal{T}$ armstronghiano come segue:

($\mathcal{V}$) Per ogni proposizione vera, $p_{\mathcal{E}}$, esiste un documento, $d$, tale che $d$ rende vero $p_{\mathcal{E}}$ se e solo se non è possibile che $d$ esista e $p_{\mathcal{E}}$ sia falso

In questo modo dovremmo ormai essere in posizione di apprezzare pienamente il significato di questo principio. Consideriamo nuovamente il nostro schema

$$(1)\ \underbrace{\mathcal{E}_{m_i}^{d_i}}_{e_i} = \mathrm{V/F}$$

Solleviamo, ora, la tipica *truthmaking question*: in virtù di cosa (1) è vero? Ebbene, in base a $\mathcal{V}$ dire che un *framework* $a$ ha estratto correttamente un metadato $m$ vuol dire che esiste un documento $d$ che "rende vera" l'estrazione $e$.

## 4. Esempio di $\mathcal{S}$

Consideriamo ora in che modo possiamo modellare i metadati estratti da due differenti articoli scientifico-giuridici usando il framework CERMINE[12]. I metadati che terremo in considerazione negli esempi sono: $m_1$ titolo, $m_2$ autore e $m_3$ rivista. Il primo documento $d_1$ riguarda l'utilizzo di modelli bayesiani nell'ambito dell'argomentazione giuridica, invece, il secondo documento $d_2$ analizza il percorso e le ragioni che hanno portato l'Unione Europea ad entrare in una nuova fase del costituzionalismo moderno (ossia il costituzionalismo digitale).

Posto un *framework* $a$, tre metadati $m_1, m_2$ e $m_3$ e due articoli scientifico-giuridici $d_1$ e $d_2$ avremo la seguente struttura $\mathcal{S} = \langle \mathcal{F}, E, \mathcal{P}_{\mathcal{E}}, M, D \rangle$:

$\mathcal{F} = \{a\}$

$E = \{e_1, \dots, e_m\}$;

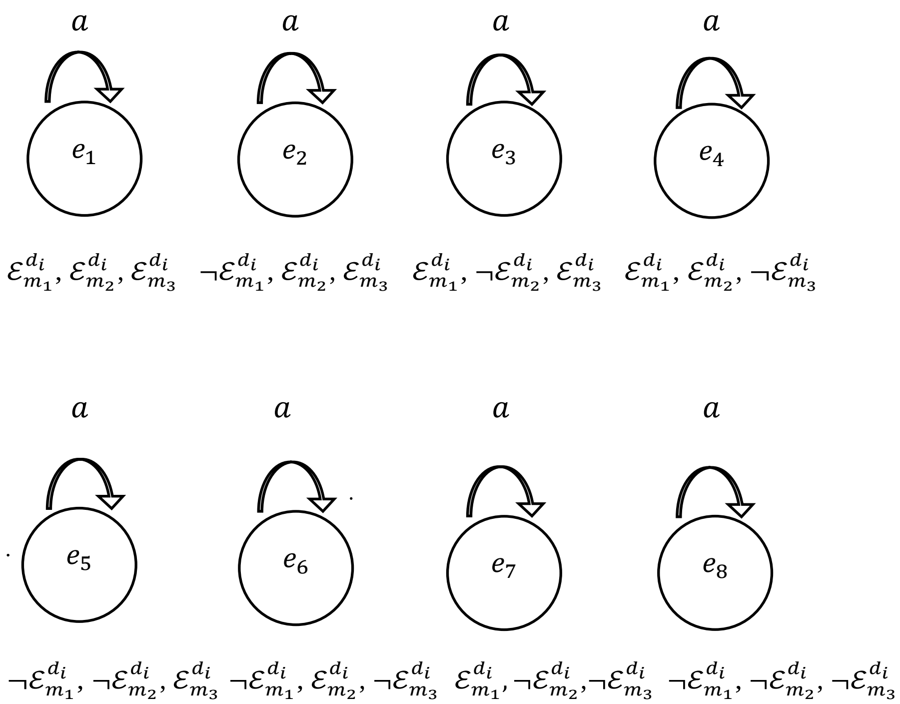
[12] Per l'estrazione dei metadati è stata utilizzata la risorsa gratuita online http://cermine.ceon.pl/index.html.

$\mathcal{P}_{\mathcal{E}} = \{p_{\mathcal{E}_1}, \dots, p_{\mathcal{E}_m}\}$

$M = \{m_1, m_2, m_3\}$

$D = \{d_1, d_2\}$

In $\mathcal{S}$ avremo il seguente universo di possibili estrazioni:




Con il primo documento $d_1$ il *framework* $a$ estrae correttamente tutti i metadati. Nella figura 3 il metadato “rivista” è evidenziato in giallo, il metadato “titolo” in verde e il metadato “autore” in rosso, mentre nella figura 4 è riportato l’XML dell’estrazione operata dal *framework* $a$ con i relativi metadati evidenziati con gli stessi colori


Artificial Intelligence and Law (2019) 27:403–430
https://doi.org/10.1007/s10506-019-09250-3

ORIGINAL RESEARCH

**Modelling competing legal arguments using Bayesian model comparison and averaging**

Martin Neil[1,4] · Norman Fenton[1,4] · David Lagnado[2] · Richard David Gill[3]

Published online: 27 March 2019
© The Author(s) 2019

**Abstract**
Bayesian models of legal arguments generally aim to produce a single integrated model, combining each of the legal arguments under consideration. This combined approach implicitly assumes that variables and their relationships can be represented without any contradiction or misalignment, and in a way that makes sense with respect to the competing argument narratives. This paper describes a novel approach to compare and 'average' Bayesian models of legal arguments that have been built independently and with no attempt to make them consistent in terms of variables, causal assumptions or parameterization. The approach involves assessing whether competing models of legal arguments are explained or predict facts uncovered before or during the trial process. Those models that are more heavily disconfirmed by the facts are given lower weight, as model plausibility measures, in the Bayesian model comparison and averaging framework adopted. In this way a plurality of arguments is allowed yet a single judgement based on all arguments is possible and rational.

**Keywords** Bayesian model comparison and averaging · Bayesian networks · Legal argumentation

Martin Neil
m.neil@qmul.ac.uk

Norman Fenton
n.fenton@qmul.ac.uk

David Lagnado
d.lagnado@ucl.ac.uk

Richard David Gill
gill@math.leidenuniv.nl

1 School of Electronic Engineering and Computer Science, Queen Mary, University of London, London, UK
2 Department of Experimental Psychology, University College London, London, UK
3 Faculty of Mathematics and Natural Sciences, Leiden University, Leiden, The Netherlands
4 Agena Ltd, Cambridge, UK

Springer

**Figura 3** Documento $d_1$

```xml
<?xml version="1.0" encoding="UTF-8"?>
<article xmlns:xlink="http://www.w3.org/1999/xlink">
  <front>
    <journal-meta>
      <journal-title-group>
        <journal-title>Artificial Intelligence and Law</journal-title>
      </journal-title-group>
    </journal-meta>
    <article-meta>
      <title-group>
        <article-title>Modelling competing legal arguments using Bayesian model compa
      </title-group>
      <contrib-group>
        <contrib contrib-type="author">
          <string-name>Martin Neil</string-name>
          <xref ref-type="aff" rid="aff0">0</xref>
          <xref ref-type="aff" rid="aff1">1</xref>
          <xref ref-type="aff" rid="aff2">2</xref>
          <xref ref-type="aff" rid="aff3">3</xref>
        </contrib>
        <contrib contrib-type="author">
          <string-name>Norman Fenton</string-name>
          <xref ref-type="aff" rid="aff0">0</xref>
          <xref ref-type="aff" rid="aff1">1</xref>
          <xref ref-type="aff" rid="aff2">2</xref>
          <xref ref-type="aff" rid="aff3">3</xref>
        </contrib>
        <contrib contrib-type="author">
          <string-name>David Lagnado</string-name>
          <xref ref-type="aff" rid="aff0">0</xref>
          <xref ref-type="aff" rid="aff1">1</xref>
          <xref ref-type="aff" rid="aff2">2</xref>
          <xref ref-type="aff" rid="aff3">3</xref>
        </contrib>
        <contrib contrib-type="author">
          <string-name>Richard David Gill</string-name>
          <xref ref-type="aff" rid="aff0">0</xref>
          <xref ref-type="aff" rid="aff1">1</xref>
          <xref ref-type="aff" rid="aff2">2</xref>
          <xref ref-type="aff" rid="aff3">3</xref>
        </contrib>
```


**Figura 4** XML estrazione documento $d_1$[13]

[13]http://cermine.ceon.pl/cermine/task.html;jsessionid=E3B413EA00B92B9A1080DD1943910FD5?task=6507409268409556231,
http://cermine.ceon.pl/cermine/download.html?type=nlm&task=6507409268409556231.

Poiché $a$ estrae correttamente tutti i metadati si verifica l'estrazione possibile $e_1$:

- $\underbrace{\mathcal{E}_{m_1}^{d_1}}_{e_1} = \mathrm{V}$
- $\underbrace{\mathcal{E}_{m_2}^{d_1}}_{e_1} = \mathrm{V}$
- $\underbrace{\mathcal{E}_{m_3}^{d_1}}_{e_1} = \mathrm{V}$

Invece, con il secondo documento $d_2$ il *framework* $a$ estrae correttamente due metadati su tre. Anche in questo caso, nella figura 5 il metadato "rivista" è evidenziato in giallo, il metadato "titolo" in verde e il metadato "autore" in rosso, mentre nella figura 6 è riportato l'XML dell'estrazione operata dal *framework* $a$ con i relativi metadati evidenziati con gli stessi colori. Si noti come il metadato $m_3$ "rivista" non sia estratto correttamente

© The Author(s) 2021. Oxford University Press and New York University School of Law. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

*The rise of digital constitutionalism in the European Union*

Giovanni De Gregorio*

*In the last twenty years, the policy of the European Union in the field of digital technologies has shifted from a liberal economic perspective to a constitution-oriented approach. This change of heart has resulted primarily from the rise of the information society which has created not only new opportunities but also challenges to fundamental rights and democratic values. Even more importantly, this technological framework driven by liberal ideas has empowered transnational corporations operating in the digital environment to perform quasi-public functions on a global scale. This article analyzes the path and the reasons that have led the European Union to enter a new phase of modern constitutionalism (i.e. digital constitutionalism). The primary goal of this article is to describe the characteristics of this new constitutional phase opposing platform powers, and to outline the potential evolution of European digital constitutionalism in the global context.*

**1. Introduction**

In the last twenty years, the policy of the European Union (EU) in the field of digital technologies has shifted from a liberal economic perspective to a constitutional approach aimed to protect fundamental rights and democratic values. In order to understand this change of heart, it is necessary to frame the debate in the information society which is increasingly subject to the power of public and private actors implementing automated decision-making technologies.[1] If the digital environment, as a new space where information and data flow, has been an opportunity to offer cross-border services and exercise individual freedoms, it has also led to serious challenges for constitutional law. Since the end of the last century, the development of digital technologies has not only challenged the protection of individuals' fundamental rights, such as freedom of expression, privacy, and data protection. Even more importantly, this technological

I•CON (2021), Vol. 19 No. 1, 41–70

**Figura 5** Documento $d_2$

```xml
<?xml version="1.0" encoding="UTF-8"?>
<article xmlns:xlink="http://www.w3.org/1999/xlink">
  <front>
    <journal-meta>
      <journal-title-group>
        <journal-title>I•CON</journal-title>
      </journal-title-group>
    </journal-meta>
    <article-meta>
      <article-id pub-id-type="doi">10.1093/icon/moab001</article-id>
      <title-group>
        <article-title>The rise of digital constitutionalism in the European Union</a
      </title-group>
      <contrib-group>
        <contrib contrib-type="author">
          <string-name>Giovanni De Gregorio</string-name>
          <xref ref-type="aff" rid="aff1">1</xref>
        </contrib>
        <aff id="aff0">
          <label>0</label>
          <institution>John Danaher, The Threat of Algocracy: Reality</institution>
          ,
          <addr-line>Resistance and Accommodation, 29(3) P</addr-line>
        </aff>
        <aff id="aff1">
          <label>1</label>
          <institution>PhD Candidate, University of Milano-Bicocca</institution>
          ,
          <addr-line>Milan</addr-line>
          ,
          <country country="IT">Italy;</country>
          <institution>Academic Fellow, Bocconi University</institution>
          ,
          <addr-line>Bocconi</addr-line>
        </aff>
      </contrib-group>
      <pub-date>
        <year>2021</year>
```


**Figura 6** XML estrazione documento $d_2$[14]

Poiché $a$ estrae correttamente i metadati $m_1$ e $m_2$ ma non il metadato $m_3$ si verifica l'estrazione possibile $e_4$:

- $\underbrace{\mathcal{E}^{d_2}_{m_1}}_{e_4} = \text{V}$
- $\underbrace{\mathcal{E}^{d_2}_{m_2}}_{e_4} = \text{V}$
- $\underbrace{\mathcal{E}^{d_2}_{m_3}}_{e_4} = \text{F}$

# 5. Conclusioni

In questo articolo, a partire dai metadati estratti da articoli scientifico-giuridici tramite sistemi di estrazione automatica di metadati basati sull'intelligenza artificiale, abbiamo usato la logica epistemica e la metaontologia per descrivere (modellare) formalmente i metadati estratti e definire un principio di veridicità come *truthmaker*. In particolare, in ambito giuridico un potenziale agente

[14]http://cermine.ceon.pl/cermine/task.html;jsessionid=99D5F9BDA2ED2DD6DCC3F9FF8EF5F9B1?task=3230711466372700550, http://cermine.ceon.pl/cermine/download.html?type=nlm&task=3230711466372700550.

automatico deve essere in grado di operare legittimamente, e di produrre decisioni e atti giuridicamente validi. Tuttavia, la capacità di istruire un agente automatico non può prescindere dalla costruzione di specifici modelli formali che consentano di utilizzare metadati ad alta veridicità e, conseguentemente, di ridurre il rischio di fondare le decisioni di un agente su informazioni non veritiere. In un mondo sempre più guidato dai dati, la modellizzazione logico-metaontologica dei metadati permette la costruzione di soluzioni per la sistematizzazione delle informazioni estratte e l'intersezione tra il diritto, la Data Science e l'intelligenza artificiale.

## Riferimenti bibliografici

[1] B. M. Abdel-Karim, N. Pfeuffer e O. Hinz. Machine learning in information systems – a bibliographic review and open research issues. *Electronic Markets*, 2021,1-28. 2021.

[2] C. Anderson. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired*. 2008. http://www.wired.com/2008/06/pb-theory/.

[3] D.M. Armstrong. *Truth and Truthmakers*. Cambridge University Press, Cambridge. 2004.

[4] F.F. Calemi. *Le radici dell'essere. Metafisica e metaontologia in David Malet Armstrong*. Armando Editore. Roma. 2013.

[5] S. Cuconato. Mondi di Wittgenstein. Metaontologia del 'Tractatus' e teoria dei 'truthmakers' di Armstrong. *Rivista Italiana di Filosofia Analitica junior*, 5:2, Special Issue: Metaphysics, 53-65, 2014.

[6] S. Cuconato. Epistemic logic for metadata modelling from scientific papers on COVID-19. *Science & Philosophy – Journal of Epistemology, Science and Philosophy*. 9 (2),83-96. 2021.

[7] A. d'Atri. *Ritorno alla Metafisica*. Bompiani, Milano, 2012.

[8] M. Frixione, S. Iaquinto e M. Vignolo. *Introduzione alle logiche modali*. Laterza, Roma-Bari, 2016.

[9] S. Galvan. *Logiche intensionali. Sistemi proposizionali di logica modale, deontica, epistemica*. Franco Angeli, Milano, 1991.

[10] E. Gettier. Is Justified True Belief Knowledge?. *Analysis*, 23,121-123, 1963.

[11] M. G. Lozano, J. Brynielsson, U. Franke, M. Rosell, E. Tjörnhammara, S. Varga e V. Vlassov. Veracity assessment of online data. *Decision Support Systems*, 2020.

[12] T. Lukoianova e V.L. Rubin. Veracity Roadmap: Is Big Data Objective, Truthful and Credible?, *Advances in Classification Research Online*, 4-15, 2014.

[13] J. Pomerantz. *Metadata*, MIT Press, Cambridge, MA, 2015.

[14] D. Snow. Adding a 4th V to BIG Data – Veracity, http://dsnowondb2.blogspot.se/ 2012/07/adding-4th-v-tobig-data-veracity.html, 2012.

[15] D. Tkaczyk, P. Szostek, P. Jan Dendek, M. Fedoryszak e Ł. Bolikowski. CERMINE — automatic extraction of metadata and references from scientific literature. Conference: 2014 11th IAPR International Workshop on Document Analysis Systems, 2014.

[16] D. Tkaczyk, P. Szostek, P. Jan Dendek, M. Fedoryszak e Ł. Bolikowski. CERMINE — automatic extraction of metadata and references from scientific literature. *International Journal on Document Analysis and Recognition (IJDAR)*, Springer, 2015.

[17] G. Turbanti. *Logica e mondi possibili*. Pisa university press, Pisa, 2020.

[18] H. van Ditmarsch, J. Halpern, W. van Der Hoek e B. Kooi. *Handbook of Epistemic Logic*. College Publications, 2015.

[19] P. van Inwagen. *Meta-Ontology*. Erkenntnis, 48, 223–250, 1998.

[20] T. Williamson. Gettier Cases in Epistemic Logic. *Inquiry: An Interdisciplinary Journal of Philosophy*, 56, 1-14, 2013.

**Science & Philosophy** Volume 10(2), 2022

# Scientific-Philosophical definition of life

Fröhlich, Klaus[*]

**Abstract**

There are about 100 different contradictory definitions of life. The definition of life based on symbiosis that is presented here differs fundamentally from them; it gives life a value. So this definition offers a basis for ethical and legal action e.g. in organ transplants. It is based on principles and is not an ad hoc model: Significant processes for life are basis for a theoretical concept. Quality criteria for definitions are employed to control the concept.
There is a graduation, not a clear division, between inanimate and animate. The graduation is based on the amount of symbiosis to be found. Life is based on symbiosis. The ideas of "ethics" and "reality" are considered in the context of this definition.
**Keywords:** Definition of Life; Biology and Ethic; Symbiotic.†.[‡]

[*]Klaus Fröhlich, ALS, Paracelsusweg 12, 30655 Hannover, Germany, froehlich@als-hannover.de

[‡]Received on July 07, 2022. Accepted on December 15, 2022. Published on December 15, 2022. DOI: 10.23756/sp.v10i1.801. ISSN 2282-7757; eISSN 2282-7765. 

# 1. Definition of the term "Life"

## Preliminary considerations

We use the term “life” frequently in everyday life. A definition as a common basis is required for scientific and philosophical work. This article draws on the author's earlier considerations, which have been integrated into a philosophical context and thus lead to fundamentally new aspects. This evolved text is distinguished by originality and novelty. (ref. Fröhlich, 2017)

Can you define life? The following "*objection ... has been advocated in particular by Francois Jacob (Jacob. 1982). Science has always progressed by focusing on questions of a limited amplitude - in contrast with, for instance, philosophy. By so doing scientists have been able to construct a solid form or knowledge. The question or life is too large to be a scientific one. ... But it does not mean that finding a definition of life is not a scientific objective."* (Morange, 2011)

Contradicting Jacobs theory that life cannot be defined scientifically the new definition of "life" based on symbiosis that is presented meets not only philosophical (comprehensive) but also scientific (exact and verifiable) requirements.

Sometimes the opinion is expressed that the term "life" cannot be defined. Then one should be so consistent not to use this term (in philosophical discussions). It should be borne in mind that definitions and guidelines are a requirement for legal certainty and can help with ethically difficult decisions.

Definitions determine future decisions. Therefore, the suitability is checked here with quality criteria for definitions. The definition of "life" is required for scientific work, ethical considerations in medicine and legal decisions. The following definition of the term "life" takes these areas into account.

Let us first narrow down the term: *"Living systems and Life are different concepts with different properties. ... Life is an attribute of living systems, or a theoretical concept about living systems in general.* (Poppa, 2009)

The starting point for the definition of life is a theoretical concept. First we consider which principles rule the basic processes of nature before taking scientific, philosophical and ethical aspects into account. Definitions based on principles have a higher quality than descriptive definitions, which may contain hidden assumptions. Definitions are part of models. Different models lead to different definitions. The symbiosis-based definition was developed within the framework of the scientific worldview.

## Cornerstones of the definition

Significant processes in nature:
A significant processes in nature is the formation of matter (information) and the laws of nature, the formation of chemical elements from elementary systems and the emergence of biological life.
Simple unicellular organisms develop to higher unicellular organisms, eukaryotes.
All of these developments have something in common, they are based on a cooperation for mutual benefit, which is called symbiosis.
The idea that life has something to do with symbioses is not new. *"Some scientists believe that life emerged as a symbiosis (mutualism) between independently developed mechanisms. ... (Schrödinger 1944, Eigen et. al. 1981)"* (Poppa, 2004)
Symbioses are scientifically important:
Simple systems come together to form a complex system with new properties (innovations) and symbiosis creates stability.
Symbiosis is logically - philosophically important:
A coalition goes hand in hand with a leap in development in knowledge. Control systems are formed (information storage, languages, laws, natural laws). Cooperation goes hand in hand with communication. Symbiosis enables experiences to be passed on.
Symbioses are of ethical - philosophical importance:
A prerequisite for the emergence of life is the consistent application of a principle that we call "love" on an emotional level. Evolutionary development is driven by the mechanism of science, which is based on the principle of truthfulness. Symbiosis defines meaning and values. (Symmetry break)
The symbiosis definition of the term "life" is based on the unity of mind and matter in the sense of systems theory. (ref. Bertalanffy, 2009)
Stable symbiosis follows mechanical and ethical rules: There is a parallelism between ethics and mechanics. The symbiosis definition of the term "Life" is based on Ethical, logical and mechanical principles. (Monism/ Elome - concept) (ref. Haeckel, 1899)

*Fröhlich, Klaus*

## Scientific-Philosophical definition of life

In the words of philosophy: "Basis of life is the principle of love. Life exists where this principle rules (symbiosis)."
In the words of biology: "Basis and indicator of life is the symbiosis."
This is a partnership in which each partner is active for the benefit of the other partner.
Colloquial: "What loves, that lives."
To sum it up
Science: Life is based on symbiosis. Philosophy: Life is based on the principle of love. Symbiosis put this principle into practice. Science: Tests control whether symbioses are present.
Why love and not justice?

Because love is biased: For the self, for the family, for life, for being, etc. (The lion feeds its child with the killed zebra.)

The basis of justice is, mathematically speaking, symmetry. A symmetry break takes place in a symbiosis, both physically and ethically. (Fröhlich, 2012)

Love defines values that must be preserved. Now the law, as a protector of values, is justified. In symbioses significance, values and meaning are defined. Symbioses, based on the principle of love, form the starting point for many processes in our world: law, stability, beauty, value, perception, knowledge, and the question of meaning.

The following example shows the importance of law as a regulator: We differentiate between wild animals, pets and farm animals as well as cute and nasty animals. The break in symmetry is also evident in our emotions. Animal welfare therefore requires ethical and legal considerations.

Effective values arise in a symbiosis, i.e. values correspond to a physical effect. Life, the stability of matter and the validity of natural laws are based on the principle of love. All stable elementary systems are subject to this principle. In a symbiotic view, the world consists of a network of relationships. In a transcendent sense, one could say that community spirit and not egoism forms the basis of life.

## Discussion

Do mechanical processes or ethical principles form the basis of our scientific worldview? This decision shapes our thinking and acting and is also reflected in the definition of the term life. In contrast to the technical definitions, the new definition assigns a value to life. This is important for treatment with living beings, especially in medicine.

The technical definition has evolved as follows: Out of the pre-scientific idea of what life is, a list of animate and uninhabited was created and searched for distinctive features.

In the words of Radu Poppa: *"A true life definition must exclude any material references and include all forms of life, (or things that may become alive)"* (Poppa, 2009)

This approach makes sense, but does not lead to the desired result, because the (arbitrary) division into animate and inanimate is not appropriate.

The definition of the term "Life" based on symbiosis expands the scope: Life is the basic principle of nature. The opposing pair animate / inanimate is replaced by an intermittent transition between the individual forms of life, which depends on the degree of symbioses. The diversity of living beings is greater than in the traditional definition. It ranges from the simple to the complex and holds, for example, culture as a form of life. The complexity depends on the degree of symbiosis.

The chemical elements and elementary particles have also characteristics of life. Therefore, the biological life has arisen not from inanimate, but from living matter.

Unicellular organisms differ so strongly from the chemical processes that one is inclined to see a contrast (animated / unanimated) that does not exist. However, the fine-grained transition from the chemical plane to the biological plane is difficult to recognize because the processes in biological forms of life (e.g. in unicellular organism) are based on a multitude of symbioses.

Biologists who produce unicellular organism from chemical substances, only develop existing creatures. We cannot create life, but we can develop it further and bring it to a higher level.

## Varied forms of life

The division into alive and lifeless ignores the principle of the unity of nature. The definition presented here replaces this opposition with a finely graduated transition between the individual forms of life, depending on the degree of symbioses.

Symbioses can be found between the creatures of a biotope (climax), in cultures, in social cohabitation (e.g. bee-keeping), in partnerships (e.g. lichen), within one species (e.g. apoptosis in yeast), in organisms (metazoan, organ formation), in eukaryotes (endosymbiosis), for simple unicellular organism (complex control systems), in non-cellular forms of life (autocatalysis / hypercycles), in the atoms of chemistry (in shell and core), for the elementary systems (e.g., photon, electron, proton)

Living beings are: elementary systems - atoms - molecules - cells - organs - creatures - cultures. In simple unicellular organisms there are very large numbers of symbiotic processes, in chemical reactions few.

In the Elome concept, ethics, logic and mechanics form the basis of the natural and human sciences and their terminology. Life and reality are one. The full vitality of nature cannot be equated with the traditional animistic conceptions. ("The raincloud cries because it is sad.") This model requires no humanization and no mystification.

The following processes also fit into the symbiosis concept:
From a fertilized ovum a person develops.
People consist of many living cells. Cells grow and die, but people live on.
The brain cells dies, the person is declared "dead".
In organ transplantation, the donor dies, but his cells and organs continue to live. According to the technical definition, they are referred to as *"cellularly active organs"* for legal reasons. (Vrselja, et al. 2019) The step-like structure of symbioses includes technology and ethics. This enables a value-based consideration.

For medical professionals, conscious ethical action is of significant importance. Technical definitions are useful tools; principle-based definitions also offer approaches for ethical decisions. Of conscious importance here is the principle of love contained in the symbiosis model.

## Further definitions of the term "life"

How does the symbiosis-based definition fit with common definitions? Radu Poppa compiled a list of 40 definitions of life in 2002.

According to Addy Pross, these definitions can be divided into four groups: Information (software), Infrastructure (hardware / energy), Enumeration of common properties of living things and Thermodynamics.

Addy Pross gives the following examples:

*"Information (software)*

*´Life is defined as a material system that can collect, store, process and use information to organize its activities.´ (Dyson, 1999)*

*Infrastructure (hardware / energy)*

*´Life is defined as a system of nucleic acid and protein polymerases with a constant supply of monomers, energy and protection.´ (Kunin, 2000)*

*List of common properties of living things*

*´Life is defined as a system that is capable of 1. self-organization;*
*2. self-replication; 3. evolution through mutation; 4. metabolism; and*
*5. an encapsulation based on it.´ (Arrhenius, 2002)*

*Thermodynamics*

*´Life is simply a special state of organized instability.´ (Hennet, 2002)"* (Pross, 2012)

A question arises with the definitions. Why can that which is defined be called "life". Addy Pross: *"The definitions above, all relatively new and all revealing in their own way, show almost no overlap. If not all of the definitions had started with the words" life is ... ",* we would be tempted to believe that there are definitions were about completely different concepts." (Pross, 2012)

## Summary

The functioning of organisms is described by various disciplines and each is referred to as a definition. This approach is controversial in the literature. The *"spiritual aspects of life"* are not taken into account. Erwin Schrödinger saw this as an essential problem in his own definition. (Schrödinger, 1944)

The symbiosis definition of the term "life" is another type of definition. It is principle-based. It gives life a value. The values in the Elome concept are effective values. They correspond to a physical effect. The symbiosis definition of the term "life" forms the basis for scientific work, ethical considerations in medicine and legal decisions. The symbiosis model includes the differentiated special models. The symbiosis-based definition is compatible with the technical definitions and goes beyond.

In the Elome concept life and reality are one. As a philosophical basis, this model requires a definition in which creatures receive the model-specific predicate "real".

| Comparison of definitions with quality criteria for definitions | |
|---|---|
| TECHNICAL DEFINITION | PRINCIPLE OF LOVE (SYMBIOSIS) |
| **Is there a definition?** | |
| Living beings have a boundary, a metabolism, grow and reproduce. They can record, save, change, and send information. They obtain their inner order by responding appropriately to internal and external influences. *) | The basis and indicator of life is the symbiosis.<br>Colloquial: "What loves, that lives." |
| **Is it intuitive?** | |
| This definition is comprehensible but not intuitive. | This definition is intuitive.<br>(Love is only observed in creatures!) |
| **Contains the essential idea it is based on?** | |
| Too technical. | Yes. |
| **Ockham's razor** | |
| Many criteria to be met. | A single requirement. |
| **Base** | |
| Observed properties | Principles. |
| **Internal consistency** | |
| Yes | Yes |

| TECHNICAL DEFINITION | PRINCIPLE OF LOVE (SYMBIOSIS) |
|---|---|
| **Consensus** | |
| No. There are many different variations. | For discussion. |
| **Extent "fertility" (How big is its range?)** | |
| Low.<br>The definition describes the earthly life known to us. | Far. Life as a basic principle of nature: It covers all known biological forms of life, but also accepts, for example, culture or elementary systems as forms of life. |
| **Comprehension (Russel's hen)** | |
| Yes, this definition can explain most properties of living things. | Yes, because one can show that love is a necessary condition for life, and can explain how symbioses create life and create higher forms of life. |
| **Accuracy et equality before the law: Copernican principle** | |
| It leaves room for interpretations. What if all the criteria are not met? | Only one criterion has to be tested. There are no exceptions to these rules. |
| **Integration into science** | |
| No, since Wöhler's urea synthesis and Darwin's theory of evolution the subdivision into animated and uninhabited no longer corresponds with scientific knowledge. | Yes. Follows the scientific method. Corresponds to the principle of unity of nature. (Depending on the number of symbioses, a fine-tuned classification of the different forms of life is possible.) |
| **Integration into philosophy** | |
| No. | Yes. In symbioses comprehension, values and meaning define each other: Life is precious. Basis for legal considerations. |
| **Practicable (including probation)** | |
| Yes, with the mentioned restrictions. | Yes, symbioses are common to science. |
| **Importance** | |
| Basis for a technical model of biological life. | Basis for a scientific - philosophical model of life. |
| The emotional level has to be supplemented. The definition based on symbiosis does not only correspond to what we call "life" in everyday life, but it also corresponds to what we perceive as "life".<br>*) Compilation of the author. (Fröhlich, 2017) | |

Table 1: Comparison of definitions of life

# 2. Philosophical Context

### Preliminary considerations: Life and Reality

In the Elome concept life and reality are one. As a philosophical basis, this model requires a definition in which creatures receive the model-specific predicate "real".

*"Is is."* (to Parmenides, 510-440 BC) For idealism, what is immutable and eternal is real. Every form of change is "opinion", "illusion", "deception". According to this concept, all the processes of our everyday life are not real and ultimately meaningless.

The idealistic definition is deeply rooted in religion and philosophy. Think, for example, of eternal and immutable souls, or the concept of indestructible atoms. (Fröhlich 2017)

The models of physics take the opposite position: the principle of action and reaction apply. Something unchanging, according to the state of knowledge, has no effect and cannot be perceived. It has no significance either for physics or for our everyday life. The scientific understanding of reality and the concept of reality of idealism are diametrically opposed.

*"Everything flows."* (to Heraclitus Ephesus, 520-460 BC) There is no consensus. *"I know that I know nothing."* (to Socrates, 463-399 BC) In other words, we don't know the basic properties of nature, we suppose we know them. The Concept of Elome supposes: Effective values are real, living beings are real and the scientific method produces the best possible explanatory models. (ref. Genz, 2002) The new biological definition expands the existing definition, the new philosophical definition stands in contrast to the traditional definition ("*paradigm shift*"). (ref. Kuhn, 1967)

| Comparison of definitions with quality criteria for definitions | |
|---|---|
| IDEALISM (PERMANENCE) | PRINCIPLE OF LOVE (SYMBIOSIS) |
| **Is there a definition?** | |
| "Is is."<br>With other words: "What is immutable and eternal is real." *) | "Real is where values correspond to a physical effect (symbiosis)."<br>Colloquial: "What loves is real." |
| **Is it intuitive?** | |
| This definition is intuitive. | This definition is intuitive. |
| **Contains the essential idea it is based on?** | |
| *What is essential?* Permanence! | *What is essential?* Values! |
| **Ockham's razor** | |
| A single requirement. (Equalization) | A single requirement. (Symbiosis) |
| **Internal consistency** | |
| Yes, if no conclusions are drawn.<br>No, when using logic. | Yes. |
| **Comprehension (Russel's hen)** | |
| Yes. - Within the framework of the specific philosophical system of idealism.<br>No. - Unchanging things have no effect on the models of physics.<br>No. According to this definition, stones, trees, giraffes, planets, atoms and our fellow human beings are not real. | Yes - Within the framework of everyday life, science and logic (philosophy).<br>Yes, because one can explain how symbiosis manages to create and maintain matter (copyable information).<br>Yes. - According to this definition, stones, trees, giraffes, planets, atoms and our fellow human beings are real, are stable through symbiosis. |

| IDEALISM (PERMANENCE) | PRINCIPLE OF LOVE (SYMBIOSIS) |
|---|---|
| **Extent "fertility" How big is its range?** | |
| Inhibiting!<br>Rejection of logic, problems with veracity. Everyday problems, ethics and life are meaningless in this model. Perceptions are not possible. | Far. Reality as the basic principle of nature: Values and meaning are defined in symbioses.<br>Basis for the pursuit of harmony, ethics and logic, law and science.<br>Helpful to meet the demands of everyday life. Consciousness and perception are taken into account. |
| **Accuracy et al equality before the law: Copernican principle** | |
| No. Rules for everyday life and ethics can not be derived.<br>Rejection of science. | Yes. To test whether an object can be assigned the model-related predicate "real", check whether a symbiosis exists.<br>Symbioses are common to science. They can be described with mathematical models.<br>Depending on the number of symbioses, a step-by-step graduation of different degrees of being real exist. |
| **Integration into science; Integration into philosophy** | |
| Basis for many philosophies and religions. Devaluation of everyday life. Was often abused. | Basis for a philosophical - ethical - scientific - practicable model of the world. |
| **Consensus** | |
| No! Paradigm shift. | |
| The emotional level has to be supplemented. The symbiosis-based definition does not only correspond to what we call "real" in everyday life, but it also corresponds to what we perceive as "real".<br>*) Compilation of the author. (Fröhlich, 2017) | |

Table 2: Comparison of definitions of real

## Scientific-Philosophical definition of the term "real"

Based on observations in practice, theoretical considerations are made: In a dynamic world, stable objects are created in the context of symbioses. Symbioses contain values and have effective rules. These objects are given the model-specific predicate "real". The basis and characteristic of reality is the symbiosis. (Symbiosis model).

The symbiosis definition of "real" is based on beliefs and, unlike the idealistic definition, contains no claim to truth. (Good models are not "true", but helpful.)

The symbiosis-based definition presented here explains the roots of permanence. It is in harmony with science and builds upon philosophical principles: "Real is where values correspond to a physical effect (symbiosis)."

Being and not being are not opposites in the symbiosis model. Between them there is a smooth transition. (That is why one can die.) In quantum physics, "possible being" forms a level between "being" and "not being". The "degree of reality" is dependent on the number and strength of the symbioses and indirectly also on the number of perceptions and possibilities of perception. On the level of elementary systems (quantum physics), it is low, higher in individual cells, even higher in humans. Cultures possess more symbioses and more possibilities and thus have a higher degree of reality than a single person.

## Further definitions of the term "real"

The following definitions can be divided into three groups: Model (Thinking); Perception; Model and Perception.

Model (Thinking):

If two terms in a statement have the same meaning (synonyms), one speaks of tautology. Parmenides of Elea believed that such statements were absolutely true:

*"Is is."  /  "What is, that is."  /  "Being is."* (to Parmenides, 510-440 BC)

The meaning of the two words "is" does not have to be the same (seemingly a tautology). Depending on the context into which a sentence is integrated, the meaning may be different. This must be checked individually for each context. Fist rule: Since tautologies are insignificant, the two seemingly synonymous terms differ from one another the more statements they can derive. The sentence of Parmenides has only the value of a presumption. (Fröhlich, 2017)

(Parmenides and the Eleatic school assigns the following properties to the term "is": a whole, one, similar, coherent, uniform, unlimited, indivisible, unchangeable, indestructible, undeveloped, imperishable, timeless, eternal.)

According to Parmenides, only thinking leads to truth. (Isn't thinking also something dynamic?)

The symbiosis-based definition connects the spiritual with the material. Attributes are assigned based on ethical considerations. To assign the object "real" to an object, it must be observed. According to the rules of system theory, it is checked whether symbioses are present.

Perception:

Another definition of reality is based on self-awareness: *"I think, therefore I am."* (Descartes, 1642) According to the definition of Descartes, only the reality of the observer is assured.

For Berkeley, observers and perceived are real and immaterial:

*"esse est percipere"– "to be is to perceive"* and

*"esse est percipi" – "to be is to be perceived"* (George Berkeley, 1710)

Model and Perception (Effective Values):

For Berkeley, observers and perceived must have the same properties. This is controlled in the symbiosis model. There, the definition of reality is based on abstract principles. In the symbiosis model, the observer and the perceived have comparable ethical, logical and mechanical properties.

In addition, abstract considerations and observations incorporate the definition into the scientific worldview: Symbioses can be observed and described with models, they go hand in hand with interactions, communication and perception. In the symbiosis model, perception is neither an illusion nor an image; perception has the properties of a model with a model-specific language.

Observable objects form the starting point for scientific models. Within the scope of these models, these objects are present or, in other words, "real".

Model and Perception (Prediction):

Einstein, Podolsky, and Rosen (EPR) have set a definition of "real" within the framework of the discussions on quantum physics:

*"A physical quantity whose value is predictable with certainty without disturbing the system on which it is measured, is an element of physical reality."* (Einstein, A. Podolsky, B. Rosen, N., 1935)

According to this definition, (only) predictable events are real. This definition has given impulses to create Bell's inequality.

I leave the reader to examine the quality of these definitions on the basis of the criteria described above.

## Clarification of the term "ethics" of life

The meaning of ethics in this concept has to be understood.

The laws of mechanics describe the immediate and imperative reaction to an action. The laws of mechanics are not sufficient for a scientific description of our world.

In order to control the processes in complex systems, other rules are required. Ethics provide a basis for formulating these rules. Their effect can only be predicted as "empirically" or "usually" (statistically). However, the recommendations for ethical action do not have to be followed.

The task of ethics is to protect life. Recommendations for action can be derived from their principles in order to stabilize symbiotic systems. (As Immanuel Kant, for example, called for with the *"categorical imperative".)* (Kant, 1781)

Ethics, logic and mechanics (Elome concept) contribute to the understanding of our world. The ethics described here are relative ethics – there are no absolute truth, no absolute ethics. The supposed possession of absolute truth leads to untruth. The supposed possession of absolute ethics leads to immoral acts. Statements and actions by religious and political fanatics clearly show this. Absolute ethics cannot be derived from logic.

See: George E. Moor: *Naturalistic Fallacy* (ref. Moor, 1903)

See: David Hume: *To Be - Should - Fallacy* (ref. Hume, 1734)

The relative ethics presented here have a defined goal. It is the maintenance of a symbiotic system / the maintenance of life. Therefore, ethical principles can (and should) be logically justified. There is a parallelism of ethics and mechanics in the Elome concept: stable systems obey the laws of ethics and mechanics.

### Life and evolution

Does an ethical definition of the term life fit Darwin's theory of evolution?

- Many people cannot understand how *"Struggle for Existence"* could produce butterflies and flower meadows. (ref. Darwin, 1895)

- For many people it is incomprehensible how a "*pitiless indifferent universe"* could produce values and emotions. (ref. Dawkins, 1995)

In contrast the Elome concept: Darwin's theory of evolution describes a scientific process of knowledge and innovation that is associated with values. (Freedom - Diversity - Truthfulness - Love - Justice) (ref. Fröhlich, 2017)

## Evolution and the principle of science

The principle of science is truthfulness applied in practice. To successfully implement the principle in practice, scientists have developed a number of methods over the centuries. These methods are based on the "mechanism of science". This mechanism is named differently for the different applications:

In the words of science: Setting up a model - Checking the model - Publishing the model.

In everyday language: Freedom of choice - Proven in practice - Sharing of experiences.

In the words of philosophy: Freedom - Truthfulness - Love In words of biology: Mutation - Selection - Multiplication.

Interestingly, the evolution mechanism is a variant of the mechanisms of science: the mutation corresponds to the development of a model and the selection corresponds to the testing of the model. The experience gained is passed on during propagation or through publications. There is a selection in each of the three steps. Evolution accomplishes something like science: models (knowledge) and technology (innovations). Natural laws, matter and life arise in an evolutionary process. (ref. Ditfurth, 1976)

The objects that arise in this process have spiritual and material properties. Perception also arises evolutionarily. Perception is neither an illusion nor an image; perception has the properties of a model.

Quantum physics describes the emergence and decay of the elements of nature. The mechanism described can also be found in quantum physics.

In words of quantum physics: Chance - Effect - Information (Matter)

Matter is evolutionarily acquired and preserved knowledge that is stored in an effective form (technology). Matter is something "spiritual" that has an effect.

The mechanism of science is important in the emergence and maintenance of natural laws and matter but also in emergence of biological life and the development of living things. The mechanism of science is involved in the development of our perception, our feeling and our thinking, languages, culture.

Living beings, laws of nature and matter arise in evolutionary processes according to the mechanism of science. Chance is insignificant. Here it has as little meaning as chance in the test series of the scientists. In the long run, principles prevail. The laws of nature are subject to similar principles as philosophy.

The basis of physics is formed by ethics and philosophy. (Freedom - Diversity - Truthfulness - Love -Justice - etc.) Science does not lead to an emotionless, meaningless worldview.

**Statement**

- Scientists combine science with philosophy and ethics.

or

- Scientists leave philosophy and ethics to the Others.

# References

Arrhenius, Gustav, 2002 "Short definitions of life." In Fundamentals of Life, edited by G. Palyi, C. Zucchi, and L. Caglioti, Elsevier, New York, pp 17-18; in: Pross, Addy, 2012, What is Life? How Chemistry becomes Biology: 40, Oxford Landmark Science

Berkeley, George, 1710, A Treatise Concerning the Principles of Human Knowledge new edition: Forgotton Books; Turbane (1957)

Bertalanffy, Ludwig von, 1949, "General System Theory." In: Biologia Generalis. 1/1949: 114–129

Darwin, Charles, 1859, The Origin of Species, London:, retrieved by John Murray 22 February 2009

Dawkins, Richard, 1995, River out of Eden, Basic Books, New York

Descartes, René, 1641, Meditationes de prima philosophia translated by Andreas Schmidt, Reclam Verlag, Dietzingen, 2020

Ditfurth, Hoimar von, 1976, Der Geist fiel nicht vom Himmel, Hamburg, Hoffmann und Campe Verlag

Dyson, Freeman, 1999, Origins of Life, Second Edition, Cambridge University Press; in: Pross, Addy, 2012, What is Life? How Chemistry becomes Biology: 40, Oxford Landmark Science

Einstein, A., Podolsky, B. and Rosen, N., 1935, "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?" In: Physical Review, 47: 777-780.

Fröhlich, Klaus, 2017, Definition of "Life" Based On Symbiosis Model: 4-9, Raleigh

Fröhlich, Klaus, 2017a, Anregungen zum Nachdenken über Wissenschaft, Philosophie, Religion und Recht, Raleigh

Genz, Henning, 2002, Wie die Naturgesetze Wirklichkeit schaffen. München, Wien

Haecke, Ernst, 1899, Die Welträtsel. Gemeinverständliche Studien über Monistische Philosophie; new edition Verlag Hofenberg; Bonn; 2016

Heraclitus Ephesus, 520-460 BC, as quoted by Plato in Cratylus; standardebooks.org/

Hennet, Remy, 2002, Short definitions of life. In: Fundamentals of Life, Elsevier, New York: 3; in: Pross, Addy, 2012, What is Life? How Chemistry becomes Biology: 40, Oxford Landmark Science

Hume, David, 1734, A Treatise of Human Nature; Google Books

Jacob, Francois, 1982, The Logic of Life: A History of Heredity; in: Morange, Michel, 2011, "Problems raised by a definition of life" In: Origin and Evolution of Life - an Astrobiology Perspective; Cambridge: 5-6

Kuhn: Thomas Samuel, 1967, Die Struktur wissenschaftlicher Revolutionen. Frankfurt a. M.

Kant, Immanuel, 1781, Kant, collected writings (1900). Prussian Academy of Sciences (New Edition)

Kunin, Viktor, 2000, "A system of two polymerases – A model for the origin of life." Orig. Life Evol. Biosphere 30: 459–466, in: Pross, Addy, 2012, What is Life? How Chemistry becomes Biology: 40, Oxford Landmark Science

Moor, George E. 1903, Principica Ethica; New Edition CreateSpace Independent Publishing Platform (2017)

Morange, Michel, 2011, "Problems raised by a definition of life" In: Origin and Evolution of Life - an Astrobiology Perspective; Cambridge: 5-6

Parmenides, 510-440 BC, On Nature; new edition on the order of nature; Aurea Vidya Collection

Pawlowski, Tadeusz, 1980, Begriffsbildung und Definition, Berlin, New York

Poppa, Radu, 2009, The Complexity-Independence of the Origin of Life: 5; pdxscholar.library.pdx.edu

Poppa, Radu, 2004, Between Necessity and Probability: Searching for the Definition and Origin of Life: 141; Springer Verlag

Pross, Addy, 2012, What is Life? How Chemistry becomes Biology: 40, Oxford Landmark Science; Oxford

Schrödinger, Erwin, 1944, What is Life? The Physical Aspect of the Living Cell, Cambridge University Press; Cambridge

Socrates, 463-399 BC, as quoted by Platon in Apology of Socrates, 21d; new edition Friedrich Schleiermachen Taschebuch, 2004

Vrselja, Z., Daniele, S.G., Silbereis, J. et al., 2019, "Restoration of brain circulation and cellular functions hours post-mortem. "In: Nature 568, 336–343

# Science & Philosophy

## Volume n. 10, Issue n. 2, 2022

*Contents*



Published by APAV (Accademia Piceno-Aprutina dei Velati in Teramo)