ZFN

Philosophical Problems in Science

ZFN

Zagadnienia Filozoficzne w Nauce

Is logic a physical variable?

Philosophical Problems in Science

Zagadnienia Filozoficzne w Nauce

Philosophical Problems in Science

Zagadnienia Filozoficzne w Nauce

LXIX ■ 2020

Book reviews

Editorial

"Is logic a physical variable?" Introduction to the Special Issue

Michał Eckstein, Bartłomiej Skowron

"Is logic a physical variable?" This thought-provoking question was put forward by Michael Heller during the public lecture "Category Theory and Mathematical Structures of the Universe" delivered on 30th March 2017 at the National Quantum Information Center in Sopot. It touches upon the intimate relationship between the foundations of physics, mathematics and philosophy. To address this question one needs a conceptual framework, which is on the one hand rigorous and, on the other hand capacious enough to grasp the diversity of modern theoretical physics. Category theory is here a natural choice. It is not only an independent, well-developed and very advanced mathematical theory, but also a holistic, process-oriented way of thinking.

Michael Heller's inspiring question provided the first impulse to organize a meeting of mathematicians, physicists and philosophers interested in exploring the applications of category theory in their research. The actual conference, co-organized by the Copernicus Center for Interdisciplinary Studies and the International Center for Formal Ontology took place on 7-9 November 2019 in Kraków, Poland. The Conference was the 23rd edition of Kraków Methodological Conferences, held regularly since 1992, devoted to fundamental problems at the junction between philosophy and sciences. The conference proceedings are collected in this volume.

The papers in the volume have been authored by philosophers, physicists and mathematicians. The common denominator uniting

these works is a *categorical cognitive perspective*, i.e. category theory, which is gaining more and more interest in Poland, is the central axis of the volume.

Category theory was developed in the early 1940s. It resulted from the combination of topological talent of the Polish-American mathematician Samuel Eilenberg and the algebraic talent of the American mathematician Saunders Mac Lane. In the sixties, the theory developed significantly, becoming a fully-fledged and independent branch of mathematics. The development of category theory often took place in directions previously unanticipated by its creators. In particular, the deep insights by William Lawvere's have shown that category theory has enough potential to even shake the very foundations of mathematics. Today, category theory is a standard tool not only for mathematicians, but also theoretical physicists and philosophers. Category theory is full of unexpected relationships, deep analogies and hidden connections between seemingly unrelated structures, arousing interest among many scientists and scholars, thus reaching far beyond pure mathematics.

Category theory implies a formal ontology that emphasizes the relationship between objects and thus diminishes the role of the objects themselves. This aspect attracts many scholars, but for many it also makes it difficult to understand what it is all about. Thus category theory requires a certain effort to change the object-oriented cognitive attitudes. Below we briefly present the content of the all articles published in the volume to help the Reader to discover the various aspects of category theory discussed in this volume.

Colin McLarty in the paper *Mathematics as a love of wisdom: Saunders Mac Lane as philosopher* points out the similarities between Aristotle's and Saunders Mac Lane's thoughts. The author describes this unexpected analogy in a very interesting way. Mac Lane, together

with Eilenberg, have established category theory. They had a great influence on the development of mathematics in the 20th century. From the beginning of his career Mac Lane was interested in philosophy, nevertheless, not as an academic philosopher, but just as a working mathematician. McLarty points out that both Aristotle and Mac Lane treated "knowledge" as the "knowledge of reasons". For Mac Lane, mathematical understanding is not only to know the proof of a given theorem, but also to know the *reasons* for that theorem.

According to George Cantor "The very essence of mathematics lies precisely in its freedom". We can understand this statement in such a way that a mathematician can introduce into mathematics the concepts (s)he wants to introduce, there are no restrictions in this process, except the logical constraints. This is very attractive and this is what makes us want to practice pure mathematics! Zbigniew Semadeni in his paper *Creating new concepts in mathematics: freedom and limitations. The case of Category Theory* ponders the question: "What influenced the emergence of category theory?", from the perspective of Platonizing constructivism. The author claims that despite the fact that the emergence of CT was a transgression (i.e. *crossing of a previously non-traversable limit of mathematical knowledge*), the development of the notion of function since the beginning of the 19th century was one of the important factors in the emergence of CT. The author also analyzes the origin of the term *functor*, claiming that it can be found already in the works of Tarski and Kotarbiński.

Structuralism in the philosophy of mathematics is a position, painting it with a broad brush, according to which mathematics concerns certain abstract structures. It seems natural to ask what logic is behind a given version of structuralism? Logic can influence the properties of given structures, e.g. metalogical properties. Is it first-order logic, higher-order logic, or maybe modal logic, or even other? Jean-

Pierre Marquis in his paper *Abstract logical structuralism* insightfully points out that, first of all, structuralism is also about logic itself: that is, structuralism should also be interested in pondering what kind of mathematical structure hides behind logic itself and what are the relations of that structure to other mathematical structures. Secondly, and more importantly, the author indicates and convincingly justifies, giving many examples of mathematical structures related to logic, that this is in fact a categorical logic. In the words of the author: "Indeed, this is what categorical logic is all about: it reveals the abstract mathematical structures of logic and it relates them to other abstract mathematical structures, revealing yet other structural features".

The theory of toposes is typically linked with intuitionistic logic, although the actual connection is more subtle. In 1995 Chris Mortensen formulated the concept of a co-topos, based on closed rather than open sets, which is related to a paraconsistent logic. In his article *On the validity of the definition of a compelement-classifier* Mariusz Stopa has a critical take on co-toposes. He analyses the definition of the "true" arrow and argues that the definition cannot be interpreted arbitrarily. As Stopa acknowledges, his analysis hinges upon the equivalence of two definitions of toposes provided by Mac Lane and Moerdijk in 1994.

Toposes are also at the core of the article *No-signaling in topos formulation and a common ontological basis for classical and non-classical physical theories* by Marek Kuś. The author provides a concise introduction to Bell-like inequalities, which triggered the foundational tests of quantum mechanics against classical, hidden variables, theories. His analysis extends also to 'post-quantum' scenarios involving the so-called no-signalling boxes introduced by Popescu and Rohrlich in 1994. The article explains the internal logic of classical,

quantum and no-signalling theories. Kuś also argues that, with the help of toposes, one can unveil a common ontological basis for all three theories.

In 1935 Niels Bohr, in response to the now famous paradox unravelled by Einstein, Podolsky and Rosen, has written: "The apparent contradiction in this fact [i.e. the impossibility of attaching definite values to both of two canonically conjugated variables] discloses only an essential inadequacy of the customary viewpoint of natural philosophy for a rational account of physical phenomena of the type with which we are concerned in quantum mechanics." In his article *Quantum contextuality as a topological property, and the ontology of potentiality* Marek Woszczek takes this viewpoint and argues that quantum theory challenges our basic intuitions on "physical definitness" and "objective reality" and calls for an "ontology of potentiality". His pivotal argument is based on the quantum contextuality, which could be seen—in a category theory spirit—as a topological property.

Quantum geometry, logic and probability is the title of the article by Shahn Majid. The author provides an inviting introduction to his original formalism of discrete quantum geometry based on graphs. The main result is a surprising connection between Markov processes and a generalized form of (noncommutative) Riemannian geometry. Then, Majid introduces a 'discrete Schrödinger process' and shows that it induces a Markov-like correction to the standard continuity equation in quantum mechanics. In the last part, the author argues in favour of a (quantum) geometric viewpoint on de Morgan duality and discusses its consequences and potential application in the quest of constructing a consistent theory of quantum gravity.

In today's world filled with virtual objects, information plays an incomparably greater role than before, when there were no computers and so many objects generated by computers. In the text *Information*

and physics Radosław Kycia and Agnieszka Niemczynowcz present, in a way that is accessible to non-specialists, studies that shed light on the relationship between the abstract image of information on the one hand and its physical (or "real") representative on the other. The authors present a solution to the Maxwell paradox using Landauer's principle. This principle links in a non-obvious manner the computational processes to the production of measurable physical consequences, i.e. to the generation of heat. In their presentation the authors refer to their research on the general method of building thermodynamic analogues of computer memory. In this method, the authors find Galois connection, which makes category theory play a key role in this research.

The mind–body dichotomy is one of the most profound philosophical problems. Although the physical structure of the brain, along with the basic mechanisms of its functioning, is systematically unravelled by modern neurobiology, the problem of how does the brain (or, rather, mind) ascribe meaning to the processed data is unsolved. Steve Awodey and Michael Heller in their though-provoking article *The homunculus brain and categorical logic* take inspiration from the intricate duality between syntax and semantics, well established within category theory. They propose a toy-model—a "homunculus brain"—with neurons modeled by categories and axons by functors. The homunculus mind, is, on the other hand, modeled by a category of theories. The authors introduce the BRAIN and MIND categories and show—using the results from categorical logic, including the syntax-semantics relationship—how the structure enhanced by adjoint functors "Lang" and "Syn" enable the creation of meanings. As the authors summarize in a typical categorical attitude of finding un-

obvious, deep and far-reaching connections: "The categories BRAIN and MIND interact with each other with their entire structures and, at the same time, these very structures are shaped by this interaction."

As the guest editors of this volume we would like to express our sincere gratitude to all of the authors for their valuable contributions, to the anonymous reviewers for detailed and instructive reviews, and to the editor-in-chief of *Philosophical Problems in Science*, Paweł Polak, for his kind support of this initiative. We are also very grateful to the editorial secretary Piotr Urbańczyk for his commitment at almost every stage of the preparation of this volume and to Roman Krzanowski for the proofreading.

Articles

Mathematics as a love of wisdom: Saunders Mac Lane as philosopher

Colin McLarty

Department of Philosophy, Case Western Reserve University

Abstract

This note describes Saunders Mac Lane as a philosopher, and indeed as a paragon naturalist philosopher. He approaches philosophy as a mathematician. But, more than that, he learned philosophy from David Hilbert's lectures on it, and by discussing it with Hermann Weyl, as much as he did by studying it with the mathematically informed Göttingen Philosophy professor Moritz Geiger.

Keywords

naturalism, philosophy, mathematics, Aristotle.

> Go, read, and disagree for yourself.
> (Mac Lane, 1946, p.390)

This note describes Saunders Mac Lane as a philosopher, and indeed as a paragon naturalist philosopher. Obviously he approaches questions in philosophy the way a mathematician would. He is one. But, more deeply, he learned philosophy by attending David Hilbert's public lectures on it, and by discussing it with Hermann Weyl, as much as he did by studying for a qualifying exam on it with the mathematically informed Göttingen Philosophy professor

Moritz Geiger (McLarty, 2007; 2020).[1] Before comparing Mac Lane
to Penelope Maddy's created naturalist, the Second Philosopher, we
relate him as a philosopher to Aristotle.

This is not to disagree with Skowron's view of Mac Lane as a
platonist in ontology (Skowron, manuscript). This is about Mac Lane
on scientific and philosophic method. He understood *philosophy* very
much as Aristotle did: Philosophy is the love of wisdom, and every
science pursues wisdom and not mere facts. I do not claim Mac Lane
got the ideas from Aristotle. So far as I know, Mac Lane himself had
no interest in or opinion of Aristotle, though his teachers Weyl and
Geiger certainly talked of Aristotle.[2]

1. Aristotle on wisdom and science

> In general the sign of knowing or not knowing is the ability to
> teach, so we hold that art rather than experience is scientific
> knowledge; some can teach, some cannot. Further, the senses
> are not taken to be wisdom. They are indeed the authority for
> acquaintance with all individual things, but they do not tell the
> why of anything, for example why fire is hot, but only that it
> is hot.
> It is generally assumed that what is called wisdom is concerned
> with the primary causes and principles, so that, as has been
> already stated, those who have experience are held to be wiser

---

[1] Mac Lane's close contact with Paul Bernays in Göttingen deserves more attention.
But my research so far has not identified strong philosophic influence from Bernays.
See Section 3.

[2] Browder and Mac Lane (1978) give a beautiful survey of mathematics which mentions Plato and Aristotle on ontology. However, the discussion of Plato and Aristotle summarizes the longer discussion by Browder (1976).

> than those who merely have any kind of sensation, the artisan
> than those of experience, the craft master than the artisan.
> (Aristotle, Metaphysics 981f.)[3]

Here Aristotle says scientific knowledge is better than acquaintance, or even experience, in two related ways:

1. scientific knowledge can be taught, and
2. scientific knowledge gives the why of things.

He says wisdom deals with primary causes and principles. And those are his touchstone for scientific knowledge: "We do not know or have scientific knowledge of objects of any methodical inquiry, in a subject that has principles, causes, or elements, until we are acquainted with those and reach the simplest elements" (Physics 184a).

Throughout his career Mac Lane tied pedagogy to research and research to craft. Among many examples see his early notes on presenting mathematical logic to university students (Mac Lane, 1939), and his impassioned argument that theoretical education prepared men and women well to do the applied mathematics which he supervised in World War II (Mac Lane, 1989; 1997).

Mac Lane also insists mathematical understanding includes knowing the reasons for a given theorem. Some proofs of a theorem may reveal the reason, while other technically sufficient proofs will not reveal the reason. In his book for philosophers, Mac Lane sketches proofs for major theorems from many subjects, like linear algebra, or complex analysis. He often describes several alternative proofs for a single theorem and then eventually singles out one as giving the real reason. That book is (Mac Lane, 1986) and some examples are on pages 145, 189, 427, 455.

---

[3] Translations of Aristotle here use "know" for *oida*, "scientific knowledge" for *episteme*, and "acquaintance" for *gnosis*. Of course "wisdom" is *sophia*.

For me, though, the deepest connection between Aristotle's and Mac Lane's loves of wisdom is how they say we gain this scientific knowledge. Both believe in foundations, or "first principles" if you prefer, but neither believes we start with those. Aristotle's theoretical demand of philosophy was a theoretical and practical demand in mathematics for Mac Lane:

> The natural way of [getting scientific knowledge] is to start from the things which are more knowable and obvious to us and proceed towards those which are clearer and more knowable by nature; for the same things are not 'knowable relatively to us' and 'knowable' without qualification. So in the present inquiry we must follow this method and advance from what is more obscure by nature, but clearer to us, towards what is more clear and more knowable by nature. (Physics 184a)

Aristotle speaks of advancing from what is initially clear to us, towards what is more knowable by nature. I am not sure if he believed there was a final point where the absolutely first principles and simplest elements are known so that they will never change. Mac Lane certainly did not believe it for mathematics.

From his early work on field theory (Mac Lane and Schilling, 1939; 1940) and for the rest of his career Mac Lane often worked to find more basic concepts in some part of mathematics. He and Eilenberg spent over a decade collaborating on ever broader uses of the concepts in their "General theory of natural equivalences" (Eilenberg and Mac Lane, 1945). They meant that paper to be the only one ever needed on this technical concept for group theory and topology, but it became the founding paper of the whole field of category theory.

Only in the 1960s, after meeting graduate student Bill Lawvere, did Mac Lane come to believe category theory could be a foundation

for all mathematics. Even then, precisely because of all the concrete mathematics that had gone into developing his ideas, Mac Lane insisted this, and any foundation for mathematics, must be seen as "proposals for the organization of mathematics" (Mac Lane, 1986, p.406). The optimal organization (i.e. the optimal foundation) will change as mathematics develops, and will help advance those developments. He warned that excessive faith in any "fixed foundation would preclude the novelty which might result from the discovery of new form" (Mac Lane, 1986, p.455).

2. Mathematics as a love of wisdom

Let us come to cases with one paradigmatically philosophical question, and one paradigmatically mathematical. For Mac Lane these questions are inseparable:

Q$_1$ What are mathematical objects, and how do we come to know them?

Q$_2$ What are solutions to a Partial Differential Equation (PDE), and how do we come to know them?

For Mac Lane Q$_2$ can only be a specific case of Q$_1$. For him, as for Aristotle, basic questions of the special sciences *are* philosophy. They cannot *not* be philosophy.

Let us be clear: A mathematician can learn a textbook answer to Q$_2$ without ever asking for a philosophy behind it. In just the same way, a philosopher can learn the currently received answers to Q$_1$ from philosophy books, without ever asking about live mathematics. Admittedly the math textbook answers will be more stable over time than the philosophically received answers. But that is not important.

For Mac Lane, both of those ways of learning are failures of under-standing. They are failures of *philosophy*. For him, an answer to either one of those questions can only be valuable to the extent that you can see what it is *good for*—for Mac Lane that cannot be either a purely technical mathematical question or a purely academic philosophic one. Think back to his work in World War II.

Mathematicians speak of solutions to PDEs in many ways:

- Smooth (or, sufficiently differentiable) function solutions.
- Symbolic solutions.
- Generalized function solutions (of various kinds...).
- Numerical solutions...

These different senses of solutions are sought in very different ways. There are well understood relations between them, but the relations are not all obvious and in particular cases they may be quite difficult, and important, to find.

Mac Lane's war work certainly involved relating different kinds of solutions to PDEs. Even when an equation has a known exact solu-tion by an easily specified smooth function, applying it also requires numerical solutions. The worker has to choose which aspects are best handled in theory, so as to direct and optimize the calculations, and when best to leave theory and begin calculating. Those choices are rarely textbook work. They are often not clear cut at all. They require exactly what Aristotle called the wisdom of the craft master. Namely, they require grasping the *why* of each kind of solution. They require knowing not only the technical definition of each kind of solution, but *what good* each one is, and especially *the good* of their relations to one another.

The craft master, having wisdom, knows the whys, can teach them, and supervise work with them.

As I write this, I imagine some practice-minded philosopher challenging: "How are philosophies like logicism, formalism, and intuitionism going to help anyone solve or apply a PDE?" Indeed. This is why Mac Lane so often deprecates those philosophies. But just to give one example, Mac Lane argued that formalism in Hilbert's hands was a step towards programmable computers. See Section 3. Those unquestionably help solve PDEs.

3. The philosophy of mathematicians in 1930s Göttingen

> *Wir dürfen nicht denen glauben, die heute mit philosophischer Miene und überlegenem Tone den Kulturuntergang prophezeien und sich in dem Ignorabimus gefallen. Für uns gibt es kein Ignorabimus, und meiner Meinung nach auch für die Naturwissenschaft überhaupt nicht. Statt des törichten Ignorabimus heiße im Gegenteil unsere Losung: Wir müssen wissen – wir werden wissen!*

> We must not believe those, who today, with philosophical bearing and deliberative tone, prophesy the fall of culture and accept the ignorabimus. For us there is no ignorabimus, and in my opinion none whatever in natural science. In opposition to the foolish ignorabimus our slogan shall be: We must know – we will know! (Hilbert, 1930, p.385)[4]

Hilbert's conclusion, *Wir müssen wissen – wir werden wissen!*, is engraved on his tomb in Göttingen.

---

[4] Before we try to defend or defeat Hilbert's slogan as an assertion in academic epistemology, it is in fact the most important statement in philosophy of mathematics of the past 150 years. I use the translation by Ewald (2005, p.1164). Hilbert's address was broadcast on the radio and a recording is available at math.sfsu.edu/smith/Documents/HilbertRadio/HilbertRadio.mp3.

Mac Lane arrived in Göttingen just at the time Hilbert was promoting this slogan. I do not recall Mac Lane quoting it in lectures or conversations. He did not have to quote Hilbert. Everything Mac Lane said illustrated this faith. As you can see in Mac Lane (1986; 2005), or McLarty (2007), Mac Lane followed Hilbert's mathematizing scientific optimism, rather than the specific finitist program (often called "formalism," though not by Hilbert) of Hilbert's famous *On the Infinite* (1926). Mac Lane did see specific, productive mathematical value in that program though, and rejected a criticism of it by Freeman Dyson (Mac Lane, 1995).

Dyson supposed Hilbert seriously meant to reduce all mathematics to formal reasoning, and said "the great mathematician David Hilbert, after thirty years of high creative achievement [. . .] walked into a blind alley of reductionism." Specifically, Dyson claimed that Hlbert "dreamed of" formalizing all mathematics, solving the decision problem for this formal logic, and "thereby solving as corollaries all the famous unsolved problems of mathematics." Mac Lane replied:

> I was a student of Mathematical logic in Göttingen in 1931-1933, just after the publication of the famous 1931 paper by Gödel. Hence I venture to reply [. . .]. Hilbert himself called this 'metamathematics.' He used this for a specific limited purpose, to show mathematics consistent. Without this reduction, no Gödel's theorem, no definition of computability, no Turing machine, and hence no computers [. . .]. Dyson simply does not understand reductionism and the deep purposes it can serve.

Mac Lane gives a concise expert review of these issues and places them in the context he knew at the time they arose. He insists Hilbert did not tie the slogan "we must know, we will know" to the decision problem. Rather, Mac Lane says, "[Hilbert] held that the problems of

mathematics can all ultimately be solved" without supposing meta-mathematics will do it. Full disclosure: I admit that after long consideration, drawing on Sieg (1999; 2013), I myself am unsure exactly how Hilbert and/or Bernays intended their work on the decision problem at various times. But however that may be, Mac Lane understood Hilbert this way. And this is Mac Lane's own far from reductionist faith, while recognizing reductionist methods for what they actually have achieved in mathematics.

Mac Lane learned a lot in frequent discussions with Bernays. But for now I have to say I see no larger trace of those discussions in Mac Lane's philosophy than is found in the letter on Dyson. Mac Lane's book on philosophy of mathematics is titled *Mathematics: Form and Function*. But this is clearly "form" as Mac Lane learned about it from talking with Weyl and studying under Geiger (McLarty, 2007). It does not refer to formalism in any sense related to Bernays. Or, at least, so it seems to me. The reader is encouraged to go, read, and disagree if they see something else.

4. Naturalism

The decisive feature marking Penelope Maddy's Second Philosopher as a *naturalist* is that:

> [She] sees fit to adjudicate the methodological questions of mathematics—what makes for a good definition, an acceptable axiom, a dependable proof technique?—by assessing the effectiveness of the method at issue as means towards the goals of the particular stretch of mathematics involved. (Maddy, 2007, p.359)

Lots of mathematicians, and essentially all leading mathematicians, do the same.[5]

The unusual thing about Mac Lane in this regard is that he was explicitly tasked by the US government to evaluate mathematics research and teaching methods in classified reports during World War II and publicly after that.[6] Those reports were explicitly directed to various different specific short-term and long-term goals. All the variety he saw, and dealt with, left Mac Lane ever more deeply impressed with the actual unity of the whole.

Precisely that background, along with his experience as Chair of the Chicago Mathematics Department, made Mac Lane diverge from another feature of Maddy's Second Philosopher:

> All the Second Philosopher's impulses are methodological, just the thing to generate good science [. . .]. (Maddy, 2003, p.98)

All Mac Lane's impulses aim at producing good science and for this reason they are not *all* methodological.

Mac Lane, like Aristotle, knows methods alone generate no science. Besides evaluating methods of reaching goals, at least some mathematicians must evaluate goals. For Aristotle, those should be the craft masters, the wise. In his vivid words: "the wise should not accept orders but give them; nor should they be persuaded, but the less wise should" (Metaphysics 982a). We will see, though, Mac Lane is less focused on command than that. He inclines more to another

---

[5] If by axioms Maddy means specifically axioms of set theory then few mathematicians ever learn those, let alone adjudicate them. Mac Lane is famously among those few.

[6] See Mac Lane (1967; 1989) and Fitzgerald and Mac Lane (1977) and Steingart (2011).

passage: "those who are more accurate and more able to teach about the causes are the wiser in each branch of knowledge" (Metaphysics 982a).

Because of his broad experience, especially evaluating both methods and goals for mathematics, Mac Lane cannot agree that "the goal of philosophy of mathematics is to account for mathematics as it is practiced, not to recommend reform." (Maddy, 1997, p.161) Just sticking to the mathematician philosophers we have already named: Hilbert, Weyl, and Mac Lane all knew reform is integral to mathematical practice. You cannot separate reform from practice if you try. And all three made explicitly philosophic arguments for their recommended reforms along with more technically mathematical ones.[7] This is important for philosophy of mathematics.

The paradigm case for anti-revisionism in philosophy of mathematics is Brouwer's intuitionism. Brouwer is by far the favorite illustration of a revisionist, and is the sole example that the *Stanford Encyclopedia of Philosophy* discusses under anti-revisionism in the article "Naturalism in the Philosophy of Mathematics" (Paseau, 2016):

> The mathematician-philosopher L.E.J. Brouwer developed intuitionistic mathematics, which sought to overthrow and replace standard ('classical') mathematics.

So it is important for philosophers to understand that the problem with Brouwer, according to all our exemplars Hilbert, Weyl, and

---

[7] Hilbert had sweeping success with his reforms. Among many philosophic works by and on him see Hilbert (1923; 1930). Weyl (1918) advocated what Weyl took to be Brouwer's philosophy, while Weyl (1927; 1949) trace his eventual, regretful conclusion that in fact Hilbert was right about this and Brouwer wrong.

Mac Lane, is not that he had philosophical motives. It is that he was wrong. Actually, for Mac Lane, Brouwer's philosophy was at best wrong. At worst it was "pontifical and obscure" (Mac Lane, 1939).

Immediately upon completing his doctorate in mathematics at Göttingen, Mac Lane put a philosophy article in *The Monist* (Mac Lane, 1935). Fifty years later he wrote a book describing, as he told me, what he wanted philosophers to know about math (Mac Lane, 1986). There he asks about the large array of mathematics he surveyed: "How does it illuminate the philosophical questions as to Mathematical truth and beauty and does it help to make judgements about the direction of Mathematical research?" (Mac Lane, 1986, p.409) There is a reason he puts these questions together.

He asks about mathematical truth and beauty knowing very well that few mathematicians want to pursue the question seriously, and knowing philosophers who speak of it rarely know much of the wealth. For Mac Lane both of those are failures of understanding and they are nothing he means to promote.[8] He means to promote mathematically informed philosophic pursuit of the question of mathematical truth and beauty. And so he does of the question on the direction of research. He seriously means to promote philosophic thought on that. Of course he does not see philosophic thought as the sole preserve of those with philosophy degrees. No more does he see philosophy of mathematics as the sole preserve of those with math degrees. Mathematics for Mac Lane, when pursued with full awareness of its worth, is philosophy.

---

[8] À propos, I consider Edna St. Vincent Millay's poem "Euclid alone has looked on Beauty bare" incredibly true to its topic, despite that she apparently studied no mathematics beyond school textbooks based on bits of Euclid's *Elements*.

Bibliography

Browder, F.E., 1976. The relevance of mathematics. *The American Mathematical Monthly* [Online], 83(4), pp.249–254. Available at: https://doi.org/10.2307/2318212 [visited on 9 November 2020].

Browder, F.E. and Mac Lane, S., 1978. The relevance of mathematics. In: L.A. Steen, ed. *Mathematics Today: Twelve Informal Essays*. New York [etc.]: Springer-Verlag, pp.323–350.

Eilenberg, S. and Mac Lane, S., 1945. General theory of natural equivalences. *Transactions of the American Mathematical Society* [Online], 58, pp.231–294. Available at: https://www.ams.org/journals/tran/1945-058-00/S0002-9947-1945-0013131-6/S0002-9947-1945-0013131-6.pdf [visited on 29 October 2020].

Ewald, W., 2005. *From Kant to Hilbert: A Source Book in the Foundations of Mathematics. Vol. 2*. Oxford: Clarendon Press.

Fitzgerald, A. and Mac Lane, S., eds., 1977. *Pure and Applied Mathematics in the People's Republic of China: A Trip Report of the American Pure and Applied Mathematics Delegation*. Washington, D.C.: National Academy of Sciences.

Hilbert, D., 1923. Die logischen Grundlagen der Mathematik. *Mathematische Annalen* [Online], 88, pp.151–165. Available at: https://eudml.org/doc/158981 [visited on 9 November 2020].

Hilbert, D., 1926. Über das Unendliche. *Mathematische Annalen* [Online], 95, pp.161–190. Available at: https://eudml.org/doc/159124 [visited on 9 November 2020].

Hilbert, D., 1930. Naturerkennen und Logik. *Die Naturwissenschaften* [Online], 18(47-49), pp.959–963. Available at: https://doi.org/10.1007/BF01492194 [visited on 10 November 2020]. I cite the reprint: Naturerkennen und Logik. *Gesammelte abhandlungen. Bd. 3: Analysis, Grundlagen der Mathematik, Physic Verschiedenes*. New York: Chelsea Publishing Company Bronx, 1965, pp.378–385.

Mac Lane, S., 1935. A Logical Analysis of Mathematical Structure: *Monist* [Online], 45(1), pp.118–130. Available at: https://doi.org/10.5840/monist19354515 [visited on 10 November 2020].

Mac Lane, S., 1939. Symbolic logic (Club topics). *The American Mathematical Monthly* [Online], 46(5), pp.289–296. Available at: https://www.jstor.org/stable/2303907 [visited on 10 November 2020].

Mac Lane, S., 1946. Review of E. T. Bell The Development of Mathematics (New York, McGraw-Hill, 1945). *The American Mathematical Monthly* [Online], 53(7), pp.389–390. Available at: https://doi.org/10.2307/2305857 [visited on 10 November 2020].

Mac Lane, S., 1967. The future role of the federal government in mathematics. *The American Mathematical Monthly* [Online], 74(1 (Part 2: Fiftieth Anniversary Issue)), pp.92–100. Available at: https://doi.org/10.2307/2314871 [visited on 10 November 2020].

Mac Lane, S., 1986. *Mathematics: Form and Function*. New York [etc.]: Springer-Verlag.

Mac Lane, S., 1989. The Applied Mathematics Group at Columbia in World War II. In: P.L. Duren, R. Askey, H.M. Edwards and U.C. Merzbach, eds. *A Century of Mathematics in America. Part III*, *History of mathematics* 3. Providence, R.I.: American Mathematical Society, pp.495–515.

Mac Lane, S., 1995. 'A Matter of Temperament'. *New York Review of Books* [Online], 45(15), p.56. Available at: https://www.nybooks.com/articles/1995/10/05/a-matter-of-temperament/ [visited on 10 November 2020].

Mac Lane, S., 1997. Requiem for the Skillful. *Notices of the American Mathematical Society*, 44(2), pp.207–208.

Mac Lane, S., 2005. *A Mathematical Autobiography*. Wellesley, Mass: A K Peters.

Mac Lane, S. and Schilling, O.F.G., 1939. Zero-dimensional branches of rank one on algebraic varieties. *Annals of Mathematics* [Online], 40(3), pp.507–520. Available at: https://doi.org/10.2307/1968935 [visited on 10 November 2020].

Mac Lane, S. and Schilling, O.F.G., 1940. Normal algebraic number fields. *Proceedings of the National Academy of Sciences* [Online], 26(2), pp.122–126. Available at: https://doi.org/10.1073/pnas.26.2.122 [visited on 10 November 2020].

Maddy, P., 1997. *Naturalism in Mathematics*. Oxford: Clarendon Press.

Maddy, P., 2003. Second Philosophy. *Journal of the Indian Council of Philosophical Research* [Online], 20, pp.73–106. Available at: http://www.socsci.uci.edu/~pjmaddy/bio/2ndphilosophy.pdf [visited on 10 November 2020].

Maddy, P., 2007. *Second Philosophy: A Naturalistic Method*. Oxford [etc.]: Oxford University Press.

McLarty, C., 2007. The Last Mathematician from Hilbert's Gottingen: Saunders Mac Lane as Philosopher of Mathematics. *The British Journal for the Philosophy of Science* [Online], 58(1), pp.77–112. Available at: https://doi.org/10.1093/bjps/axl030 [visited on 10 November 2020].

McLarty, C., 2020. Saunders Mac Lane: From Principia Mathematica through Göttingen to the Working Theory of Structures. In: E. Reck and G. Scheimer, eds. *The Prehistory of Mathematical Structuralism* [Online]. New York: Oxford University Press, pp.215–237. Available at: https://doi.org/10.1093/oso/9780190641221.003.0009 [visited on 10 November 2020].

Paseau, A., 2016. Naturalism in the philosophy of mathematics. *The Stanford Encyclopedia of Philosophy* [Online]. winter 2016. Stanford: Metaphysics Research Lab, Stanford University. Available at: https://plato.stanford.edu/archives/win2016/entries/naturalism-mathematics/ [visited on 10 November 2020].

Sieg, W., 1999. Hilbert's Programs: 1917–1922. *Bulletin of Symbolic Logic* [Online], 5(1), pp.1–44. Available at: https://doi.org/10.2307/421139 [visited on 10 November 2020].

Sieg, W., 2013. *Hilbert's programs and beyond*. Oxford [etc.]: Oxford University Press.

Skowron, B., manuscript. *Was Saunders Mac Lane a platonist?* Manuscript made available in private communication. [unpublished].

Steingart, A., 2011. *Conditional inequalities : American pure and applied mathematics, 1940-1975* [Online]. Thesis. Massachusetts Institute of Technology. Available at: https://dspace.mit.edu/handle/1721.1/84367 [visited on 10 November 2020].

Weyl, H., 1918. *Das Kontinuum: Kritische Untersuchungen über die Grundlagen der Analysis*. Leipzig: Verlag von Veit&Comp.

Weyl, H., 1927. *Philosophie der Mathematik und Naturwissenschaft*, *Handbuch der Philosophie*. München; Berlin: Druck und Verlag von R. Oldenbourg.

Weyl, H., 1949. *Philosophy of Mathematics and Natural Science* (O. Helmer, Trans.). Princeton: Princeton University Press. This book makes many additions to the 1927 German version.

Creating new concepts in mathematics: freedom and limitations. The case of Category Theory

Zbigniew Semadeni

Institute of Mathematics, University of Warsaw

Abstract

In the paper we discuss the problem of limitations of freedom in mathematics and search for criteria which would differentiate the new concepts stemming from the historical ones from the new concepts that have opened unexpected ways of thinking and reasoning.

We also investigate the emergence of category theory (CT) and its origins. In particular we explore the origins of the term *functor* and present the strong evidence that Eilenberg and Carnap could have learned the term from Kotarbiński and Tarski.

Keywords

categories, functors, Eilenberg-Mac Lane Program, mathematical cognitive transgressions, phylogeny, platonism.

1. Introduction

The celebrated dictum of Georg Cantor that "The *essence* of *mathematics* lies precisely in its *freedom*" expressed the idea that in mathematics one can freely introduce new notions (which may, how-

ever, be abandoned if found unfruitful or inconvenient).[1] This way
Cantor declared his opposition to claims of Leopold Kronecker who
objected to the free introduction of new notions (particularly those
related to the infinite).

Some years earlier Richard Dedekind stated that—by forming, in
his theory, a cut for an irrational number—we *create* a new number.
For him this was an example of a constructed notion which was a free
creation of the human mind (Dedekind, 1872, § 4).

In 1910 Jan Łukasiewicz distinguished *constructive notions* from
empirical *reconstructive* ones. He referred (with reservation) to
Dedekind's statement and pointed out that a consequence of our
"creation" of those notions is the spontaneous emergence of countless
relations which no more depend on our will (Łukasiewicz, 1910).

Until the discovery of non-Euclidean geometries, geometry was
regarded as an abstraction of the spatial reality. The freedom of cre-
ation in geometry was limited by this reality. Hilbert advocated a
formal point of view, broadening the freedom of choosing the ax-
ioms, while Poincaré maintained that the axioms of geometry are
conventions.

Clearly, the freedom of mathematics is limited by logical con-
straints. At the same time, logical inference yields deep meaning to
mathematics. As Michał Heller put it, "If I accept one sentence, I
must also accept another sentence. Why must I? Who forces me?
Nobody. Yet, I must. Generally, we bear badly any restrictions of our
liberty, but in the case of mathematical deduction inevitability of the

---

[1] The italics in the original sentence "Das *Wesen der Mathematik* liegt gerade in ihrer
Freiheit" are Cantor's. The first version was published in 1879, reprinted in (Cantor,
1883, p.34), discussed by Ferreirós (1999, p.257).

conclusion gives us the feeling of safety (I have not deviated from the way) and of the accompanying intellectual comfort, sometimes even great joy" (Heller, 2015, p.21).

A mathematician trying to prove a theorem knows the feeling of an invisible wall which blocks some intended arguments. Also new concepts must be consistent with earlier ones and must not lead to contradiction or ambiguity. Moreover, in mathematical practice, only intersubjective mental constructions are accepted.

The purpose of this paper[2] is to look for restraints and patterns in the historical development of mathematics.[3] Some types of paths will be distinguished, first generally, and then they will be used to highlight some features of the rise of category theory (CT).

Michael Atiyah, in his *Fields Lecture* at the World Mathematical Year 2000 in Toronto, expressed his view that

> it is very hard to put oneself back in the position of what it was like in 1900 to be a mathematician, because so much of the mathematics of the last century has been absorbed by our culture, by us. It is very hard to imagine a time when people did not think in those terms. In fact, if you make a really important discovery in mathematics, you will then get omitted altogether! You simply get absorbed into the background. (Atiyah, 2002, p.1)

This statement may appear startling, as the mathematical meaning of a text from around 1900 is generally believed to be time-proof. Yet what Atiyah had in mind was not the meaning of published texts—definitions, theorems and proofs— but the way mathematicians thought at that time.

---

[2] The present paper is based on a talk delivered at XXIII Kraków Methodological Conference 2019: *Is Logic a Physical Variable?*, 7-9 November 2019.

[3] The significant question of degree of freedom in mathematical conceptualization of physical reality is not considered here.

It was 75 years ago when the celebrated paper by Eilenberg and Mac Lane was published.[4] This event marked the rise of category theory (CT). As the development of mathematics accelerated in the second half of the 20th century, it may be hard to fully imagine how mathematicians thought in 1945. Their definitions, theorems and comments are clear, but one should be aware that a reconstruction of their ideas, their thinking may be specifically biased by our present understanding of the mathematical concepts involved.

2. Background conceptions

The main ideas of this paper—which includes a very wide spectrum of examples, from ancient Greek mathematics to modern, from children's counting to CT—are the following:

- A mathematical concept, no matter how novel, is never independent of the previous knowledge; it is based on a reorganization of existing ideas.
- A radically new mathematical idea never germinate in somebody's mind without a period of incubation, usually a lengthy one.
- There is a long distance to cover between a spontaneous, unconscious use of a mathematical idea or structure in a concrete setting, and a conscious, systematic use of it (Piaget and García, 1989, p.25).
- A person who has achieved a higher level of mathematical thinking is often unable to imagine thinking of a person from

---

[4] Mac Lane earlier in his life, in particular in (Eilenberg and Mac Lane, 1945) and (Mac Lane, 1950) wrote his name as MacLane. Later he began inserting a space into his surname, in particular in (Mac Lane, 1971).

another epoch or of a present learner and, consequently, may unconsciously attribute to him/her an inappropriate (too high or too low) level of thinking.

Transgressions

A *mathematical cognitive transgression* (or briefly: a *transgression*) is defined as crossing—by an individual or by a scientific community—of a previously non-traversable limit of own mathematical knowledge or of a previous barrier of deep-rooted convictions (Semadeni, 2015). Moreover, it is assumed that:

1. the crossing concerns a (broadly understood) mathematical idea and the difficulty is inherent in the idea,
2. the crossing is critical to the development of the idea and related concepts,
3. it is a passage from a specified lower level to a new specified upper level,
4. the crossing is a result of conscious activity (the activity need not be intentional and purposely orientated towards such crossing; generally such effect is not anticipated in advance and may even be a surprise).

In the history of mathematics there were numerous transgressions of different importance, some great ones and many "mini-transgressions". Usually they were not single acts—they involved a global change of thinking which matured for years or even generations and they were based on the work of many people. In ancient times two transgressions were the most significant:

- The transition from practical dealing with specific geometric shapes to deductive geometry.
- The celebrated discovery (in the 5th century BC) that the diagonals of a square are incommensurable with their sides, that they are ἄλογος (*a-logos*), without a ratio, irrational (in modern setting, foreign to Greek thinking, it was the irrationality of $\sqrt{2}$) (Baszmakowa, 1975, pp.80-81). The Pythagorean paradigm was undermined by this ἀπορία (*aporia*). Their understanding of mathematics was eroded. However, nothing certain is known about this discovery. Stories presented in popular books are based on doubtful legends from sources written seven or eight centuries later (Knorr, 1975, p.21, 51). This incommensurability could not be a single discovery by an individual. It must have been a lengthy process. Never in the historical development of mathematics such a major change occurred in a short time.

In modern history there were many transgressions. Let us list some of the best known:

- Acceptance of negative numbers.
- The transition from potential infinity (infinity at a *process* level) to the *actual* infinity (infinity as an object).
- The emergence of projective geometry.
- The discovery of non-Euclidean geometries.

We will discuss the case of CT, arguing in particular that creating the theory of *elementary topoi* in CT should be regarded as a major transgression.

Phylogeny and ontogeny

The term *phylogeny* refers here to the evolutionary history of mathematics (or rather to its modern reconstructions), from ancient times on. *Ontogeny* denotes the development of basic mathematical concepts and structures in the mind of an individual person, from early childhood. Phylogeny and ontogeny are in some sense complementary descriptions (Freudenthal, 1984; Piaget and García, 1989, pp.4-29).

In case of mathematics, the oft-quoted phrase: *ontogeny recapitulates phylogeny* implies that one can learn from the history of old mathematics for the sake of present teaching. The history of mathematics sometimes gives useful hints, e.g. one may argue that since the historical process of forming the general concept of a function took centuries (from Descartes, if not much earlier, to Peano and Hausdorff), we should not expect that a secondary school student can grasp it—learning Dirichlet's description—after a few lessons. The general concept of the function was not yet quite clear to mathematicians of the first half of the 19th century (Lakatos, 1976, Appendix 2; Youschkevitch, 1976; Ferreirós, 1999, pp.27–30).

On the other hand, the idea that ontogeny recapitulates phylogeny may be misleading. Piaget always stressed that arithmetic cognition results from logico-mathematical experience with concrete objects, pebbles say, and is educed from *the child's actions* rather than from heard words. It is abstracted from a coordination of intentional motions and accompanying thoughts. Nevertheless, Piaget was in favour of the phylogeny-ontogeny parallelism and reasoned roughly as follows. Since one-to-one correspondence preceded numerical verbal counting in the very early periods of human civilization (evidenced by artefacts such as notched bones and also found in crude unlettered tribes), the same should apply to children. Cantor's theory of cardinal

numbers confirmed this thinking (Beth and Piaget, 1966, pp.259–260). Consequently, Piaget and many educators insisted on one-to-one correspondence as a foundation of early school arithmetic, neglecting the fact that nowadays children learn number names early, often together with learning to speak, and moreover counting is now deeply rooted culturally. Research evidence shows that counting, rather than one-to-one correspondence, is a basis of the child's concept of number (Gelman and Gallistel, 1978, pp.77–82). In this way the phylogeny–ontogeny parallelism adversely affected early mathematics education in the time of the 'New Math' movement.

Hans Freudenthal, in the context of mathematics, suggested the converse idea: *What can we learn from educating the youth for understanding the past of mankind?* This reverses the traditional direction of inference in the phylogeny–ontogeny parallelism. In particular, one may ask whether contemporary knowledge of the difficulties in the transition from the concrete to more abstract mathematical reasoning of children may be helpful in better understanding of limitations of our reconstructions of the development of the early Greek mathematics.

In the sequel, certain aspects of the development will be traced both in phylogeny and in ontogeny, inextricably intertwined with the the mathematical questions themselves.

Platonizing constructivism in mathematics

The theoretical framework of the paper is platonizing constructivism in mathematics (Semadeni, 2018). It is assumed that:

- each of the three major positions in the philosophy of mathematics from the beginning of the 20th century: *platonism*, *constructivism*, *formalism* describes some inherent, complementary features of mathematics;

- these three major positions can be reconciled provided that they are regarded as *descriptive*, as an account of some inherent features of mathematics, and *not normative*, i.e., when one leaves out the eliminating words (as 'only', 'oppose') which explicitly deny other standpoints. Moreover, various versions of the three positions often overlap.

In the sequel, the term *constructivism* will not be understood as in papers on foundations of mathematics, but rather in a way akin to its meaning in research on mathematics education, related to post-Piagetian psychological versions of constructivism. Briefly, one assumes here that humans construct mathematical concepts in their minds and discover their properties. A concept develops its necessary structure as a consequence of its context and—in the long term— becomes *cristalline* in the sense of David Tall (2013, p.27); then its properties appear independent of our will. This phenomenon may be traced both in phylogeny and in ontogeny. Moreover, in each essential progression, new mental structures are build on the preceding ones and are always integrated with previous ones (Piaget and García, 1989, pp.22–29).

By platonizing constructivism we mean an analysis—in constructivistic terms—of sources and consequences of the platonistic attitude of a majority of mathematicians and contrasting them with the well-known difficulties of consistent platonism in the philosophy of mathematics.

3. Developmental successors

After the introductory examples we now look for ways to distinguish between:

- mathematical concepts which—historically—were natural successors to previous ones,
- mathematical concepts which could be conceived and defined only after opening new paths of thought and reasoning.

The following metaphorical labels will be used: *onward development*, *branching-off*, *upward development*, *downward development*. These labels will be interpreted with examples. We do not expect to find clear criteria, but the ensuing discussion may be illuminating.

The *emergence of numerals* in the Late Stone Age is evidenced by tally marks (in the form of notched bones). Ethnologists have found that early tribes had only two counting words: *one* and *two*, followed by *many*. Also in present Indo-European languages these two numbers and their ordinal counterparts are linguistically different from the following numbers. It has been suggested that the proto-Indo-European number **trei* (three) was derived from the verb **terh* (meaning: *pass*); thus, the word *three* is related to *trans*. This may be a hint of a very ancient mental obstacle between numbers *two* and *three*.

One may conjecture that *after the passage from 2 to 3 there was no notable obstacle to gradual development of unlimited counting*. Of course, the actual development took centuries, if not millennia. Anyway, for present children there is no hurdle between 2 and 3, as they are taught counting very early. Moreover, counting starts to make sense with three items.

Onward development

Onward development of indefinite counting includes its *developmental successors*: simple addition of natural numbers (which develops through a stage called *count all* and then a more advanced stage *count on*), subtraction (as taking away), multiplication, division (originally there are two kinds of it: *equal sharing* and *equal grouping*), and even simple powers, all within some range of natural numbers.

These concepts are included in the onward development of counting, by virtue of the following features:

- *no branching*: each new concept naturally comes after the previous ones;
- *ontological stability*: each concept (e.g., number 17 or the product 3×6), remains essentially the same object, although the related ideas are enriched after each extension of the scope of arithmetic and—in the historical development—are subject to evolutionary changes.

The conception of developmental successors, outlined here, does not take into account a relative difficulty of concepts; what is crucial is whether these concepts follow the previous lines of thinking.

Branching-off

The concept of "branching-off" arises from a negation of the first requirement in the description of an onward development. An example of it are fractions, which *branch off* from natural numbers; it is not onward development, although there are many ties between natural numbers and fractions.

Fractions could be introduced to children in two ways. In the first, some idealized whole is divided into m equal parts and then n of them are taken. In the second, n whole things are equally divided into m parts. They are two main aspects of the concept of a fraction. For instance, $\frac{3}{4}$ of pizza may be obtained by cutting it into 4 parts and taking 3 such parts (thus $\frac{3}{4} = 3 \times \frac{1}{4}$). A more advanced way of thinking of $\frac{3}{4}$ is 3 divided by 4; the latter may be explained with the example of 3 pizzas to be divided among 4 persons.

The distinction looks quite elementary. Yet, it was significant in the phylogeny of fractions. First procedure is akin to that of ancient Egyptians, the second—to Greek ratios; both were inherited by Arabic mathematicians. In the ontogeny the two ways are always present, but not necessarily noticed. The following reminiscence by William Thurston (1946-2012), written 8 years after he had received Fields Medal, describes his discovery of the identification of previously different objects.

> I remember as a child, in fifth grade, coming to the amazing (to me) realization that the answer to 134 divided by 29 is $134/29$ (and so forth). What a tremendous labor-saving device! To me, '134 divided by 29' meant a certain tedious chore, while $134/29$ was an object with no implicit work. I went excitely to my father to explain my major discovery. He told me that of course this is so, a/b and a divided by b are just synonyms. To him it was just a small variation in notation (Thurston, 1990, p.848).

The fraction $\frac{a}{b}$ becomes identified with the result of division $a \div b$ and—in this synthesis—they both form a single mathematical object. Philosophically, however, it is an ontological change: two different beings, results of two different mental constructions, become regarded as a single one.

Onward branching-off can be traced in many parts of mathematics. Calculus branches off from a theory of the field of real numbers (axiomatic or based on a construction). Infinite sequences of real numbers branch off from elementary algebra of real numbers. Limits of sequences branch off from general theory of sequences. These examples vividly show that the question of distinguishing branching-off from onward development is delicate, as the criteria are far from being precise, but it may contribute to better understanding the historical development of mathematics.

Upward development and downward development

By *upward development* of a piece of mathematics we mean passing from some concepts and relations between them to a more abstract version of them. Examples:

- Transition from practical addition (verbalized as, e.g., *two and three make five*) to symbolic version (e.g., $2 + 3 = 5$) took centuries (the sign $+$ appeared in some 15th century records; the first occurrence of the equality sign $=$ was found in a text by a Welsh mathematician Robert Recorde from 1557).
- Transition from the space $\mathbb{R}^n$ to an axiomatically given vector space over $\mathbb{R}$.
- Transition from a vector space over $\mathbb{R}$ to a vector space over a field.
- Transition from group theory to the category **Grp**.
- Transition from a category to a metacategory (in the sense of (Mac Lane, 1971, pp.7–11)).

Downward development is—in some sense—an inverse process, to more concrete questions or to a lower level of abstraction, so the above examples may be used the other way round. Also some typical applications of mathematics may be included here, e.g., the passage from abstract Boolean algebras to a description of certain types of electrical circuits, as conceived by Claude Shannon (1936).

In the 20th century the mathematics grew rapidly and the upward development became much easier mentally as a result of both: a general change of the attitude of mathematicians toward abstraction and the routine of expressing all concepts in the language of set theory. Branching-off were so frequent that the above metaphors are of little use. There is, however, a notable exception: a new theory which opened a new direction of thinking, so its beginnings may be discussed in a way akin to that used with respect to distant past.

4. The rise of category theory (CT)

A very special feature of CT is that it has a pretty precise date of its official birth: the publication of the paper by Samuel Eilenberg and Saunders Mac Lane (1945). It was presented at a meeting of the American Mathematical Society in 1942 and published in 1945.

According to the Stanford Encyclopedia of Philosophy, *CT "appeared almost out of nowhere"*. Not quite so. As in any mathematical theory, some CT ideas had been conceived much earlier, particularly in algebraic topology, and some of them can be traced to the 19th century.

Many conceptual transformations—either explicit and well recognized or used implicitly, without awareness—contributed to the rise of CT. Going back, a significant factor was the historic development of the mathematical concept of a function.

Until the beginning of the 19$^{\text{th}}$ century, a general symbol for a function (f or φ) was almost non-existent (Youschkevitch, 1976). A significant step toward CT was the general notion of a mapping introduced by Dedekind in 1888. In his *Enklärung* (*explanation*) he did not define the concept of *Abbildung* φ (literally: *image* or *representation*) from a set (*System*) S into a set S', but interpreted it generally as an arbitrary law (*Gesetz*) according to which to each element s there corresponds (*gehört*) a certain thing (*Ding*) $S' = \varphi(s)$, called the image (*Bild*) of s. He also defined a composed mapping (*zusammengesetzte Abbildung*) $\vartheta(s) = \psi(s') = \psi(\varphi(s))$ of two given ones, denoted as $\varphi.\psi$ or $\varphi\psi$, defined injective mappings (*ähnlich* or *deutlich*), the inverse mapping and proved their main properties (Dedekind, 1888, §2–4; Ferreirós, 1999, p.88–90, 228–229).

Emmy Noether in her lectures in the 1920s emphasised the role of homomorphisms in group theory. Before her, groups were understood as generators and relations (in modern terms, as quotients of free groups). She also argued that the homology of a space is a group, is an algebraic system rather than a set of numbers assigned to the space. Her lectures and the lectures of Emil Artin formed a basis for the celebrated book by van der Waerden (1930) on modern algebra. This current of thought led to CT.

Generally, in the symbol of the type $f(x)$, the part f was always understood as fixed and x was a variable. At the end of the 1920s, however, in functional analysis and related fields, a new way of thinking emerged. In certain situations the roles of symbols in $f(x)$ reversed: the point x was regarded as fixed while the function f became a vari-

able (as an element of a function space), e.g., $x \in [0, 1]$ was fixed and f was a variable in the space $C([0, 1])$ of continuous functions on the interval $[0, 1]$. In this new role, the point x became a functional δ_x on $C([0, 1])$. Such a change of the roles function-element became crucial, e.g., in the Potryagin duality of locally compact abelian groups (Hewitt and Ross, 1963) and in Gelfand's theory of commutative Banach algebras. It was also used by Eilenberg and Mac Lane in their first example (finite dimensional vector spaces and their dual spaces) motivating the concept of a natural equivalence.

A crucial example of a contravariant functor was the adjoint T^* of a linear operator T on a Hilbert space, introduced in 1932 by John von Neumann.[5]

According to Mac Lane, abstract algebra, lattice theory and universal algebra were necessary precursors for CT. However, he also suggested that certain notational devices preceded the definition of a category. One of them was the fundamental idea of representing a function by an arrow $f\colon X \to Y$, which first appeared in algebraic topology about 1940, probably introduced by Polish-born topologist Witold Hurewicz (Mac Lane, 1971, p.29; 1988, p.333). Originally, it looked as just another symbol, but from a later perspective the use of such symbol was one of the key changes. Thus, a notation (the arrow) led to a concept (category). Such new symbols got absorbed into the background of mathematical thinking, and was later used as something obvious. Together with commutative diagrams, which were probably also first used by Hurewicz, they paved the way to CT.

---

[5] Mac Lane (1988, p.330) tells a story of how Marshall Stone advised von Neumann to introduce the symbol T^* and how it changed the publication. He also mentions a fact which may interests philosophers: in 1929 von Neumann lectured in Göttingen and presented his axiomatic definition of a Hilbert space, while David Hilbert—listening to it—evidently thought of it as of the concrete space ℓ^2, not in the axiomatic setting.

Mac Lane often accented two features of mathematics: computational and conceptual. He noted that the initial discovery of CT came directly from a problem of calculation in algebraic topology (Mac Lane, 1988, p.333).

Eilenberg and Mac Lane were aware that they introduced very abstract mathematical tools, which did not fit any algebraic system in the Garrett Birkhoff's universal algebra. It might seem too abstract and was certainly off beat and a "far out" endeavour. Although it was carefully prepared, it might not have seen the light of day (Mac Lane, 2002, p.130).

The origin of the term *functor*

Mac Lane has written *Categories, functors, and natural transformations were discovered by Eilenberg–Mac Lane in 1942* (Mac Lane, 1971, p.29). The word "discovered" may be regarded as an indication of a hidden Platonistic attitude of Mac Lane, in spite of his verbal declarations against Platonism (Mac Lane, 1986, pp.447–449; Król, 2019; Skowron, manuscript). He also wrote:

> Now the discovery of ideas as general as these is chiefly the willingness to make a brash or speculative abstraction, in this case supported by the pleasure of purloining words from the philosophers: "Category" from Aristotle and Kant, "Functor" from Carnap (*Logische Syntax der Sprache*) (Mac Lane, 1971, pp.29–30).

This sentence has been taken very seriously by several authors. However, the way it was phrased suggests that it was rather intended to be a

delicate joke.[6] Attributing the origin of the term *category* to Aristotle and Kant is clear, although in 1899 René-Louis Baire (in his *Thèse*) introduced—in another context—the word *category* to mathematics.[7] Concerning the origin of the term *functor* in CT, one may recall the following facts:[8]

- In Rudolf Carnap's book *Abriss der Logistik* (Carnap, 1929) the term "Funktor" does not appear.
- In 1929 the Polish term "funktor" was used in propositional calculus by Tadeusz Kotarbiński in his book (Kotarbiński, 1929).
- Alfred Tarski often emphasised that Kotarbiński had been his teacher (Feferman and Feferman, 2004, part 2).
- Carnap met Tarski in Vienna in February 1930 and visited Warsaw in November 1930; he learned much from Tarski.

---

[6] Let us note that the noun *brash* means a mass of fragments; according to *Cambridge International Dictionary of English* (1995), the adjective *brash* is disapproving, referring to people who show too much confidence and too little respect, while *Webster's New World Dictionary* (1984) lists—as meanings of *brash*—also *hasty and reckless*, *offensively bold*. On the other hand, the word *purloining* means *stealing* or *borrowing without permission*. Such a comment (with the word *pleasure*) by Mac Lane concerning CT could not be serious. On the other hand, in 2002 Mac Lane came back to Carnap, adding: "Also the terminology was largely purloined: "category" from Kant, "natural" from vector spaces and "functor" from Carnap. (It was used in a different sense in Carnap's influential book *Logical Syntax of Language*; I had reviewed the English translation of the book (in the Bulletin AMS 1938) and had spotted some errors; since Carnap never acknowledged my finding, I did not mind using his terminology)" (Mac Lane, 2002, pp.130–131).

[7] A subset A of a topological space X is called *a set of first category* (*un ensemble de première catégorie*) in X iff A is the union of a countable family of nowhere dense sets; otherwise it is *a set of the second category* (*Menge erster und zweiter Kategorie*). The celebrated *Baire category theorem* states, in a generalized form, that a complete metric space is not a set of the first category (Hausdorff, 1914, p.328; Kuratowski, 1933, § 10). The clumsy term *set of the first category* was later replaced by the term a *meager set* (Kelley, 1955, p.201). In the 1930s Baire category theorem was a very popular tool in the Warsaw school of topology, so Eilenberg must have known it.

[8] The author is indebted to Professor Jan Woleński for the relevant information.

- In 1933 Tarski, in the Polish version of his famous paper *On the concept of truth in formal languages* Tarski (1933) used the term "funktor" and mentioned that he owed the term to Kotarbiński.
- Carnap used the term "Funktor" in his book (1934) (quoted by Mac Lane) in a sense more general than that of Kotarbiński and Tarski.
- Eilenberg studied mathematics in Warsaw from 1930. In 1931 he attended Tarski's lectures on logic (Feferman and Feferman, 2004, part 3 and 12). He left Warsaw in 1939.

This evidence strongly suggests that both Carnap and Eilenberg could have learned, independently, the term from Kotarbiński and Tarski.

Was the original CT an onward development of previous mathematical theories?

Using a metaphor explained above, one may argue that the definition of a category and of a functor were within a major onward development of part of mathematics of the first half of the 20th century, that is set theory, algebra, topology etc. Indeed, for a person working in group theory, say, a natural continuation should be to think of all groups, their homomorphisms, isomorphisms, and the composites as of a single whole: *group theory*. Similarly one could think of vector spaces with linear maps as of another whole. Some analogies between theories were obvious. Moreover, the axioms of CT are reminiscent of those of semigroup theory. The concept of a covariant functor was a natural analogue of homomorphisms of algebras. Contravariant functors had been present in various duality theories (e.g., in Pontryagin's

- In 1933 Tarski, in the Polish version of his famous paper *On the concept of truth in formal languages* Tarski (1933) used the term "funktor" and mentioned that he owed the term to Kotarbiński.
- Carnap used the term "Funktor" in his book (1934) (quoted by Mac Lane) in a sense more general than that of Kotarbiński and Tarski.
- Eilenberg studied mathematics in Warsaw from 1930. In 1931 he attended Tarski's lectures on logic (Feferman and Feferman, 2004, part 3 and 12). He left Warsaw in 1939.

This evidence strongly suggests that both Carnap and Eilenberg could have learned, independently, the term from Kotarbiński and Tarski.

Was the original CT an onward development of previous mathematical theories?

Using a metaphor explained above, one may argue that the definition of a category and of a functor were within a major onward development of part of mathematics of the first half of the 20th century, that is set theory, algebra, topology etc. Indeed, for a person working in group theory, say, a natural continuation should be to think of all groups, their homomorphisms, isomorphisms, and the composites as of a single whole: *group theory*. Similarly one could think of vector spaces with linear maps as of another whole. Some analogies between theories were obvious. Moreover, the axioms of CT are reminiscent of those of semigroup theory. The concept of a covariant functor was a natural analogue of homomorphisms of algebras. Contravariant functors had been present in various duality theories (e.g., in Pontryagin's

duality mentioned above). CT provided general concepts applicable to all branches of abstract mathematics, contributed to the trend towards uniform treatment of different mathematical disciplines, provided opportunities for the comparison of constructions and of isomorphisms occurring in different branches of mathematics, and may occasionally suggest new results by analogy (Eilenberg and Mac Lane, 1945, p.236).

The great achievement of Eilenberg and Mac Lane was the idea that a formalization of various evident analogies was worth systematizing and publishing. The initial neglect of the paper of Eilenberg and Mac Lane (1945) by mathematicians was very likely a result of the fact that it was regarded as a long paper within onward development of known part of mathematics, with many rather simple definitions and examples, tedious verification of easy facts, and no theorem with an involved proof. Ralf Krömer, in his book on the history and philosophy of CT, has outright stated that Eilenberg and Mac Lane needed to have remarkable courage to write and submit for publication the paper almost completely concerned with conceptual clarification (Krömer, 2007, p.65).

A novelty of (Eilenberg and Mac Lane, 1945, p.272), which at first appeared insignificant, was regarding elements p_1, p_2 of a single quasi-ordered set P as objects of a category, with a unique morphism $p_1 \rightarrow p_2$ iff $p_1 \leq p_2$ and no morphisms otherwise. This opened a way to a series of generalizations, in particular regarding certain commutative diagrams as functors on small categories.

One may argue that this achievement was still within onward development of CT as it was within the scope of previous knowledge. Let us recall that the difficulty and the originality of a theorem are not taken into account; what is crucial is whether the concepts involved are natural extension of the previous knowledge and thinking.

The category axioms represent a very weak abstraction (Goldblatt, 1984, p.25). In spite of this fact, a few years later the conceptual clarification turned out to be highly effective in the book *Foundations of Algebraic Topology* written by Eilenberg together with Norman Steenrod (1952). The latter admitted in a conversation that the 1945 paper on categories had a more significant impact on him than any other research paper, it changed his way of thinking.

The Eilenberg-Mac Lane Program

This program has been formulated as follows:

> The theory [. . .] emphasizes that, whenever new abstract objects are constructed in a specified way out of given ones, it is advisable to regard the construction of the corresponding induced mappings on these new objects as an integral part of their definition. The pursuit of this program entails a simultaneous consideration of objects and their mappings (in our terminology, this means the consideration not of individual objects but of categories). [. . .]

> The invariant character of a mathematical discipline can be formulated in these terms. Thus, in group theory all the basic constructions can be regarded as the definitions of co- or contravariant functors, so we may formulate the dictum: The subject of group theory is essentially the study of those constructions of groups which behave in a covariant or contravariant manner under induced homomorphisms. More precisely, group theory studies functors defined on well specified categories of groups, with values in another such category. This may be regarded as a continuation of the Klein Erlanger Programm, in the sense that a geometrical space with its group of transformations is generalized to a category with its algebra of mappings.

> [...] such examples as the "category of *all* sets", the "category of *all* groups" are illegitimate. The difficulties and antinomies are exactly those of ordinary intuitive *Mengenlehre*; no essentially new paradoxes are involved. [...] we have chosen to adopt the intuitive standpoint, leaving the reader free to insert whatever type of logical foundation (or absence thereof) he may prefer. [...]

> It should be observed first that the whole concept of a category is essentially an auxiliary one; our basic concepts are essentially those of a functor and of a natural transformation. [...] The idea of a category is required only by the precept that every function should have a definite class as domain and a definite class as range, for the categories are provided as the domains and ranges of functors. (Eilenberg and Mac Lane, 1945, p.236–237, 246–247)

The quoted comparison of CT to the celebrated program of Felix Klein shows vividly that the authors regarded their work as significant. An important novelty of (Eilenberg and Mac Lane, 1945) was to use the same letter to denote both: the object component of a functor and its morphism component. This had not been a common practice, even when these correspondences were dealt with in a single paper. This novelty and the whole program fit well Atiyah's conception (quoted above) of ideas *absorbed by mathematicians' culture*.

CT is both a specific domain of mathematics and at the same time a conceptual framework for a major part of modern theories.

5. The amazing phenomenon of unexpected branchings-off in CT

Up to this point CT might be regarded as being within onward development of earlier theories: algebra, topology, functional analysis.

However, a branching-off in (Eilenberg and Mac Lane, 1945) is the concept of a *natural equivalence* (central in the title of the paper), with an essential use of commutative diagrams.[9] It was a completely new idea, but its initial impact was limited.

After 1945 CT—as a general theory—lay dormant till the emergence of significant new concepts and a breaking series of major branchings-off in the second half of the 1950s. One of their outstanding features was a new type of a definition, formulated in the form of a *unique factorization problem*.

First such explicit definition appeared in the paper by Pierre Samuel (1948), a member of the Bourbaki group, on free topological groups, albeit it was still in the language of set theory, without arrows. Two years later commutative diagrams were demonstrated as a convenient tool in such problems by Mac Lane (1950). The concept of a *dual category*, formulated by Eilenberg and Mac Lane in (1945, p.259) and further developed by Mac Lane (1950), had its conceptual roots in various duality theories, particularly in that of projective geometry.

The *product* of two categories was an analogue of that for groups and various algebras. Mac Lane analysed the concept of duality, stressed diagrammatic dualities of various pairs of concepts and presented the definitions of *direct* and *free products* in group theory (later generalized to the concepts of categorial *products* and *coproducts*, respectively).

---

[9] It is not clear why Eilenberg and Mac Lane refrained from setting the concept in the general form of a *natural transformation* (examples abounded). Perhaps they felt they should not pursue a still more general setting without accompanying results.

A metamorphosis from Eilenberg–Mac Lane Program to mature CT

A turning point in the development of CT was the seminal paper by Daniel Kan (1958) on *adjoint functors*. In much the same time period, independently, several closely related concepts and results were worked out: *representable functors*, *universal morphisms*, *Yoneda lemma*, various types of *limits* and *colimits* (Mac Lane, 1988, pp.345–352). Special cases of them had a long earlier history in specific situations in algebra and topology (e.g., Freudenthal's theorems on *loops* and *suspensions* in homotopy theory proved in 1937). This confirms a known phenomenon that mathematicians may use an idea spontaneously, without being conscious of it in a more abstract setting.

Mac Lane (1950) also opened the way to the study of categories with additional structure, which some years later developed to the study of abelian categories. This topic was developed—in a remarkably short time—due to the work of Alexandre Grothendieck, David Buchsbaum, Pierre Gabriel, Max Kelly and authors of two monographs: Peter Freyd (1964) and Barry Mitchell (1965).

Within 20 years CT, originally conceived as a useful language for mathematicians, became a developed, mature theory, something totally unexpected by its founders.

Set theory without elements

The results of the work of William Lawvere turned out to be not only a new branch of CT, but also opened new perspectives in mathematics, logic, foundations of mathematics, and philosophy. In his Ph.D. thesis at Columbia University in New York, supervised by

Eilenberg (defended in 1963, known from various copies, with full text published 40 years later) many new ideas were presented, including a categorical approach to algebraic theories (Lawvere, 1963).

Lawvere also tackled the general question as to what conditions a category must satisfy in order to be equivalent to the category **Set**. The idea looked analogous to the so-called *representation theorems*, i.e., propositions asserting that any model of the axioms for a certain abstract structure must be (in some prescribed sense) isomorphic to a specific type of models of the theory or to one particular concrete model.[10] However, Lawvere's case was unique and controversial in the sense that his 'sets' were conceived *without elements*. The theory did not have the primitive notion "element of". And it did work.

Specifically, Lawvere characterized **Set** (up to *equivalence* of categories) as a category $\mathcal{C}$ with the following: an *initial* object **0**; a *terminal* object **1** (which gives rise to *elements* of A defined as morphisms from **1** to A); *products* and *coproducts* of finite families of objects; *equalizers* and *coequalizers*; for any two objects there is an *exponential*; existence of a specific object **N** with morphisms $0\colon \mathbf{1} \to \mathbf{N}$ and $s\colon \mathbf{N} \to \mathbf{N}$ yielding the successor operation s on **N** and a simple recursion for sequences; axiom that **1** is a *generator* (if parallel morphisms f, g are not equal, then there is an element $x \in A$ such that $xf \neq xg$); axiom of choice; three additional elementary axioms of this sort (everything in the language of CT). This was augmented with one non-elementary axiom: $\mathcal{C}$ has products and coproducts for any *indexing infinite set*. A coproduct of copies of **1** played the role of a set (Lawvere, 1964; Mac Lane, 1986, pp.386–407; 1988, pp.341–345).

---

[10] The oldest theorems of this type are: Cayley's theorem that every (abstract) group is isomorphic to a group of bijections of a set; Kuratowski's theorem that every partially ordered set is order-isomorphic to a family of subsets of a set, ordered by inclusion; theorem that every group with one free generator is isomorphic to $\mathbb{Z}$. Analogous examples are known in many theories.

The point was not to avoid membership relation completely, but (instead of taking as the starting point the primitive notions of *set*, *element* and membership ∈) *function* may be taken as a primitive notion of the theory (with suitable axioms, using elementary logic, but avoiding any reference to sets) and then one derives membership and most concepts of set theory as a special case from there.

In the second half of the 1960's Lawvere opened a way to a new theory of *elementary toposes* (called also *elementary topoi*, with Greek plural $\tau\acute{o}\pi o\iota$ of the noun $\tau\acute{o}\pi o\varsigma$). Unexpected territories of mathematics were discovered (Lawvere, 1972; Mac Lane, 1988, pp.352–359; Krömer, 2007).

CT became a contender for a foundation of mathematics, although the hope that it undermine the overwhelming role of set theory turned out unfounded and most working mathematicians keep away from CT and toposes. CT yields new tools to study many formal mathematical theories and mutual relations between them, from a perspective different from that set theory.

6. Recapitulation of some points

Let us recall Atiyah's remark (quoted in the Introduction) that really important discoveries get later omitted altogether as they become absorbed by the general mathematical culture. This thought fits particularly well with the case of CT. Most of the ideas presented by Eilenberg and Mac Lane in 1945 have been absorbed as a natural language of advanced mathematical thinking. Once mathematicians learnt the definitions of a functor and a natural transformation, these concepts became a major tool of mathematical reasoning in many ab-

stract theories of the second half of the 20th century. However, it took several years to realise the scope of the change. Freyd commented as follows:

> MacLane's definition of "product" (1950) as the solution of a universal mapping problem was revolutionary. So revolutionary that it was not immediately absorbed even by most category minded people.

> [...] In a new subject it is often very difficult to decide what is trivial, what is obvious, what is hard, what is worth bragging about (Freyd, 1964, p.156).

Mac Lane, however, used the definition of a product and its dual only in the case of groups (general or abelian). He did not formulate it *mutatis mutandis* in the general case of a category, although he had several simple examples at hand. Freyd also told the story of the term *exact sequence*, a technical definition in homological algebra. In the late 1950's, when he was a graduate student at Brown University, he was brought up to think in terms of exactness of maps. This concept seemed to him as fundamental as the notion of continuity must seem to an analyst. And later he was astonished to hear that when Eilenberg and Steenrod wrote their fundamental book (1952) they defined this very notion, recognized the importance of the choice of a suitable name for it, yet could not invent any satisfactory word. Consequently, they wrote the word "blank" throughout most of the manuscript, ready to replace it before submitting the book for publication. After entertaining an unrecorded number of possibilities they settled on "exact" (Freyd, 1964, p.157).

One may argue that the 1945 definitions of a category and of a functor were within a major onward development of abstract algebra and other advanced topics. In fact, originally they were not regarded

as a novelty. Eilenberg and Mac Lane were not even certain whether their paper will be accepted for publication (it was long and lacked theorems with substantial proofs). However, they were genuinely convinced of the significance of their conceptual clarification and took pains to write the paper clearly to attract the reader.

After this publication for almost ten years CT appeared dormant. The groundbreaking papers on abelian categories by Buchsbaum and Grothendieck marked a far-reaching change. And then—in the 1960's—CT unexpectedly started to grow rapidly, with astonishing results (Mac Lane, 1988, pp.338–339, 341–361).

Thus, from the present perspective, in spite of the previous arguments, one can say that the emergence of CT was undoubtedly a major transgression in mathematics. It was a crossing of a previously non-traversable barrier of deep-rooted habits to think of mathematics. A vivid argument is the fact that—even after publication of the main ideas—it was so difficult to overcome the previous inhibition and widespread tradition.[11]

The creators of CT and their followers could choose their definitions freely, nobody could forbid that. And yet the previous way of thinking was an obstacle for potential authors and for prospective readers. Great insight of Eilenberg and Mac Lane of what is significant in mathematics turned out a crucial factor.

---

[11] Many mathematicians, in USA and elsewhere, expressed disinclination about CT. Karol Borsuk, an outstanding topologist, the teacher of Eilenberg in Warsaw and coauthor of their joint paper published in 1936, was later critical of CT and the categorical methods in mathematics (Jackowski, 2015, p.30). Jerzy Dydak, a student of Borsuk, recalled after years: *My own PhD thesis written under Borsuk in 1975 makes extensive use of category theory and I was asked by him to cut that stuff out. Only after I assured him that I spent many months trying to avoid abstract concepts, he relinquished and the thesis was unchanged* (Dydak, 2012, p.92).

Bibliography

Atiyah, M., 2002. Mathematics in the 20th century. *Bulletin of the London Mathematical Society* [Online], 34(1), pp.1–15. Available at: https://doi.org/10.1112/S0024609301008566 [visited on 29 October 2020].

Baszmakowa, I.G., 1975. Grecja starożytna. Kraje hellenistyczne i imperium rzymskie. In: A.P. Juszkiewicz, ed. *Historia matematyki od czasów najdawniejszych do początku XIX stulecia*. Vol. 1: *Od czasów najdawniejszych do początku czasów nowożytnych* (S. Dobrzycki, Trans.). Warszawa: Państwowe Wydawnictwo Naukowe, pp.64–168.

Beth, E.W. and Piaget, J., 1966. *Mathematical Epistemology and Psychology*. Dordrecht: D. Reidel Publishing Company.

Cantor, G., 1883. Über unendliche, lineare Punktmannigfaltigkeiten. *Mathematische Annalen* [Online], 21(4), pp.545–591. Available at: https://doi.org/10.1007/BF01446819 [visited on 29 October 2020].

Carnap, R., 1929. *Abriss der logistik: mit besonderer Berücksichtigung der relationstheorie und ihrer Anwendungen*, *Schriften zur Wirtschaftswissenschaftlichen Forschung* Bd. 2. Wien: Verlag von Julius Springer.

Carnap, R., 1934. *Logische Syntax der Sprache*, *Schriften zur Wissenschaftlichen Weltauffassung* Bd. 8. Wien: Verlag von Julius Springer.

Dedekind, R., 1872. *Stetigkeit und irrationale Zahlen* [Online]. Braunschweig: Friedrich Vieweg & Sohn. Available at: http://www.rcin.org.pl/publication/13029.

Dedekind, R., 1888. *Was sind und was sollen die Zahlen?* [Online]. Braunschweig: Friedrich Vieweg & Sohn. Available at: http://www.digibib.tu-bs.de/?docid=00024927 [visited on 30 October 2020].

Dydak, J., 2012. Ideas and Influence of Karol Borsuk. *Wiadomości Matematyczne* [Online], 48(2), p.81. Available at: https://doi.org/10.14708/wm.v48i2.305 [visited on 30 October 2020].

Eilenberg, S. and Mac Lane, S., 1945. General theory of natural equivalences. *Transactions of the American Mathematical Society* [Online], 58, pp.231–294. Available at: https://doi.org/10.1090/S0002-9947-1945-0013131-6 [visited on 30 October 2020].

Eilenberg, S. and Steenrod, N.E., 1952. *Foundations of Algebraic Topology*, *Princeton Mathematical Series* 15. Princeton: Princeton University Press.

Feferman, A.B. and Feferman, S., 2004. *Alfred Tarski: Life and Logic*. Cambridge [etc.]: Cambridge University Press.

Ferreirós, J., 1999. *Labyrinth of Thought: A History of Set Theory and Its Role in Modern Mathematics*, *Science Networks. Historical Studies* vol. 23. Basel: Birkhäuser.

Freudenthal, H., 1984. The Implicit Philosophy of Mathematics: History and Education. In: Z. Ciesielski and C. Olech, eds. *Proceedings of the International Congress of Mathematicians, August 16-24, 1983, Warszawa*. Warszawa; Amsterdam [etc.]: PWN-Polish Scientific Publishers; North-Holland, pp.1695–1709.

Freyd, P., 1964. *Abelian Categories: An Introduction to the Theory of Functors*. New York: Harper & Row.

Gelman, R. and Gallistel, C.R., 1978. *The Child's Understanding of Number*. Cambridge Mass; London: Harvard University Press.

Goldblatt, R., 1984. *Topoi – The Categorial Analysis of Logic*. North-Holland.

Hausdorff, F., 1914. *Grundzüge der Mengenlehre* [Online]. Leipzig: Von Veit & Comp. Available at: http://gallica.bnf.fr/ark:/12148/bpt6k9736980b [visited on 30 October 2020].

Heller, M., 2015. *Bóg i geometria: gdy przestrzeń była Bogiem*. Kraków: Copernicus Center Press.

Hewitt, E. and Ross, K.A., 1963. *Abstract Harmonic Analysis. Vol. 1: Structure of Topological Groups, Integration Theory, Group Representations*, *Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen mit Besonderer Berücksichtigung der Anwendungsgebiete* Bd. 115. Berlin: Springer.

Jackowski, S., 2015. Samuel Eilenberg – wielki matematyk z Warszawy. *Wiadomości Matematyczne* [Online], 50(1), pp.21–43. Available at: https://doi.org/10.14708/wm.v50i1.651 [visited on 1 November 2020].

Kan, D.M., 1958. Adjoint functors. *Transactions of the American Mathematical Society* [Online], 87(2), pp.294–294. Available at: https://doi.org/10.1090/S0002-9947-1958-0131451-0 [visited on 1 November 2020].

Kelley, J.L., 1955. *General Topology*, *University Series in Higher Mathematics*. Princeton, N.J.: D. Van Nostrand Co.

Knorr, W.R., 1975. *The Evolution of the Euclidean Elements: A Study of the Theory of Incommensurable Magnitudes and Its Significance for Early Greek Geometry*, *Synthese Historical Library* vol. 15. Dordrecht: D. Reidel.

Kotarbiński, T., 1929. *Elementy teorji poznania, logiki formalnej i metodologji nauk*. Lwów: Wydawn. Zakładu Narodowego im. Ossolińskich.

Krömer, R., 2007. *Tool and Object: A History and Philosophy of Category Theory*, *Science Networks. Historical Studies* vol. 32. Basel [etc.]: Birkhäuser.

Król, Z., 2019. Category theory and philosophy. In: M. Kuś and B. Skowron, eds. *Category Theory in Physics, Mathematics, and Philosophy* [Online], *Springer Proceedings in Physics* 235. Springer International Publishing, pp.21–32. Available at: https://doi.org/10.1007/978-3-030-30896-4 [visited on 1 November 2020].

Kuratowski, K., 1933. *Topologie. 1, Espaces Métrisables, Espaces Complets*, *Monografje Matematyczne* t. 3. Warszawa; Lwów: s.n.

Lakatos, I., 1976. *Proofs and Refutations: The Logic of Mathematical Discovery*. Ed. by J. Worrall and E. Zahar. Cambridge [etc.]: Cambridge University Press.

Lawvere, F.W., 1963. Functorial semantics of algebraic theories. *Proceedings of the National Academy of Sciences of U.S.A.* [Online], 50, pp.869–872. Available at: https://www.pnas.org/content/pnas/50/5/869.full.pdf [visited on 29 October 2020].

Lawvere, F.W., 1964. An elementary theory of the category of sets. *Proceedings of the National Academy of Sciences of the U.S.A.* [Online], 52(6), pp.1506–1511. Available at: https://doi.org/10.1073/pnas.52.6.1506 [visited on 1 November 2020].

Lawvere, F.W., ed., 1972. *Toposes, Algebraic Geometry and Logic*, *Lecture Notes in Mathematics* 274. Berlin; Heidelberg; New York: Springer.

Łukasiewicz, J., 1910. Über den Satz des Widerspruchs bei Aristoteles. *Bulletin International de l'Académie des Sciences de Cracovie*, 1-2, pp.15–38.

Mac Lane, S., 1950. Duality for groups. *Bulletin of the American Mathematical Society* [Online], 56(6), pp.485–516. Available at: https://projecteuclid.org/euclid.bams/1183515045 [visited on 29 October 2020].

Mac Lane, S., 1971. *Categories for the Working Mathematician*, *Graduate Texts in Mathematics* 5. New York: Springer-Verlag.

Mac Lane, S., 1986. *Mathematics: Form and Function*. New York [etc.]: Springer-Verlag.

Mac Lane, S., 1988. Concepts and Categories in Perspective. In: P. Duren, R. Askey and U. Merzbach, eds. *A Century of Mathematics in America, Part I* [Online], *History of Mathematics* 1. Providence, R.I.: American Mathematical Society, pp.323–365. Available at: https://www.ams.org/publicoutreach/math-history/hmath1-maclane25.pdf [visited on 29 October 2020].

Mac Lane, S., 2002. Samuel Eilenberg and categories. *Journal of Pure and Applied Algebra* [Online], 168(2-3), pp.127–131. Available at: https://doi.org/10.1016/S0022-4049(01)00092-5 [visited on 29 October 2020].

Mitchell, B., 1965. *Theory of Categories*, *Pure and Applied Mathematics. A Series of Monographs and Textbooks* 17. New York; London: Academic Press.

Piaget, J. and García, R.V., 1989. *Psychogenesis and the history of science* (H. Feider, Trans.). New York: Columbia University Press.

Samuel, P., 1948. On universal mappings and free topological groups. *Bulletin of the American Mathematical Society* [Online], 54(6), pp.591–598. Available at: https://projecteuclid.org/euclid.bams/1183512049 [visited on 29 October 2020].

Semadeni, Z., 2015. Transgresje poznawcze jako istotna cecha rozwoju matematyki. In: R. Murawski, ed. *Filozofia matematyki i informatyki*. Kraków: Copernicus Center Press, pp.65–90.

Semadeni, Z., 2018. Platonizujący konceptualizm w matematyce. In: R. Murawski and J. Woleński, eds. *Problemy filozofii matematyki i informatyki*. Poznań: Wydawnictwo Naukowe UAM, pp.77–95.

Shannon, C.E., 1936. *A Symbolic Analysis of Relay and Switching Circuits* [Online]. Thesis. Massachusetts Institute of Technology. Available at: https://dspace.mit.edu/handle/1721.1/11173 [visited on 2 November 2020].

Skowron, B., manuscript. *Was Saunders Mac Lane a platonist?* [unpublished].

Tall, D., 2013. *How Humans Learn to Think Mathematically: Exploring the Three Worlds of Mathematics*, *Learning in Doing: social, cognitive, and computational perspectives*. Cambridge: Cambridge University Press.

Tarski, A., 1933. *Pojęcie prawdy w językach nauk dedukcyjnych*, *Prace Towarzystwa Naukowego Warszawskiego. Wydział III: Nauk Matematyczno-Fizycznych* 34. Warszawa: nakładem Towarzystwa Naukowego Warszawskiego, z zasiłku Ministerstwa Wyznań Religijnych i Oświecenia Publicznego.

Thurston, W.P., 1990. Mathematical Education. *Notices of the Amer- ican Mathematical Society*, 37(7), pp.844–850.

Waerden, B.L.v.d., 1930. *Moderne Algebra. T. 1*, *Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen mit Besonderer Berücksichtigung der Anwendungsgebiete* Bd. 33. Berlin: Verlag von Julius Springer.

Youschkevitch, A.P., 1976. The concept of function up to the middle of the 19th century. *Archive for History of Exact Sciences* [Online], 16(1), pp.37–85. Available at: https://doi.org/10.1007/BF00348305 [visited on 2 November 2020].

Abstract logical structuralism[*]

Jean-Pierre Marquis
Department of Philosophy, University of Montreal

Abstract

Structuralism has recently moved center stage in philosophy of mathematics. One of the issues discussed is the underlying logic of mathematical structuralism. In this paper, I want to look at the dual question, namely the underlying structures of logic. Indeed, from a mathematical structuralist standpoint, it makes perfect sense to try to identify the abstract structures underlying logic. We claim that one answer to this question is provided by categorical logic. In fact, we claim that the latter can be seen—and probably should be seen—as being a structuralist approach to logic and it is from this angle that categorical logic is best understood.

Keywords

philosophy, logic, structuralism, categorical logic.

1. Introduction

In their recent booklet *Mathematical Structuralism*, Hellman and Shapiro (2019), give a list of eight criteria to evaluate var-

---

ious strands of mathematical structuralism. The very first criterion is the background logic used to express the version of structuralism examined: is it first-order, second-order, higher-order, modal? The claim made by Hellman & Shapiro is that the logic has a direct impact on the philosophical thesis: for instance, the existence of non-standard models in first-order logic seems to affect some of the claims made by a mathematical structuralist. Be that as it may, it is assumed that any variant of mathematical structuralism is based on an underlying logic and the chances are that a change of logic will modify the type of structuralism defended. One then has to weigh the pros and the cons of adopting a specific logic to defend a kind of mathematical structuralism.

This is all well and good, and I don't intend to discuss this assumption in this paper. Let me rather turn the question of the underlying logic on its head. A mathematical structuralist not only believes that *pure* mathematics is about abstract structures, but also that *pure* logic can be seen that way. In other words, a mathematical structuralist aspires to know what are the abstract structures underlying a given logic. Can logic itself be given a structuralist treatment? In other words, is it possible to identify the *abstract mathematical structures* from which the standard logical systems can be derived? In the same way that the natural numbers or the real numbers can be seen as specific structures arising from the combination of particular structures and properties, specific logical systems, e.g. intuitionistic first-order logic, classical first-order logic would result from the combination of particular abstract structures and properties. And important logical theorems—completeness theorems, incompleteness theorems, definability, undecidability, etc.—would be instances of abstract structural features, together with, presumably, singular properties.

Our main claim in this paper is that categorical logic is one way to give precise answers to these questions.[1] Indeed, this is what categorical logic is all about: it reveals the abstract mathematical structures of logic and it relates them to other abstract mathematical structures, revealing yet other structural features. It also shows how the main results of logic are a combination of abstract structural facts and specific properties of logical systems. It is in this sense that one can say that *pure* category theory *is* logic.[2] This is one of the main *points* of categorical logic. Thus, when mathematical structuralism is developed within the metamathematical framework of category theory, it is possible to give a positive answer to our new challenge. And as far as I know, it is at the time the only metamathematical framework that allows us to do it in such generality.

Of course, we are not claiming that the abstract analysis is *superior* in a strong sense to the other presentations of logical systems. It brings a certain perspective, a certain understanding and opens up certain connections that are otherwise unavailable. Furthermore, in as much as a logic is applied, be it in foundational studies or in computer science, one also wants to look at it from a different point of view. But it has to be perfectly clear that these are completely different issues. We are now positioning ourselves in a structuralist framework, and will therefore ignore the aspects of a logic that become prevalent when one look at it for its applications. Ours is a *philosophical* goal, not a technical one.

---

[1] In his (1996), often quoted as representing the categorical perspective on mathematical structuralism, Awodey did include a presentation of the basic features of logic from a categorical point of view, more specifically the internal logic of a topos. Given how he treated logic in his paper, none of his critics saw that he was including logic itself as an object of mathematical structuralism (see, for instance, Hellman, 2001; 2003). We approach the issues from a different angle.

[2] Which is not to say that it is *all* of logic. That is *not* the claim.

It goes without saying that categorical logic, seen as a search for the abstract structures of logic, did not come out of the blue. It was part of a larger mathematical movement, namely the structuralist movement in mathematics, which culminated in Bourbaki's *Éléments de mathématique* (for more on the prehistory and the history of mathematical structuralism, see Corry, 2004; Reck and Schiemer, 2020). It will therefore be worth our while to look briefly at the birth of categorical logic in the 1960s and early 1970s. We will merely indicate the major landmarks of the story, to see how indeed categorical logic was, right from the start, understood as being a part of that movement. We will then briefly look at some of the basic abstract structures that correspond to logic and then how some of the main theorems follow from features of abstract structures. At this point, to call the latter 'mathematical' or 'logical' is a matter of choice.

In the end, this paper should be taken as a challenge. We claim that one should add a ninth criterion of Hellman & Shapiro's list: how does a given form of structuralism treat logic itself? Does it reveal the structures of logic? Are these structures related to other mathematical structures in a natural manner? Are these structures on the same level as the other fundamental structures of mathematics? One could, and we suggest that one should, evaluate different forms of mathematical structuralism according to this criterion.

2. The abstract structures of logic: setting the stage

Although category, functors and natural transformations were introduced explicitly in 1945 by Eilenberg and Mac Lane, category *theory* came to maturity only fifteen years later, thus at the beginning of the

1960s (for more on this history, see, for instance Landry and Marquis, 2005; Krömer, 2007; Marquis, 2009; Rodin, 2014). At that point, the major concepts of category theory were in place, e.g. adjoint functors, representable functors, constructions on categories, including the crucial construction of functor categories, abstract categories like additive categories, abelian categories, tensor or monoidal categories, etc.[3] But it is only in the early 1960s that the connection with logic and the foundations of mathematics was made and it was mostly the work of one person, namely Bill Lawvere. We will not present Lawvere's early work here, for our goal is not to survey the history of the subject, but rather to make a conceptual point.[4]

Of course, by that time, connections between classical propositional logic and Boolean algebras, intuitionistic propositional logic and Heyting algebras, as well as others were known. Already in (Birkhoff, 1940), the main relations are presented.[5] The first links between logic and lattice theory were restricted to propositional logic. The chase for the identification of the abstract structures corresponding to first-order, higher order logics, non-classical logics, as well as algebraic proofs of the main theorems of logic was taken up by, among others, Tarski and his school, Halmos, and the Polish school,

---

[3] The *locus classicus* of the time is (Mac Lane, 1965).

[4] See (Lawvere, 1963; 1966; 1967; 1969a; 1970; 1971). He was rapidly joined by (Freyd, 1966; Linton, 1966; Lambek, 1968a,b; 1969; 1972). For a detailed history, see (Marquis and Reyes, 2012).

[5] We could argue that already for propositional logics, the structures arising from the logical systems naturally live in categories, e.g. the category of distributive lattices, the category of Boolean algebras, the category of Heyting algebras, the category of S4-algebras, etc., and that the main results are also naturally expressed in categorical language. It is in fact an important point to make, for once the higher order logical systems find their place in this landscape, the fact that we move to bicategories to express and prove the results of first-order logic is easily understood. But we will not dwell on that.

e.g. Łoś, Mostowski, Rasiowa and Sikorski.[6] The main contenders to the title of abstract algebraic structures corresponding to first order logic at the time were cylindric algebras and polyadic algebras. The main problem, so to speak, were the quantifiers $\forall$ and $\exists$. It was not a technical problem. They were treated properly in each case. But their treatment, as algebraic operators, was somewhat *ad hoc*, in the sense that they did not arise as an instance of abstract operators in an algebraic context. The resulting algebras were therefore somewhat *ad hoc* also, in as much as they did not belong to a family of abstract structures that arose naturally in other mathematical contexts. In other words, the abstraction proposed via the concepts of cylindric algebras or polyadic algebras were not genuine mathematical abstractions, for they were merely the algebraic transcription of the quantifiers and solely of the latter.[7] This is in stark contrast with the case of propositional logic, where the abstract algebraic structures capturing the logic and its main properties have instances in a variety of completely different mathematical fields. Distributive lattices, Boolean algebras, Heyting algebras, etc., are genuine mathematical abstractions.

It therefore came as a complete surprise that the quantifiers, as well as the propositional connectives, could be seen as being instances of adjoint functors on very simple categories, the concept of adjoint functors being one of the core concepts of category theory, introduced by Kan in the context of algebraic topology in 1958. This was one of Lawvere's crucial observations. Three additional crucial facts

---

[6] The list of references is long, but clearly indicates that it was a very active area of research in the 1950s as well as in the 1960s. See, for instance, (McKinsey and Tarski, 1944; 1946; 1948; Jónsson and Tarski, 1951; 1952; Henkin and Tarski, 1961; Henkin, Monk and Tarski, 1971; Halmos, 1954; 1956a,b,c,d; 1962; Mostowski, 1949; Rasiowa, 1951; 1955; Rasiowa and Sikorski, 1950; 1953; 1955; Rasiowa and Sikorski, 1963).

[7] The reader might wonder what we mean by "genuine mathematical abstraction". We refer her to (Marquis, 2015; 2016).

had been established by Lawvere in his Ph.D. thesis. First, Lawvere
showed how algebraic theories, in the standard logical sense of that
expression, could themselves be captured by specific categories. Sec-
ond, the models of algebraic theories, again in the standard logical
sense of that expression, could be described in the language of cate-
gories, functors and natural transformations. Third, the classical links
between the syntax and the semantics of these theories could receive
an adequate categorical treatment, and at the core of this treatment
one finds adjointness. Thus, it seemed possible to put all the structures
of logic in the theoretical framework of categories, the latter being,
of course, an abstraction of a central fact of modern mathematics.
The overall plan was presented by Lawevere in (1969a). That paper
articulates in very broad strokes how the syntax, the semantics and
their relationships could be captured in a categorical framework. Here
are, in a nutshell, the main ingredients of this ambitious program.

A few words about the philosophical framework underlying Law-
vere's program are in order. Lawvere identifies two fundamental as-
pects to all of mathematics, namely the formal and the conceptual,
roughly the manipulation of symbols, on the one hand, and what these
symbols refer to, their content. Lawvere is aware of the work done
in algebraic logic when he writes his paper. Indeed, he refers to it
explicitly in the opening section: "[...] Foundations may conceptual-
ize the formal aspect of mathematics, leading to Boolean algebras,
cylindric and polyadic algebras [...]" (Lawvere, 1969a, p.281). He is
also presenting the introduction of categories in the analysis of logic
as a structural approach, based on the notion of adjoints: "Specifi-
cally, we describe [...] the notion of cartesian closed category, which
appears to be the appropriate abstract structure for making explicit

[...]. The structure of a cartesian closed category is entirely given by adjointness, as is the structure of a 'hyperdoctrine', which includes quantifiers as well." (Lawvere, 1969a, p.281)

We will now focus on the final section of the paper, which is really programmatic. In this last section, Lawvere is describing what he himself characterizes as a globalized Galois connection, and indeed, it also contains the main ideas that have driven the development of duality theory in a categorical framework. But as far as logic is concerned, we are offered the following picture.

1. Logical operations should arise from an elementary context as adjoint operations. From a categorical point of view, a logical doctrine, that is an abstract mathematical structure encapsulating a logical framework, should be given by adjoint functors.[8]

2. A theory **T**, in the standard logical sense of the term, should be constructed as a category, in the same way that a propositional theory in classical logic can be turned into a Boolean algebra via the Lindenbaum-Tarski construction. A theory **T**, seen as a special type of category, is conceived by Lawvere as being the invariant notion of a theory, that is, independent of a choice of primitive symbols or specific axioms. We thus have abstract mathematical structures corresponding to the formal.

3. The models of **T** should form a category. Lawvere, having himself developed the case of algebraic theories earlier in his

---

[8] We have to point out that categorical logic does cover logical situations in which certain logical operations are not given by adjoint operations. Although they do not constitute logical *doctrines* in the sense of Lawvere, they are part and parcel of categorical logic. We should also mention that the syntactic aspects of logic, which are pushed in the background in Lawvere's early work, occupy nonetheless an important part of categorical logic, for instance via the notion of sketch, introduced by Ehresmann and his school or the various formal graphical languages developed mostly in the context of monoidal categories.

thesis, generalizes from his work and proposes to make the category of models of a theory a functor category. We will be more specific in later sections. These provide the abstract mathematical structures emerging from the conceptual.

4. Last, but certainly not least, since everything is a category now, the links between the formal and the conceptual should also be given by (adjoint) functors, and we have yet again a new type of abstract mathematical structure, in Lawvere's mind a globalized Galois connection, arising from that situation.

Lawvere was of course guided by his own work on algebraic theory, but also explicitly by Grothendieck's work in algebraic geometry. As I said, at the time, it was a program, some would say a vision. It became a reality in the following decade and is still the basis of important developments in the field.

We have to explain why we claim that we are then in a structuralist framework. It is not only because we are in fact dealing with abstract mathematical structures—this is of course a necessary step— but these abstract mathematical structures can be characterized up to 'isomorphism', where the latter notion is derived from the abstract structures themselves. Each and every one of these abstract structures comes with a notion of homomorphism and, in particular, a notion of 'isomorphism'. Therefore, it becomes possible to develop logic with respect to the structuralist principle: if X is a structure of a given kind, and it has property $P(X)$, then given any other structure Y of the same kind such that X is isomorphic to Y, $X \simeq Y$, in the appropriate sense of isomorphism, we should be able to prove that $P(Y)$. As it can be seen, the key component of this desideratum is the appropriate sense of isomorphism. In some cases, we are dealing with the usual set-theoretical sense of isomorphism, in others, it becomes an equivalence of categories and in still others, it is a 2-categorical equiv-

alence. It is the very possibility of having the appropriate sense of isomorphism that allows us to claim that we are dealing with *abstract* mathematical structures.

I want to emphasize again, at this point, that I am not claiming that categorical logic, as I will present it succinctly below, is the only possible answer for a structuralist nor is it the final answer. But it is one clear answer and one of the very few that provides a comprehensive answer.

3. Categories as abstract logical structures

At this stage, we would have to give a long list of definitions and examples to illustrate how certain categories correspond to the abstract mathematical structures of certain logics. I will assume that the reader knows the notions of category, functor, natural transformation, adjoint functors, etc., for otherwise this paper would be terribly long and boring. We will try to put some flesh on Lawvere's program described in the foregoing section. We assume, however, that the description of the logical connectives, including the quantifiers, as adjoint functors is understood[9]. We will sketch how the other three steps are filled[10]. A warning is necessary. Each following section would require a careful and systematic exposition to be ultimately convincing. It is impossible to do justice to the field in such a short paper. We will provide a more detailed presentation in the next section only and merely gloss over the abstract mathematical structures involved in the other sections. We

---

[9] Mac Lane's textbook, (1998), is still a good reference. All the standard concepts and examples can also be found in (Riehl, 2017).

[10] A more detailed presentation can be found in (Marquis, 2009), chapter 6. Our exposition here is adapted from the latter, but the philosophical point is different and therefore the presentation found there might not be optimal for our present purpose.

apologize for the opacity that might result from the lack of details and clarifications, but a much longer paper would be required to present and motivate adequately the main mathematical ideas involved.

A theory as a category and a category as a theory

Let us start with the goal that was in the minds of logicians and mathematicians in the 1950s, that is finding the appropriate abstract mathematical structure that correspond to a first-order theory.

Let us fix the logical context first. We consider formal systems with many sorts, which is a simple generalization of the standard first-order logic which is done over a single sort. A *similarity type* or *alphabet* A, often called a language in the literature, is given by:

1. A collection of sorts $S_1, S_2, S_3, \ldots$;[11]
2. A collection of relation symbols $R_1, R_2, R_3, \ldots$, each of which is given with the sorts of its arguments;
3. A collection of function symbols $f_1, f_2, f_3, \ldots$ each of which is given with the sorts of its arguments and the sort of its target; we denote a function symbol f as $f \colon S_1 \times \cdots \times S_n \to S$ if f takes n arguments of sorts $S_1, \ldots, S_n$ respectively to a value of sort S.
4. A collection of constants $c_1, c_2, c_3, \ldots$ each with a specified sort; we denote a constant c by $c \colon 1 \to S_i$ to indicate that the constant c is of sort S_i.

---

[11] Or types, if you prefer. We are dealing here with first-order logic. We will say a few words about type theory later. It does not affect our basic general point. Type theories can also be analyzed as instances of abstract mathematical structures.

This is the standard definition extended to a many-sorted context. To obtain a *formal system* L_A in the alphabet A, we add the usual elements:

1. Each sort S_i comes with infinitely many variables $x_1, x_2, x_3, \ldots$; we write $x \colon S_i$ to indicate that the variable x is of sort S_i;

2. Each sort has an equality relation $=_S$; notice immediately that this means that equality is not treated as a universal or purely logical relation and that in the interpretation, whatever will correspond to a sort will have to come equipped with a criterion of identity or equality for *its* objects;

3. The usual logical symbols and two propositional constants, $\top$ and $\bot$;

4. The usual deductive machinery for a predicate logic (say, intuitionistic predicate logic; if any other deductive procedures are assumed, they are made explicit).

Terms (of a given sort) and atomic formulas are defined as usual.

Here is the first original result obtained by searching for abstract mathematical structures corresponding to theories in a given logic. Some fragments of first-order logic and some extensions of first-order logic turn out to have significant properties, properties that would not have been identified otherwise[12]. We can immediately identify the following fragments.

1. A formula φ is said to be *regular* if it is obtained from atomic formulas by applying finite conjunction and existential quantification.

---

[12] We will not be exhaustive here. There are other infinitary fragments that are important, but we will ignore them.

2. A formula φ is said to be *coherent* if it is obtained from atomic formulas by applying finite conjunction, disjunction and existential quantification.

3. A formula is said to be *geometric* if it is obtained from atomic formulas by applying *finite* conjunction, *finite* existential quantification and *infinite* disjunction.

Intuitionistic formulas and classical (or Boolean) formulas are defined in the obvious manner. We point out immediately that many metalogical results about intuitionistic and classical logic follow directly from results about the foregoing fragments. This is a genuine discovery that could not have been foreseen beforehand.

An *implication* of regular (resp. coherent, geometric, etc.) formulas φ and ψ has the form

$$\forall x_1 \ldots \forall x_n (\varphi(x_1, \ldots, x_n) \Rightarrow \psi(x_1, \ldots, x_n))$$

where $\varphi(x_1, \ldots, x_n)$ and $\psi(x_1, \ldots, x_n)$ are regular (resp. coherent, geometric, etc.) formulas. A theory $\mathbf{T}$ in the given language L is said to be a *regular theory* (resp. coherent, geometric, etc.) if all its axioms are implications of regular (resp. coherent, geometric, etc.) formulas. Many mathematical theories can be expressed in the form of regular theories, or coherent theories, etc.

Apart from the fact that we have assumed a many sorted context and cut the fragments of first-order logic in ways that might seem arbitrary, the foregoing presentation is squarely in a standard logical context. We now move to the abstract mathematical context.

Given a theory $\mathbf{T}$ in one of the foregoing languages, we construct a category denoted by $[\mathbf{T}]$, out of it. The latter is sometimes called the category of concepts, sometimes the syntactic category, and it is basically an extension of the Lindenbaum-Tarski construction, but for

theories expressed in first-order logic (and others, as the reader can now guess). It is constructed from the language and the axioms of $\mathbf{T}$ as follows.[13]

Remember that we are constructing a category, thus a web of objects connected by morphisms. To get the objects of this category, we start with *formal sets* $[\mathbf{x}; \varphi(\mathbf{x})]$, where $\mathbf{x}$ denotes a n-tuple of distinct variables containing all free variables of φ and φ is a formula of the underlying formal system L. Two such formal sets, $[\mathbf{x}; \varphi(\mathbf{x})]$ and $[\mathbf{y}; \varphi(\mathbf{y})]$ are equivalent if one is the alphabetic variant of the other, that is if $\mathbf{x}$ and $\mathbf{y}$ have the same length and sorts and $\varphi(\mathbf{y})$ is obtained from $\varphi(\mathbf{x})$ by substituting $\mathbf{y}$ for $\mathbf{x}$ (and changing bound variables if necessary). This is an equivalence relation and it is therefore possible to consider equivalence classes of such formal sets. An object of the category of concepts $[\mathbf{T}]$ is such an equivalence class of formal sets $[\mathbf{x}; \varphi(\mathbf{x})]$, where φ is a formula of the formal system L. The objects of $[\mathbf{T}]$ are the *equivalence classes of these formal sets*, for all formulas of L. Notice this last important point: we take *all* formulas of the language, not only those which appear in $\mathbf{T}$. Thus, in a sense, the space of objects is the collection of all possible properties and sentences expressible in that language, thus all possible theories in the given formal system. No logical relationship is considered at this stage. The next step introduces the structure corresponding to the structure of that particular theory $\mathbf{T}$. This is just as one would expect in a categorical framework: the structure of $\mathbf{T}$ is captured by the morphisms we will define and the properties resulting therefrom.

It is easier to motivate the definition of morphism with an eye on the semantics, although the properties of the morphisms, e.g., that they form a category, have to be proved with the syntactical

---

[13] See also (Makkai and Reyes, 1977, chap. 8) or (Mac Lane and Moerdijk, 1994, chap. X, § 5) for more details and proofs or again (Johnstone, 2002).

features of the theory (unless one has a completeness theorem at hand). The basic idea is this: a functor from $[\mathbf{T}]$ to $\mathbf{Set}$ should transform the objects of $[\mathbf{T}]$ into *genuine sets* and the morphisms of $[\mathbf{T}]$ into *genuine functions* automatically. These functions should be functions that are definable in $\mathbf{T}$—i.e., for which we can prove in $\mathbf{T}$ that they are indeed functions. Furthermore, $[\mathbf{T}]$ should contain all of them. By sending a formal set $[\mathbf{x}; \varphi(\mathbf{x})]$ to the set of n-tuples satisfying the formula, i.e. $\{(x_1, \ldots, x_n) \mid \varphi(\mathbf{x})\}$, a morphism from $[\mathbf{x}; \varphi(\mathbf{x})]$ to $[\mathbf{y}; \psi(\mathbf{y})]$ should become a genuine function between genuine sets $\{(x_1, \ldots, x_n) \mid \varphi(\mathbf{x})\}$ and $\{(y_1, \ldots, y_m) \mid \psi(\mathbf{y})\}$ respectively. Such a morphism should simply be given by a formula of the theory $\mathbf{T}$ that defines such a function, that is a formula $\theta(\mathbf{x}, \mathbf{y})$ of $\mathbf{T}$ that is provably functional. The only trick in the construction is to construct a morphism between two (equivalence classes of) formal sets $[\mathbf{x}; \varphi(\mathbf{x})]$ and $[\mathbf{y}; \psi(\mathbf{y})]$ in such a way that, when interpreted, it yields the *graph* of the function, in the standard set-theoretical sense of that expression, between the actual sets $\{(x_1, \ldots, x_n) \mid \varphi(\mathbf{x})\}$ and $\{(y_1, \ldots, y_m) \mid \psi(\mathbf{y})\}$. Thus, all definable functions in $\mathbf{T}$ will be represented by a morphism in $[\mathbf{T}]$.

Formally, consider a triple $(\mathbf{x}, \mathbf{y}, \gamma)$, where $\mathbf{x}$ and $\mathbf{y}$ are disjoint tuples of distinct variables and γ is a formula with free variables possibly among $\mathbf{x}$ and $\mathbf{y}$. Such a triple defines a *formal function* if the following formulas are provable:

$$\mathbf{T} \vdash \forall \mathbf{x} \forall \mathbf{y} (\gamma(\mathbf{x}, \mathbf{y}) \Rightarrow (\varphi(\mathbf{x}) \wedge \psi(\mathbf{y})));$$
$$\mathbf{T} \vdash \forall \mathbf{x} (\varphi(\mathbf{x}) \Rightarrow \exists \mathbf{y} (\gamma(\mathbf{x}, \mathbf{y})));$$
$$\mathbf{T} \vdash \forall \mathbf{x} \forall \mathbf{y} \forall \mathbf{y}' (\gamma(\mathbf{x}, \mathbf{y}) \wedge \gamma(\mathbf{x}, \mathbf{y}') \Rightarrow \mathbf{y} = \mathbf{y}');$$

where we have used some obvious abbreviations. The underlying motivation should be clear: these formulas will be true in any interpretation of $\mathbf{T}$ in which γ is indeed a morphism.

We now define an equivalence relation $(\mathbf{x}, \mathbf{y}, \gamma) \sim (\mathbf{u}, \mathbf{v}, \eta)$ if

$$\mathbf{T} \vdash \forall \mathbf{x} \forall \mathbf{y} (\gamma \Leftrightarrow (\eta(\mathbf{x}/\mathbf{u}, \mathbf{y}/\mathbf{v}))).$$

The equivalence relation guarantees that for every model M of $\mathbf{T}$, the functions corresponding to γ and to η will coincide. We can now stipulate that a formal function is an *equivalence class* of the foregoing equivalence relation. Given a representative $(\mathbf{x}, \mathbf{y}, \gamma)$ of such an equivalence class, we denote the equivalence class containing it by $\langle \mathbf{x} \mapsto \mathbf{y} \colon \gamma \rangle$. Thus, a *formal morphism* in $[\mathbf{T}]$ is denoted by:

$$\langle \mathbf{x} \mapsto \mathbf{y} \rangle \colon [\mathbf{x} \colon \varphi] \to [\mathbf{y} \colon \psi].$$

We need two more ingredients to get to a category. Firstly, for each formal set $[\mathbf{x} \colon \varphi(\mathbf{x})]$, the identity morphism is provided by the formal morphism $\langle \mathbf{x} \mapsto \mathbf{y} \colon (\mathbf{x} = \mathbf{y}) \wedge \varphi \rangle$. Secondly, given two formal morphisms $\langle \mathbf{x} \mapsto \mathbf{y} \colon \gamma \rangle \colon [\mathbf{x} \colon \varphi] \to [\mathbf{y} \colon \psi]$ and $\langle \mathbf{y} \mapsto \mathbf{z} \colon \eta \rangle \colon [\mathbf{y} \colon \psi] \to [\mathbf{z} \colon \zeta]$, their composition is defined by the formal morphism $\langle \mathbf{x} \mapsto \mathbf{z} \colon \mu \rangle \colon [\mathbf{x} \colon \varphi] \to [\mathbf{z} \colon \zeta]$ where $\mu = \exists \mathbf{y} (\gamma \wedge \eta)$. These two definitions satisfy the usual requirements of a category. Thus, $[\mathbf{T}]$ is a category and we have an abstract mathematical structure corresponding to a given theory.

Notice that $[\mathbf{T}]$ is *not* a category of structured sets and structure-preserving functions! A lot of information about $\mathbf{T}$ is lost when all we have at our disposal is $[\mathbf{T}]$. It is, for instance, impossible to know which atomic formulas are involved in specific formal sets or what were the primitive symbols of the language $L_{\mathbf{T}}$. Furthermore, two different theories $\mathbf{T}$ and $\mathbf{T}'$ can very well yield isomorphic categories of concepts, thus essentially the same category. We are squarely in a structuralist framework: the category $[\mathbf{T}]$ is given up to an isomor-

phism of categories.[14] As we have seen, the syntactic logical opera-
tions, i.e., quantifiers and connectives, become categorical operations
in the category and this part of the structure is not lost. Again, moving
from a theory **T** to its category of concepts [**T**] is an *abstraction*: the
specific formulas with specific variables are abstracted from when
we move to the equivalence classes. The category of concepts is the
category of all definable sets and functions of a theory **T**. Thus, in a
sense, it contains all the formally expressible concepts of **T**, whence
its name.

When we start with, for example, a regular theory, the foregoing
construction yields a category with additional structures and proper-
ties. For instance, it is automatically a category with finite limits.

T and [**T**] are interchangeable in the following sense: for, given a
(small) category $\mathcal{C}$, at least with finite limits, it is possible to associate
or construct the language $L_{\mathcal{C}}$ of $\mathcal{C}$ as follows. We first have to identify
the alphabet of $L_{\mathcal{C}}$. The sorts are given by the objects $X, Y, Z, \ldots$ of $\mathcal{C}$.
Every morphism $f \colon X \to Y$ of $\mathcal{C}$ becomes a function symbol of $L_{\mathcal{C}}$.
(In particular, a constant $c \colon 1 \to X$ is seen as 0-ary function symbols.)
This is called the *canonical language* of $\mathcal{C}$. Notice that $L_{\mathcal{C}}$ is obtained
as if we had taken $\mathcal{C}$ and destroyed its categorical structure, retaining
only the symbols, and keeping in mind that function symbols are
sorted. It is possible to extend this language to reflect the structure of
$\mathcal{C}$ more closely. Although subobjects of $\mathcal{C}$ can be denoted naturally by
formulas of $L_{\mathcal{C}}$, it is possible to introduce relation symbols for each
subobject $R(x_1, \ldots, x_n) \rightarrowtail X_1 \times \cdots \times X_n$ and n-ary function symbol
for morphisms $f \colon X_1 \times \cdots \times X_n \to X$. This is called the *extended*
canonical languageof $\mathcal{C}$ (see Makkai and Reyes, 1977, chap. 2, sec. 4).
In order to get the *internal theory* $\mathbf{T}_{\mathcal{C}}$ of $\mathcal{C}$ in its canonical language,

---

[14] Notice that we are talking about isomorphism here and not an equivalence of
categories.

$\mathcal{C}$ has to have more structure than just finite limits. It has to be at least a *regular* category, which we will define shortly. In this case, it is possible to give a list of *regular* axioms $\Sigma_{\mathcal{C}}$, that is a set of regular formulas, and prove that $\mathbf{T}_{\mathcal{C}}$ is sound in $\mathcal{C}$ (see Makkai and Reyes, 1977, chap. 3). The internal theory $\mathbf{T}_{\mathcal{C}}$ is related to $\mathcal{C}$ by two expected properties:

1. There is a canonical interpretation G of $\mathbf{T}_{\mathcal{C}}$ in $\mathcal{C}$;
2. For any model M of $\mathbf{T}_{\mathcal{C}}$ in a regular category $\mathcal{D}$, in any reasonable sense of the term 'model', there is a unique regular functor $I\colon \mathcal{C} \to \mathcal{D}$ such that I applied to G is equal to M.

It is of course possible to complete the circle: starting with a regular category $\mathcal{C}$, construct its internal theory $\mathbf{T}_{\mathcal{C}}$ and then move to its category of concepts $[\mathbf{T}_{\mathcal{C}}]$. How are $\mathcal{C}$ and $[\mathbf{T}_{\mathcal{C}}]$ related? They are in fact *equivalent as categories*, which means that they share the same categorical properties. In other words, as abstract mathematical structures, they are indistinguishable. From this, it is possible to conclude a very important result that every (small) regular category is equivalent to a category of concepts for some theory $\mathbf{T}$.

The Architecture of Logical Theories

We now have sketched how a theory in a logical framework can be turned into an instance of an abstract mathematical structure. It should not come as a surprise to learn that a regular theory $\mathbf{T}$ (resp. a coherent, geometric, etc.) yield a specific kind of abstract category, namely a regular category (resp. a coherent, geometric, etc.). We will fill in some blanks here, for we want to emphasize the existence

of kinds of abstract mathematical structures. The existence of these abstract mathematical structures explains why we have introduced these fragments of first-order logic. Logic itself can be organized from the perspective of these structures. We thus get what we call the "architecture of logical theories" or the "architecture of logic".[15]

We will simply state the definition without explaining all the technical details. We refer the reader to the literature.

A *regular* category $\mathcal{C}$ is a category with finite limits[16], such that

1. Every morphism has a kernel pair;
2. Every kernel pair has a coequalizer;
3. The pullback of a regular epimorphism along any morphism exists and is a regular epimorphism.

This is the purely abstract mathematical structure corresponding to a regular theory **T**, but the abstract notion was not abstracted from that construction. It has an independent mathematical existence. The definition does not show automatically how regular logic can be interpreted in a regular category or that a regular theory yields, as its category of concepts, a regular category. But of course, in both cases, it does.

In a structuralist framework, one has to specify the criterion of identity for the abstract structures given. Thus, we first have to specify what a *regular functor* between regular categories is. Of course, it is a functor that preserves the appropriate structure. In this particular

---

[15] Again, we are being very selective, and our goal is not to be exhaustive. The picture is much more elaborate than what we are presenting here. This is but the tip of the iceberg.

[16] There are various equivalent definitions of regular categories in the literature. We are following (Borceux, 1994).

case, a functor $F : \mathcal{C} \to \mathcal{D}$ between regular categories is *regular* if it preserves finite limits and regular epimorphisms. The criterion of identity for regular categories is given by the notion of *equivalence* of regular categories, that is by a pair of regular functors $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{C} \to \mathcal{D}$ such that $G \circ F \simeq 1_{\mathcal{C}}$ and $F \circ G \simeq 1_{\mathcal{D}}$.

In the same vein, corresponding to coherent theories, we have:

A *coherent* category $\mathcal{C}$ is a regular category such that

1. Every subobject meet semilattice $S(X)$ is a lattice;
2. Each $f^* \colon S(Y) \to S(X)$ is a lattice homomorphism.

A coherent functor between coherent categories is a functor preserving the coherent structure, and the criterion of identity for coherent categories is extracted from that context.

And, of course, we can add more structure and properties to get other abstract mathematical structures. Let us simply mention a few more abstract structures that are directly related to logic.

A *Heyting category* $\mathcal{C}$ is a coherent category in which each $f^* \colon S(Y) \to S(X)$ has a right adjoint, denoted by $\forall_f$. The last condition is sufficient to entail that each $S(X)$ is a Heyting algebra, that f is a homomorphism of Heyting algebras and that the right adjoint is also stable under substitution. Heyting categories are common: for any small category $\mathcal{P}$, the functor category $\mathbf{Set}^{\mathcal{P}}$ is a Heyting category. They correspond to theories in intuitionistic predicate logic. Heyting functors are defined in the expected manner.

A *Boolean category* $\mathcal{C}$ is a coherent category such that every $S(X)$ is a Boolean algebra, i.e., every subobject has a complement. Boolean functors between Boolean categories are functors preserving the Boolean structure.

A *pretopos* $\mathcal{C}$ is a coherent category having (1) quotients of equivalence relations and (2) finite disjoint sums. Pretopos functors can be defined.[17] The notion of pretopos occupies a central place in the picture, since many of the important theorems about logic can be hooked to that notion. Finally, we have to mention at this stage the notion of Grothendieck topos, which surprisingly sits right at the center of the development of first-order logic and its many variants.[18] There is no need to go on for our purposes.

Let us immediately point out that the category **Set** is regular, coherent, Heyting, Boolean, a pretopos and a Grothendieck topos. And it is even more than just those.

There is an interesting and immediate application of the above constructions. In a classical logical framework, the notion of a interpretation or translation of one category into another one is delicate and complicated. Once we move from a theory **T** to its category of concepts [**T**], there is a very simple and direct way to define it. Indeed, a structure-preserving functor $I \colon [\mathbf{T}] \to [\mathbf{T}']$ between (small) categories of concepts is called an *interpretation* of [**T**] in [**T**']. (When [**T**] and [**T**'] have been constructed from theories, one can verify that it is a legitimate notion of interpretation (see Makkai and Reyes, 1977, chap. 7, p.196).

---

[17] As we have already mentioned, for many important results, it is enough to consider weaker functors between some of the categories involved, e.g. coherent functors, for they preserve, in these particular contexts, the additional structure.

[18] As already pointed out by Makkai & Reyes in (1977), Giraud's theorem can be interpreted as giving a *logical characterization* of the notion of a Grothendieck topos (see Makkai and Reyes, 1977, chapter 1, section 4; for a more recent look at the role of Grothendieck toposes in logic, see Caramello, 2018; see also Caramello, 2014).

The models of a theory as a category

We now have to consider how a theory $\mathbf{T}$ can be interpreted in a category $\mathcal{C}$. As it is often the case when facing such a situation, the easiest solution is to translate what one does in sets, but express it in the language of the category of sets and finally move to an arbitrary category with the adequate structure and properties. It can indeed be done and what we get is a genuine generalization of Tarski's notion of satisfaction or model. The classical notions of satisfaction, interpretation, model and truth transfer directly to this new context. However, instead of presenting the nuts and bolts of these definitions, we will jump immediately to the next step.

Since we have constructed $[\mathbf{T}]$ from $\mathbf{T}$, and since $[\mathbf{T}]$ is a category, we can look directly at the interpretations of the latter. We will illustrate the situation with a coherent theory $\mathbf{T}$, but starting with the (small) coherent category $[\mathbf{T}]$ constructed from it. A coherent functor $M \colon [\mathbf{T}] \to \mathbf{Set}$ is called a (set-)*model* of $[\mathbf{T}]$. Since $\mathbf{Set}$ is a coherent category, this makes sense.

It is natural to consider to category of all such models, that is the functor category $\mathrm{Mod}([\mathbf{T}], \mathbf{Set})$, the *category of all (set-)models* of $[\mathbf{T}]$. The objects of this category are the models of $[\mathbf{T}]$, that is coherent functors $M \colon [\mathbf{T}] \to \mathbf{Set}$, and the morphisms are the natural transformations between models $\eta \colon M_1 \to M_2$. These are the *homomorphisms* of models of $[\mathbf{T}]$ and they are the traditional model-theoretic structure-preserving functions between models. More generally, for any coherent category $\mathcal{C}$, the category $\mathrm{Mod}([\mathbf{T}], \mathcal{C})$ of models of $[\mathbf{T}]$ in $\mathcal{C}$ is defined in the same way. We therefore have a flexibility that was not available previously.

The category $\mathrm{Mod}([\mathbf{T}], \mathbf{Set})$ of models of $[\mathbf{T}]$ in $\mathbf{Set}$ *is* certainly an instance of an abstract mathematical structure.[19] It has, in fact, a lot of structure. It is, among other things, a Grothendieck topos, an important type of abstract mathematical structure.

A theory and its models: moving up the ladder

We have identified some of the abstract mathematical structures that arise from the traditional logical notions. We now have, on the one hand, abstract mathematical structures corresponding to what Lawvere referred to as the "formal", and, on the other hand, abstract mathematical structures corresponding to what Lawvere referred to as the "conceptual". Of course, these have to be connected and these connections constitute the core of classical logic.

These connections are themselves part of an abstract mathematical structure. Since $[\mathbf{T}]$ is a category—the formal side of mathematics—and $\mathrm{Mod}([\mathbf{T}], \mathbf{Set})$ is a category—the conceptual side of mathematics—, we can investigate the functors between them. But there is more. There are also functors between theories $[\mathbf{T}] \to [\mathbf{T}']$, functors between set-models of theories

$$\mathrm{Mod}([\mathbf{T}], \mathbf{Set}) \to \mathrm{Mod}([\mathbf{T}'], \mathbf{Set}),$$

functors between models of theories in different categories

$$\mathrm{Mod}([\mathbf{T}], \mathcal{C}) \to \mathrm{Mod}([\mathbf{T}], \mathcal{D}),$$

---

[19] This is also true when we take category different from $\mathbf{Set}$. But then, the structure of the resulting category depends directly on the structure of $\mathcal{C}$.

functors between all those and natural transformations between some of these functors![20] In fact, we are in a 2-category, which is a *genuinely new structure*. A 2-category is not merely a category with additional data. Thus, once again, we are in a realm of abstract mathematical structures and many of the results we are interested in will be consequences of this abstract mathematical structure *together with* some specific properties inherent in the situation we are dealing with.

Here are some questions that can now be investigated. The main point here is that these questions make perfect sense, they are entirely natural, whereas it is hard to imagine how they could have arisen outside this mathematical context.

1. Given an interpretation $I \colon [\mathbf{T}] \to [\mathbf{T}']$ between theories, one can transfer models of $\mathbf{T}'$ to models of $\mathbf{T}$ by composing with I, that is given a model $M \colon \mathbf{T}' \to \mathbf{Set}$, we get by composition with I a model $M \circ I \colon \mathbf{T} \to \mathbf{Set}$. Hence, there is a functor $I^* \colon \mathrm{Mod}(\mathbf{T}', \mathbf{Set}) \to \mathrm{Mod}(\mathbf{T}, \mathbf{Set})$. The natural questions to ask pertain to the relations between I^* and I. More specifically, are there properties of I^* that imply properties of I? In particular, is it possible that I^* being an equivalence of categories imply that I is? In words, what are the properties of the conceptual that affect the properties of the formal?

2. Given a functor $F \colon \mathcal{C} \to \mathcal{D}$ of the right type (that preserves the right kind of structure in each case), we get a functor $F^* \colon \mathrm{Mod}([\mathbf{T}], \mathcal{C}) \to \mathrm{Mod}([\mathbf{T}], \mathcal{D})$ by composing models M with F. One question here focuses on the categories $\mathcal{C}$ and $\mathcal{D}$, more specifically on $\mathcal{C}$ and the abstract mathematical structure

---

[20] We are not being careful here. Some of these are covariant functors, while others are contravariant. We simply want to point at the possibilities at this juncture. We are not developing the theory as such. Again, these details do not affect our main point.

both these categories are instances of. Thus, is there an abstract mathematical type of structure such that all the models of $[\mathbf{T}]$ in $\mathcal{D}$ arise from models of $[\mathbf{T}]$ in $\mathcal{C}$ and functors $\mathcal{C} \to \mathcal{D}$?

Other questions can be formulated, but these are not unlike questions that arise in other mathematical domains, thus relating this formulation of logic with comparable frameworks. To be able to identify what is the common abstract core of logic with other mathematical domains and what is specific to logic is one of the gains of the abstract structuralist approach.

4. Metalogical theorems from an abstract structural standpoint

From a structuralist standpoint, once the abstract mathematical structures have been identified, one hopes to be able to prove standard theorems from that vantage. And, indeed, one can. One of the epistemic gains expected from these theorems is the identification of the abstract components involved in various proofs and thus see what is the core structural component upon which these results are grounded. Another expected benefit is the possibility to get genuinely new results which were impossible to get in the classical framework, even impossible to formulate adequately.

Completeness and conceptual completeness

Let us start with what can be considered the pillar of logic in general, namely completeness results. Completeness results for various propositional logics are equivalent to representation theorems for

various algebras, e.g., in the case of classical propositional logic, the completeness theorem is equivalent to Stone's representation theorem for Boolean algebras. As we have already mentioned, the most natural context to prove this result is already the context of the category of Boolean algebras and the theorem is done up to isomorphism.

Moving to first-order logic, it is to be expected that the completeness theorems would amount to representation theorems for certain categories, e.g. regular, coherent, pretoposes, Heyting, Boolean, etc. Indeed, the classical (Gödel) completeness theorem is equivalent to a representation theorem for coherent categories, which can be stated thus: for any small coherent category C, there is a (small) set I and a conservative coherent functor $F \colon \mathcal{C} \to \mathbf{Set}^I$. A functor $F \colon \mathcal{C} \to \mathcal{D}$ is said to be *conservative* if it reflects isomorphisms, i.e., if $F(f)$ is an isomorphism in $\mathcal{D}$, then f was already an isomorphism in $\mathcal{C}$. Needless to say, the key property is precisely that of being conservative. For what it amounts to is the fact that for any diagram in $\mathcal{C}$ such that its image under F in $\mathcal{D}$ is a diagram of a universal morphism, then the original diagram was already a diagram of a universal morphism in $\mathcal{C}$.

As we have already mentioned, the category $\mathbf{Set}$ is coherent and so is the functor category $\mathbf{Set}^I$. Since the functor $F \colon \mathcal{C} \to \mathbf{Set}^I$ is conservative, it follows that $\mathcal{C}$ shares all the coherent properties of $\mathbf{Set}^I$, and in fact of $\mathbf{Set}$. The equivalence between the representation theorem and the completeness theorem can be established as follows. Assuming the representation theorem, we start with a coherent theory $\mathbf{T}$ and construct the category of concepts $[\mathbf{T}]$ of $\mathbf{T}$, which is a coherent category. Applying the representation theorem to $[\mathbf{T}]$, we obtain the completeness theorem. To prove the other direction, we assume the completeness theorem and start with a coherent category $\mathcal{C}$. Using the internal language of $\mathcal{C}$, one constructs as above the coherent theory $\mathbf{T}_{\mathcal{C}}$

of $\mathcal{C}$. The models of $\mathbf{T}_{\mathcal{C}}$ are then constructed so that they are identical with functors $\mathcal{C} \to \mathbf{Set}$. The representation theorem then follows from the completeness theorem for $\mathbf{T}_{\mathcal{C}}$.

Two important elements have to be added to the picture. First, the representation theorem for coherent categories is but one representation theorem for a whole collection of relevant categories: regular categories, pretoposes, Heyting categories and Boolean categories. Second, these results in fact follow a general pattern. Indeed, the foregoing representation theorem takes a general, purely categorical form, in other words, there is a crucial part that is purely based on the abstract mathematical structures. Given any categories $\mathcal{S}$ and $\mathcal{C}$, we can always consider the repeated functor category $\mathcal{S}^{(\mathcal{S}^{\mathcal{C}})}$.[21] In this situation, there is a canonical functor, the evaluation functor

$$e \colon \mathcal{C} \to \mathcal{S}^{(\mathcal{S}^{\mathcal{C}})}$$

for which, given any object X of $\mathcal{C}$, and any functor $F \colon \mathcal{C} \to \mathcal{S}$, $e(X)(F)$ is simply $F(X)$, the evaluation of F at X. For any subcategory $\mathcal{D}$ of $\mathcal{S}^{\mathcal{C}}$, the same functor $e \colon \mathcal{C} \to \mathcal{S}^{\mathcal{D}}$ can be defined. It is then possible to show that the representation theorem for coherent categories is equivalent to the claim that the functor $e \colon \mathcal{C} \to \mathcal{S}^{\mathrm{Mod}(\mathcal{C})}$ is conservative. The fact that the evaluation functor is coherent holds on purely general grounds. We therefore have a purely categorical description of the representation theorem. Moreover, in the early seventies Joyal demonstrated that the functor e preserves all existing instances of the Heyting structure in $\mathcal{C}$. This automatically yields a representation theorem for Heyting categories and, in turn, a canonical completeness theorem for intuitionistic logic.

---

[21] This is not an unusual construction in mathematics. Think of the double dual of a finite-dimensional vector space, for instance.

The categorical set-up allows is to consider a stronger claim, called the *conceptual completeness*. Given a functor $I : \mathbf{T} \to \mathbf{T}'$ and an equivalence of categories between $\mathrm{Mod}(\mathbf{T}', \mathbf{Set})$ and $\mathrm{Mod}(\mathbf{T}, \mathbf{Set})$, when is it possible to conclude that I is also an equivalence of categories? From a categorical point of view, the assumption means that the category of models of $\mathbf{T}'$ is indistinguishable from the category of models of $\mathbf{T}$. We can think of the functor I as a translation of $[\mathbf{T}]$ into $[\mathbf{T}']$, thus as a case when the latter theory can in principle be more expressive than the former. In a sense, the conceptual completeness can be interpreted as saying that adding new concepts to $\mathbf{T}$ simply does not modify in any essential way what it can express. This means that $\mathbf{T}$ has some sort of completeness and in this context it makes perfect sense to say that it is *conceptually complete*. Thus, we say that $\mathbf{T}$ is *conceptually complete* whenever the following is satisfied: if the functor $I^*\colon \mathrm{Mod}(\mathbf{T}', \mathbf{Set}) \to \mathrm{Mod}(\mathbf{T}, \mathbf{Set})$ is an equivalence of categories, then the functor $I\colon \mathbf{T} \to \mathbf{T}'$ was one already. This literally means that by moving to $\mathbf{T}'$, we did not add anything essentially new to $\mathbf{T}$, although we might have thought we had, and this information was obtained by looking at the *categorical structure of the category of models of the theories*. We can conclude that a certain logical framework, say an equational theory, is enough to characterize a type of structures, from the categorical structure of the category of models. Conceptual completeness is in fact equivalent to a standard result of model theory, namely Beth definability theorem. However, one of the advantages of working in the categorical framework is that categorical methods make it possible to prove results which might not be accessible otherwise, for instance, a constructive proof of this result for intuitionistic logic.[22]

---

[22] See (Pitts, 1989) for a categorical proof of conceptual completeness of intuitionistic first-order logic.

It is possible to strengthen the conceptual completeness theorem. In the latter, we assume as given a functor $I : \mathbf{T} \to \mathbf{T}'$. Is it possible to start with an equivalence of categories $\mathrm{Mod}(\mathbf{T}, \mathbf{Set}) \to \mathrm{Mod}(\mathbf{T}', \mathbf{Set})$ and construct from it an equivalence $\mathbf{T} \to \mathbf{T}'$? This is much stronger theorem, but it can be proved under certain circumstances. It says that a logical theory is completely characterized by the categorical structure of its category of models. In some sense, the conceptual determines the formal, up to equivalence. In the case of propositional logic, it amounts to a form of Stone duality, the latter being formulated entirely *within* the category of Boolean algebras, and not as the existence of an equivalence of categories between the category of Boolean algebras and the category of Stone spaces. The strong conceptual completeness asserts that the Lindenbaum-Tarski algebra of a propositional theory can be recovered from its space of models— the ultrafilters on the given Boolean algebra. A theory for which the theorem can be proved is said to be *strongly conceptually complete*. A different way to formulate this result is to say that if $\mathrm{Mod}(\mathbf{T}, \mathbf{Set})$ and $\mathrm{Mod}(\mathbf{T}', \mathbf{Set})$ are equivalent, then $\mathbf{T}$ and $\mathbf{T}'$ are equivalent too. Whereas conceptual completeness is a local phenomenon, since it depends on the interpretation I, strong conceptual completeness is a global phenomenon, since there is no underlying interpretation at hand. The construction of $\mathbf{T}$ can be thought of as a case of abstracting certain data out of another, more "concrete", situation. Finite limit categories of concepts are strongly conceptually complete[23], although the original result applied to (Boolean) pretoposes. If the category of models is adequately enriched in a precise technical sense, then in these circumstances first-order classical logic is strongly conceptually complete (see Makkai, 1988; 1990; for a a different proof which is

---

[23] Thus they are the so-called Barr-exact categories. A Barr-exact category is a regular category in which every equivalence relation is a kernel pair (see Makkai, 1990).

build with higher-dimensional categories in mind, see Lurie, 2019). Notice that it is hard to see how this theorem could even be formulated outside the context of category theory.

It is impossible not to mention the fact that strong conceptual completeness theorems are closely related to dualities. In fact, they are equivalent in a precise technical sense to dualities. The only thing we want to underline is that these results are proved in the context of 2-categories. Thus, it is not only that the natural set-up involves 2-categories, but that important theorems require 2-categorical (even bicategorical) concepts. We cannot, in such a short paper, present these in any comprehensible manner.

Syntax and abstract completeness

From the above considerations, the reader might feel that we have entirely left behind syntactical considerations, more specifically formal deductions. Therefore, it might seem like the categorical completeness results are not quite the same as the classical results which assert that semantical consequences of a theory are provable in a fully specified formal system. This is not the case. For one thing, we have not abandoned the syntax, nor the formal systems in these investigations. But there is an additional point to make, for it brings to the fore a way of dealing with the syntax of theories that emerged naturally from the context of categories, namely the idea of a sketch and its generalizations.

From Lawvere's thesis, category theorists toyed with the idea that a theory could be presented directly in the form of a category or some graphical variant thereof. In this spirit, sketches were introduced by Charles Ehresmann in the early 1960s and developed afterwards by

him and his school (see Ehresmann, 1967; 1968; Lair, 2001; 2002; 2003, for instance). A sketch, which is a specific kind of (oriented) graph, is a new kind of syntax, specifically tailored to do categorical logic. It gives directly in a graphical way the syntactical and proof theoretical content of a theory.

We will give one definition of the notion of sketch[24]. A sketch $S = (G, D, L, C)$ is given by a graph G, a set D of diagrams in G, a set L of cones in G and a set C of cocones in G. We can consider the category of sketches by stipulating that a morphism of sketches is a homomorphism of graphs which preserves the diagrams, the cones and the cocones. It is easy to see that any category $\mathbb{C}$ has an underlying sketch $S_{\mathbb{C}}$. A *model* of a sketch S in a category $\mathbb{C}$ takes all the diagrams of S to commutative diagrams, all the cones of S to limits of $\mathbb{C}$ and all cocones of S to colimits of $\mathbb{C}$. A morphism of models is a natural transformation. Thus, we can reproduce what we did above with theories, namely we can construct the category of models $\mathrm{Mod}(S, \mathbb{C})$ of a sketch S in a category $\mathbb{C}$.

It is natural to consider sketches in which there are no cocones and only discrete and finite cones, or in which there are no cocones and only finite cones, etc. Sketches organized themselves with respect to these natural choices and they correspond to various logical theories. Thus, there are finite product sketches (a FP-sketch), left exact sketches (a LE-sketch), regular sketches, coherent sketches, etc.

In this framework, it is natural to ask which categories are sketchables: is it possible to characterize categories that are equivalent to categories of models of a type of sketch? There are positive answers to that question and it naturally brings us, when the most general kinds of sketches are considered, to infinitary logic $L_{\infty,\infty}$ (see Lair, 1981; Makkai and Paré, 1989).

---

[24] See (Barr and Wells, 1990; 2005) for an introduction to sketches.

Generalization of the notion of sketch has led Makkai to develop a categorical proof theory and establish a completeness theorem along the classical lines, that is proving that a formula is formally provable in a theory if and only if it is true in all models of the theory. Interestingly enough, the set-up still rests upon the categorical representation theorems, but it is enriched with a categorical notion of formal proof in the set-up of (generalized) sketches. (see Makkai, 1997a,b,c).

Incompleteness

We have to say a few words about Gödel's incompleteness theorems. Is it possible to identify abstract mathematical structures that underly these theorems? Is it possible to deduce these theorems from a theorem or theorems about these abstract mathematical structures? There are some pieces in place, although the complete picture—no pun intended—has still to be presented.

First, already in the 1960s, Lawvere presented a categorical analysis of various phenomena related to Gödels's incompleteness theorems. In his (1969b), Lawvere presents what he takes to be the abstract mathematical structure underlying Cantor's theorem that there is no surjection $\mathbb{X} \to 2^{\mathbb{X}}$ and its variants in the heads of Russell, Gödel and Tarski. The starting point here is the notion of a cartesian closed category. A cartesian closed category is a nice example of a categorical doctrine since it can be given entirely by stipulating the existence of certain adjoint functors to elementary, that is first-order, functors. More precisely, a cartesian closed category $\mathbb{C}$ is a category such that

1. The functor $! : \mathbb{C} \to 1$ has a right adjoint;
2. The diagonal functor $\Delta : \mathbb{C} \to \mathbb{C} \times \mathbb{C}$ has a right adjoint, namely the product functor;

3. For each object X of $\mathbb{C}$, the functor $X \times (-) : \mathbb{C} \to \mathbb{C}$ has a right adjoint $(-)^X$.

Given a cartesian closed category $\mathbb{C}$, it is possible to then define what D. Pavlovic has called a "paradoxical structure" on an object of $\mathbb{C}$ that satisfies a fixed-point property. By specifying the adequate cartesian closed category $\mathbb{C}$ and the paradoxical structure on objects of $\mathbb{C}$, it is possible to prove Cantor's theorem, Russell's paradox, Gödel's first incompleteness theorem, Tarski's theorem of the impossibility of defining truth in a theory and many others (see Pavlović, 1992 and Yanofsky, 2003).

In a series of unpublished lectures presented in the 1970s, André Joyal introduced another abstract mathematical structure, in a precise sense weaker than Lawvere's proposal, to pursue the analysis of Gödel's incompleteness results, in particular the second theorem, namely what he called 'arithmetical universes'. Very roughly, an arithmetic universe $\mathbb{U}$ is a pretopos such that the free category object constructed from a graph object in $\mathbb{U}$ exists. This is the abstract structure in which one can do recursive arithmetic and prove versions of the two incompleteness theorems (see Maietti, 2010 for a different definition).

Type theories

We will be very succinct, not because this area is not important, quite the contrary, but simply because there is no need to cover everything in details given the goal of this paper.

The first and well-known result in this area is the correspondence between cartesian closed categories with a natural number object and typed λ-calculus. We finally get to elementary toposes. These were

introduced by Lawvere and Tierney in 1970 to provide an elementary treatment of sheaves over a site, thus of Grothendieck toposes. It is another remarkable example of a categorical doctrine. An *elementary topos* $\mathcal{E}$ is a category with finite limits, cartesian closed, and has a subobject classifier. These three conditions do amount to the existence of certain adjoint functors to given (elementary) functors. As is well known now, it is possible to construct an intuitionistic type theory from a given topos $\mathcal{E}$ and, conversely, it is possible to specify an intuitionistic type theory such that its conceptual category is an elementary topos and it can be interpreted in an elementary topos. There is then a correspondance between categorical properties of the topos and logical properties of the type theory (see Boileau and Joyal, 1981; Lambek and Scott, 1988).

The same can be said about homotopy type theory. Homotopy type theory comes form Martin-Löf's intensional type theory (see The Univalent Foundations Program, 2013). It has models in various categories, but a homotopy type theory ought to correspond to a kind of abstract categories. It has been conjectured, by Steve Awodey, that homotopy type theory should correspond to the internal logic of higher-dimensional elementary toposes. As of this writing, the full conjecture has still to be proved, although certain advances have been made (see Kapulkin and Lumsdaine, 2018).

One last thing...

Last but not least, connections between linear logic and category theory appeared almost immediately after the creation of linear logic by Girard in (1987) (see Lafont, 1988; Seely, 1989). It took almost twenty years of research before a consensus emerged as to what

constitutes a categorical model of linear logic (see Bierman, 1995; Blute and Scott, 2004; Melliès, 2009; de Paiva, 2014). We will not introduce nor discuss the categorical framework here. It would require defining and explaning a lot of categorical structures, e.g. symmetric monoidal categories, symmetric monoidal adjunctions, etc., as well as an explanation of how the various frameworks proposed converge towards a basic structure. The point is: we may be seeing the beginning of a stable picture that will allow us to start building a conceptual interpretation of linear logic. In as much as homotopy type theory seems to be intimately connected to the basic constituents of spaces, the "atoms of space" to use Baues's expression in (2002), namely homotopy types, linear logic seems to be intimately tied to generalized vector spaces and the mathematics inherent to the latter (see Melliès, n.d.). If this reading is correct, it may lead to new interpretations and developments of conceptual spaces and the categorical structures would naturally find their place in that context. But this is sheer speculation at this point and it does not affect our main point.

5. Conclusion

We insist that this way of framing logic and metalogic is a direct continuation of mathematical logic as it developed in the first half of the 20th century and the rise of the abstract axiomatic method at the same time. Category theory itself is an offspring of this period, and as such, does not constitute a radical methodological change. It is, undoubtedly, a rise in abstraction. It reveals new types of abstract mathematical structures.

We hope we have convinced the reader that it is possible to identify the abstract mathematical structures underlying (fragments of, and

extensions of) first-order logic and type theories. It is also possible to see how the important metalogical results correspond to theorems on these abstract mathematical structures. Finally, the invariance property at the core of any abstract mathematical structuralism comes naturally and automatically in this framework. Thus, the standard logical systems—and some non-standard logical systems as well—find a natural place in this structuralist context. Pure logic is seen as a specific type of abstract structure. We can thus answer our own challenge positively and precisely. We leave to other structuralists to provide their answer to our challenge.

Bibliography

Awodey, S., 1996. Structure in mathematics and logic: a categorical perspective. *Philos. Math. (3)* [Online], 4(3), pp.209–237. Available at: https://doi.org/10.1093/philmat/4.3.209.

Barr, M. and Wells, C., 1990. *Category Theory for Computing Science* [Online], *Prentice Hall International Series in Computer Science*. New York: Prentice Hall International, pp.xvi+432. Available at: https://mathscinet. ams.org/mathscinet-getitem?mr=1094561.

Barr, M. and Wells, C., 2005. Toposes, triples and theories. *Repr. Theory Appl. Categ.* [Online], (12), pp.x+288. Available at: https://mathscinet. ams.org/mathscinet-getitem?mr=2178101.

Baues, H.-J., 2002. Atoms of topology. *Jahresber. Deutsch. Math.-Verein.*, 104(4), pp.147–164.

Bierman, G.M., 1995. What is a categorical model of intuitionistic linear logic? *Typed Lambda Calculi and Applications (Edinburgh, 1995)* [Online]. Vol. 902, *Lecture Notes in Comput. Sci.* Springer, Berlin, pp.78–93. Available at: https://doi.org/10.1007/BFb0014046.

Birkhoff, G., 1940. *Lattice Theory* [Online]. New York: American Mathematical Society, pp.v+155. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=0001959.

Blute, R. and Scott, P., 2004. Category theory for linear logicians. *Linear Logic in Computer Science* [Online]. Vol. 316, *London Math. Soc. Lecture Note Ser.* Cambridge: Cambridge Univ. Press, pp.3–64. Available at: https://doi.org/10.1017/CBO9780511550850.002.

Boileau, A. and Joyal, A., 1981. La logique des topos. *J. Symbolic Logic* [Online], 46(1), pp.6–16. Available at: https://doi.org/10.2307/2273251.

Borceux, F., 1994. *Handbook of Categorical Algebra. 2* [Online]. Vol. 51, *Encyclopedia of Mathematics and its Applications*. Cambridge: Cambridge University Press, pp.xviii+443. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=1313497.

Caramello, O., 2014. Topologies for intermediate logics. *MLQ. Mathematical Logic Quarterly* [Online], 60(4-5), pp.335–347. Available at: https://doi.org/10.1002/malq.201200059.

Caramello, O., 2018. *Theories, sites, toposes*. Oxford: Oxford University Press.

Corry, L., 2004. *Modern Algebra and the Rise of Mathematical Structures*. Second. Basel: Birkhäuser Verlag, pp.xvi+451.

Ehresmann, C., 1967. Sur les structures algébriques. *C. R. Acad. Sci. Paris Sér. A-B* [Online], 264, A840–A843. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=217140.

Ehresmann, C., 1968. Esquisses et types des structures algébriques. *Bul. Inst. Politehn. Iaşi (N.S.)* [Online], 14(18)(fasc., fasc. 1-2), pp.1–14. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=238918.

Eilenberg, S. and Mac Lane, S., 1945. A General Theory of Natural Equivalences. *Trans. Amer. Math. Soc.*, 58, pp.231–294.

Freyd, P., 1966. Representations in Abelian Categories. *Proceedings of the Conference on Categorical Algebra, La Jolla, Calif., 1965*. New York: Springer-Verlag, pp.95–120.

Girard, J.-Y., 1987. Linear logic. *Theoret. Comput. Sci.* [Online], 50(1), p.101. Available at: https://doi.org/10.1016/0304-3975(87)90045-4.

Halmos, P.R., 1954. Polyadic Boolean algebras. *Proc. Nat. Acad. Sci. U.S.A.* [Online], 40, pp.296–301. Available at: https://doi.org/10.1073/pnas.40.5.296.

Halmos, P.R., 1956a. Algebraic logic. I. Monadic Boolean algebras. *Compositio Math.* [Online], 12, pp.217–249. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=78304.

Halmos, P.R., 1956b. Algebraic logic. II. Homogeneous locally finite polyadic Boolean algebras of infinite degree. *Fund. Math.* [Online], 43, pp.255–325. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=86029.

Halmos, P.R., 1956c. Algebraic logic. III. Predicates, terms, and operations in polyadic algebras. *Trans. Amer. Math. Soc.* [Online], 83, pp.430–470. Available at: https://doi.org/10.2307/1992883.

Halmos, P.R., 1956d. The basic concepts of algebraic logic. *Amer. Math. Monthly* [Online], 63, pp.363–387. Available at: https://doi.org/10.2307/2309396.

Halmos, P.R., 1962. *Algebraic Logic* [Online]. New York: Chelsea Publishing Co., p.271. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=0131961.

Hellman, G., 2001. Three varieties of mathematical structuralism. *Philos. Math. (3)* [Online], 9(2), pp.184–211. Available at: https://doi.org/10.1093/philmat/9.2.184.

Hellman, G., 2003. Does category theory provide a framework for mathematical structuralism? *Philos. Math. (3)* [Online], 11(2), pp.129–157. Available at: https://doi.org/10.1093/philmat/11.2.129.

Hellman, G. and Shapiro, S., 2019. *Mathematical Structuralism* [Online], *Elements in the Philosophy of Mathematics*. Cambridge: Cambridge University Press, p.92. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=3967722.

Henkin, L., Monk, J.D. and Tarski, A., 1971. *Cylindric Algebras. Part I. With an Introductory chapter: General Theory of Algebras* [Online]. Amsterdam; London: North-Holland Publishing Co., pp.vi+508. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=0314620.

Henkin, L. and Tarski, A., 1961. Cylindric algebras. *Proc. Sympos. Pure Math., Vol. II* [Online]. Providence, R.I.: American Mathematical Society, pp.83–113. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=0124250.

Johnstone, P.T., 2002. *Sketches of an Elephant: A Topos Theory Compendium. Vol. 1* [Online]. Vol. 43, *Oxford Logic Guides*. New York: The Clarendon Press, Oxford University Press, pp.xxii+468+71. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=1953060.

Jónsson, B. and Tarski, A., 1951. Boolean algebras with operators. I. *Amer. J. Math.* [Online], 73, pp.891–939. Available at: https://doi.org/10.2307/2372123.

Jónsson, B. and Tarski, A., 1952. Boolean algebras with operators. II. *Amer. J. Math.* [Online], 74, pp.127–162. Available at: https://doi.org/10.2307/2372074.

Kapulkin, K. and Lumsdaine, P.L., 2018. The homotopy theory of type theories. *Adv. Math.* [Online], 337, pp.1–38. Available at: https://doi.org/10.1016/j.aim.2018.08.003.

Krömer, R., 2007. *Tool and Object* [Online]. Vol. 32, *Science Networks. Historical Studies*. Basel: Birkhäuser Verlag, pp.xxxvi+367. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=2272843.

Lafont, Y., 1988. The linear abstract machine. Vol. 59, 1-2, pp.157–180. Available at: https://doi.org/10.1016/0304-3975(88)90100-4.

Lair, C., 1981. Catégories modelables et catégories esquissables. *Diagrammes* [Online], 6, pp.L1–L20. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=684535.

Lair, C., 2001. Éléments de théorie des esquisses. I. Graphes à composition. *Diagrammes* [Online], 45/46, pp.3–33. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=2102295.

Lair, C., 2002. Éléments de théorie des esquisses. II. Systèmes tensoriels et systèmes enrichis de graphes à composition. *Diagrammes* [Online], 47/48, p.34. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=2104013.

Lair, C., 2003. Éléments de théorie des esquisses. III. Esquisses. *Diagrammes* [Online], 49/50, p.58. Available at: https : / / mathscinet . ams . org / mathscinet-getitem?mr=2104014.

Lambek, J., 1968a. A Fixpoint Theorem for Complete Categories. *Math. Z.*, 103, pp.151–161.

Lambek, J., 1968b. Deductive Systems and Categories. I. Syntactic Calculus and Residuated Categories. *Math. Systems Theory*, 2, pp.287–318.

Lambek, J., 1969. Deductive Systems and Categories II. Standard Constructions and Closed Categories. In: P.J. Hilton, ed. *Category Theory, Homology Theory, and their Applications, I*. Vol. 86, Lecture Notes in Mathematics. Berlin: Springer-Verlag, pp.76–122.

Lambek, J., 1972. Deductive Systems and Categories III. Cartesian Closed Categories, Intuitionistic Propositional Calculus, and Combinatory Logic. In: F.W. Lawvere, ed. *Toposes, Algebraic Geometry and Logic*. Vol. 274, Lecture Notes in Mathematics. Berlin: Springer-Verlag, pp.57–82.

Lambek, J. and Scott, P.J., 1988. *Introduction to Higher Order Categorical Logic* [Online]. Vol. 7, *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, pp.x+293. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=939612.

Landry, E. and Marquis, J.-P., 2005. Categories in Context: Historical, Foundational and Philosophical. *Philos. Math. (3)*, 13(1), pp.1–43.

Lawvere, F.W., 1963. Functorial semantics of algebraic theories. *Proc. Nat. Acad. Sci. U.S.A.* [Online], 50, pp.869–872. Available at: https://doi.org/10.1073/pnas.50.5.869.

Lawvere, F.W., 1966. Functorial Semantics of Elementary Theories. *Journal of Symbolic Logic*, 31, pp.294–295.

Lawvere, F.W., 1967. Theories as Categories and the Completeness Theorem. *Journal of Symbolic Logic*, 32, p.562.

Lawvere, F.W., 1969a. Adjointness in Foundations. *Dialectica*, 23, pp.281–296.

Lawvere, F.W., 1969b. Diagonal arguments and cartesian closed categories. *Category Theory, Homology Theory and their Applications, II (Bat-*

telle Institute Conference, Seattle, Wash., 1968, Vol. Two) [Online]. Berlin: Springer, pp.134–145. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=0242748.

Lawvere, F.W., 1970. Equality in hyperdoctrines and comprehension schema as an adjoint functor. *Applications of Categorical Algebra (Proc. Sympos. Pure Math., Vol. XVII, New York, 1968)* [Online]. Providence, R.I.: Amer. Math. Soc., pp.1–14. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=0257175.

Lawvere, F.W., 1971. Quantifiers and sheaves. *Actes du Congrès International des Mathématiciens (Nice, 1970), Tome 1* [Online], pp.329–334. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=0430021.

Linton, F.E.J., 1966. Some aspects of equational categories. *Proc. Conf. Categorical Algebra (La Jolla, Calif., 1965)*. New York: Springer, pp.84–94.

Lurie, J., 2019. *Ultracategories*. [unpublished].

Mac Lane, S., 1965. Categorical Algebra. *Bull. Amer. Math. Soc.*, 71, pp.40–106.

Mac Lane, S., 1998. *Categories for the Working Mathematician*. 2nd ed. Vol. 5, *Graduate Texts in Mathematics*. New York: Springer-Verlag.

Mac Lane, S. and Moerdijk, I., 1994. *Sheaves in Geometry and Logic* [Online], *Universitext*. New York: Springer-Verlag, pp.xii+629. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=1300636.

Maietti, M.E., 2010. Joyal's arithmetic universe as list-arithmetic pretopos. *Theory Appl. Categ.* [Online], 24, No. 3, 39–83. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=2610177.

Makkai, M., 1988. Strong conceptual completeness for first-order logic. *Ann. Pure Appl. Logic* [Online], 40(2), pp.167–215. Available at: https://doi.org/10.1016/0168-0072(88)90019-X.

Makkai, M., 1990. A theorem on Barr-exact categories, with an infinitary generalization. *Ann. Pure Appl. Logic* [Online], 47(3), pp.225–268. Available at: https://doi.org/10.1016/0168-0072(90)90036-2.

Makkai, M., 1997a. Generalized sketches as a framework for completeness theorems. I. *J. Pure Appl. Algebra* [Online], 115(1), pp.49–79. Available at: https://doi.org/10.1016/S0022-4049(96)00007-2.

Makkai, M., 1997b. Generalized sketches as a framework for completeness theorems. II. *J. Pure Appl. Algebra* [Online], 115(2), pp.179–212. Available at: https://doi.org/10.1016/S0022-4049(96)00008-4.

Makkai, M., 1997c. Generalized sketches as a framework for completeness theorems. III. *J. Pure Appl. Algebra* [Online], 115(3), pp.241–274. Available at: https://doi.org/10.1016/S0022-4049(96)00009-6.

Makkai, M. and Paré, R., 1989. *Accessible Categories: The Foundations of Categorical Model Theory* [Online]. Vol. 104, *Contemporary Mathematics*. Providence, R.I.: American Mathematical Society, pp.viii+176. Available at: https://doi.org/10.1090/conm/104.

Makkai, M. and Reyes, G.E., 1977. *First Order Categorical Logic* [Online], *Lecture Notes in Mathematics, Vol. 611*. Berlin; New York: Springer-Verlag, pp.viii+301. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=0505486.

Marquis, J.-P., 2009. *From a Geometrical Point of View* [Online]. Vol. 14, *Logic, Epistemology, and the Unity of Science*. Springer, Dordrecht, pp.x+309. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=2730089.

Marquis, J.-P., 2015. Mathematical abstraction, conceptual variation and identity. In: P.-E.B.G.H.W.H.P. Schroeder-Heister, ed. *Logic, Methodology and Philosophy of Science, Proceedings of the Fourteen International Congress*. London: College Publications, pp.299–322.

Marquis, J.-P., 2016. Stairway to Heaven: The Abstract Method and Levels of Abstraction in Mathematics. *The Mathematical Intelligencer*, 38(3), pp.41–51.

Marquis, J.-P. and Reyes, G.E., 2012. The history of categorical logic: 1963–1977. *Sets and Extensions in the Twentieth Century* [Online]. Vol. 6, *Handb. Hist. Log.* Amsterdam: Elsevier/North-Holland, pp.689–800. Available at: https://doi.org/10.1016/B978-0-444-51621-3.50010-4.

McKinsey, J.C.C. and Tarski, A., 1944. The algebra of topology. *Ann. of Math. (2)* [Online], 45, pp.141–191. Available at: https://doi.org/10.2307/1969080.

McKinsey, J.C.C. and Tarski, A., 1946. On closed elements in closure algebras. *Ann. of Math. (2)* [Online], 47, pp.122–162. Available at: https://doi.org/10.2307/1969038.

McKinsey, J.C.C. and Tarski, A., 1948. Some theorems about the sentential calculi of Lewis and Heyting. *J. Symbolic Logic* [Online], 13, pp.1–15. Available at: https://doi.org/10.2307/2268135.

Melliès, P.A., n.d. *A Functorial Excursion Between Algebraic Geometry and Linear Logic*.

Melliès, P.-A., 2009. Categorical semantics of linear logic. In: P.-L. Curien, H. Herbelin, J.-L. Krivine and P.-A. Méllies, eds. *Interactive models of computation and program behaviour*, *Panoramas et Synthèses* 27. Société Mathématique de France, pp.1–196.

Mostowski, A., 1949. Sur l'interpretation géométrique et topologique des notions logiques. *Library of the Tenth International Congress of Philosophy, Amsterdam, August 11–18, 1948, Vol. I, Proceedings of the Congress* [Online]. publisher unknown, pp.767–769. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=0028277.

de Paiva, V., 2014. Categorical semantics of linear logic for all. *Advances in Natural Deduction* [Online]. Vol. 39, *Trends Log. Stud. Log. Libr.* Springer, Dordrecht, pp.181–192. Available at: https://doi.org/10.1007/978-94-007-7548-0_9.

Pavlović, D., 1992. On the structure of paradoxes. *Arch. Math. Logic* [Online], 31(6), pp.397–406. Available at: https://doi.org/10.1007/BF01277482.

Pitts, A.M., 1989. Conceptual completeness for first-order intuitionistic logic: an application of categorical logic. *Ann. Pure Appl. Logic* [Online], 41(1), pp.33–81. Available at: https://doi.org/10.1016/0168-0072(89)90007-9.

Rasiowa, H., 1951. Algebraic treatment of the functional calculi of Heyting and Lewis. *Fund. Math.* [Online], 38, pp.99–126. Available at: https://doi.org/10.4064/fm-38-1-99-126.

Rasiowa, H., 1955. Algebraic models of axiomatic theories. *Fund. Math.* [Online], 41, pp.291–310. Available at: https://doi.org/10.4064/fm-41-2-291-310.

Rasiowa, H. and Sikorski, R., 1950. A proof of the completeness theorem of Gödel. *Fund. Math.* [Online], 37, pp.193–200. Available at: https://doi.org/10.4064/fm-37-1-193-200.

Rasiowa, H. and Sikorski, R., 1953. Algebraic treatment of the notion of satisfiability. *Fund. Math.* [Online], 40, pp.62–95. Available at: https://doi.org/10.4064/fm-40-1-62-95.

Rasiowa, H. and Sikorski, R., 1955. An application of lattices to logic. *Fund. Math.* [Online], 42, pp.83–100. Available at: https://doi.org/10.4064/fm-42-1-83-100.

Rasiowa, H. and Sikorski, R., 1963. *The Mathematics of Metamathematics* [Online], *Monografie Matematyczne, Tom 41*. Państwowe Wydawnictwo Naukowe, Warsaw, p.522. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=0163850.

Reck, E.H. and Schiemer, G., eds., 2020. *The Prehistory of Mathematical Structuralism*. Oxford: Oxford University Press.

Riehl, E., 2017. *Category Theory in Context*. Mineola: Dover Publications.

Rodin, A., 2014. *Axiomatic Method and Category Theory* [Online]. Vol. 364, *Synthese Library. Studies in Epistemology, Logic, Methodology, and Philosophy of Science*. Cham: Springer, pp.xii+285. Available at: https://doi.org/10.1007/978-3-319-00404-4.

Seely, R.A.G., 1989. Linear logic, ∗-autonomous categories and cofree coalgebras. *Categories in Computer Science and Logic (Boulder, CO, 1987)* [Online]. Vol. 92, *Contemp. Math.* Amer. Math. Soc., Providence, RI, pp.371–382. Available at: https://doi.org/10.1090/conm/092/1003210.

The Univalent Foundations Program, 2013. *Homotopy Type Theory—Univalent Foundations of Mathematics* [Online]. Princeton, N.J.: The Univalent Foundations Program; Institute for Advanced Study (IAS), pp.xiv+589. Available at: https://mathscinet.ams.org/mathscinet-getitem?mr=3204653.

Yanofsky, N.S., 2003. A universal approach to self-referential paradoxes, incompleteness and fixed points. *Bull. Symbolic Logic* [Online], 9(3), pp.362–386. Available at: https://doi.org/10.2178/bsl/1058448677.

On the validity of the definition of a complement-classifier

Mariusz Stopa

Institute of Philosophy, Jagiellonian University
Dominican College of Philosophy and Theology

Abstract

It is well-established that topos theory is inherently connected with intuitionistic logic. In recent times several works appeared concerning so-called complement-toposes (co-toposes), which are allegedly connected to the dual to intuitionistic logic. In this paper I present this new notion, some of the motivations for it, and some of its consequences. Then, I argue that, assuming equivalence of certain two definitions of a topos, the concept of a complement-classifier (and thus of a co-topos as well) is, at least in general and within the conceptual framework of category theory, not appropriately defined. For this purpose, I first analyze the standard notion of a subobject classifier, show its connection with the representability of the functor Sub via the Yoneda lemma, recall some other properties of the internal structure of a topos and, based on these, I critically comment on the notion of a complement-classifier (and thus of a co-topos as well).

Keywords

category theory, topos theory, categorical logic, Heyting algebras, co-Heyting algebras, intuitionistic logic, dual intuitionistic logic, complement-classifier.

1. General introduction

Category theory, and especially topos theory, have changed the way we think about the role of logic in mathematics, and, through mathematics, perhaps also in physics. It is common knowledge that toposes are intrinsically connected to intuitionistic logic. However, recently there appeared several works concerning so-called complement-toposes (co-toposes), which are, supposedly, connected to a certain type of paraconsistent logic called dual to intuitionistic, or anti-intuitionistic, logic, which algebra is a co-Heyting one.[1] It is known that, indeed, some toposes having co-Heyting structures exhibit some aspects of dual to intuitionistic logic (see Reyes and Zolfaghari, 1996, and Section 5 here), and more in this matter may be discovered in the future, but here I want to examine only some aspects of the notion of a complement-topos.

In (Mac Lane and Moerdijk, 1994, p.161f and p.163) two definitions of a topos are given, which are said to be equivalent. The whole reasoning of my paper hinges on this equivalence and in what follows I assume its validity (although I have some reasons to doubt it, and I plan to examine this in the future). Here I argue that the way co-toposes are defined in the above-mentioned works is, at least from the point of view of category theory, inappropriate. I will not give the complete exposition of the role a dual to intuitionistic logic, or, from an algebraic point of view, a co-Heyting structure, plays in category theory. This is rather still a work in progress and I would like to elaborate on this subject more extensively in future. Here I give

---

[1] About co-toposes see (Mortensen, 1995; 2003; James, 1996; Estrada-González, 2010; 2015), about a dual to intuitionistic logic see e.g. (Goodman, 1981; Czermak, 1977; Urbas, 1996; Kamide, 2003).

only some arguments, I think quite convincing, against the considered notion of a co-topos. It should be stressed that if not stated otherwise I am considering only zero-order logic, i.e. propositional logic.

First, in Section 2, I give some motivations behind the introduction of the concept of a co-topos, I provide the definition of this concept, and, also for this purpose, of the notion of the so-called complement-classifier. Later, assuming the validity of these concepts, I show some of their consequences. In Section 3, I analyze the original concept of a subobject classifier and show how it is connected with the requirement that the functor Sub is representable. In the next section, I critically comment on the notion of a co-topos basing on both the above analysis and some other properties of toposes. In the last section, I conclude and give some further short remarks on the presence and relevance of the co-Heyting structure in toposes and other areas of mathematics.

2. Introduction to the alleged co-toposes

As far as I know, the first definition of a co-topos appeared in (Mortensen, 1995, Chapter 11) (this is suggested in (James, 1996, p.80)), written together with Peter Lavers. After that publication there appeared subsequent publications dealing with co-toposes, see e.g. (James, 1996; Mortensen, 2003; Estrada-González, 2010; 2015).

Let us first see what the motivations for co-toposes are. The logic of toposes is known to be intuitionistic logic (IL), although one should be careful with such a simplification because, as Colin McLarty pointed out, "topos logic coincides with no intuitionist logic studied before toposes" (see McLarty, 1995, p.vii).[2] In one of the

---

[2] In short, higher-order intuitionistic logic of toposes agrees with traditional Heyting's rules of inference for connectives and quantifiers, but the disjunction and existence

approaches, a topos is considered as a generalization of a topological space. The algebra of open sets (of some topological space) is the Heyting algebra and therefore it forms a semantics for IL (cf. Stone, 1938; Tarski, 1938) on topological interpretation of IL). However, a topology can be equivalently specified by the family of closed sets. Incidentally, it is closure operation and closed sets, rather than interior operation and open sets, that were first analyzed: McKinsey and Tarski first considered closure algebra as an algebra of topology (see (McKinsey and Tarski, 1944; 1946)) and the first general and explicit definition of a sheaf on a space was described by Leray in terms of the closed sets of that space (cf. Mac Lane and Moerdijk, 1994, p.1). In this context we can better understand Mortensen's motivation when he writes (see Mortensen, 1995, p.102):

> Specifying a topological space by its closed sets is as natural as specifying it by its open sets. So it would seem odd that topos theory should be associated with open sets rather than closed sets. Yet this is what would be the case if open set logic were the natural propositional logic of toposes. At any rate, there should be a simple 'topological' transformation of the theory of toposes, which stands to closed sets and their logic, as topos theory does to open sets and intuitionism. Furthermore, the logic of closed sets is paraconsistent.

---

properties, which are a traditional part of intuitionism, do not hold in toposes in general (see McLarty, 1990, p.154). Although toposes are intrinsically connected with (higher-order) intuitionistic logic, they were not simply designed to agree with it (see McLarty, 1990, p.152f). Lambek and Scott even claim: "Nothing could have been further from the minds of the founders of topos theory than the philosophy of intuitionism" (see Lambek and Scott, 1994, p.125).

Mortensen gives the following definition of a complement-classifier (see (Mortensen, 1995, p.104f), I keep the exact same wording, changing only F to *false*, a to A, and b to X in order to standardize the notation in this paper):

> A *complement-classifier* for a category $\mathcal{E}$ with terminal object
> 1, is an object Ω together with an arrow *false* : $1 \to \Omega$ satisfying the condition that for every monic arrow $f : A \rightarrowtail X$
> there exists a unique arrow $\overline{\chi}_f$ such that
>
> $$\begin{array}{ccc} A & \stackrel{f}{\rightarrowtail} & X \\ {\scriptstyle !}\downarrow & & \downarrow{\scriptstyle \overline{\chi}_f} \\ 1 & \underset{false}{\longrightarrow} & \Omega \end{array}$$
>
> is a pullback. $\overline{\chi}_f$ is the *complement-character* of f.

Then it is stated what a *complement-topos* is (I quote from (Mortensen, 1995, p.105)):

> An (elementary) *complement-topos* is a category with initial
> and terminal objects, pullbacks, pushouts, exponentiation, and
> a complement classifier.

In what follows, I shall use the term complement-topos (or in short co-topos) meaning the notion as it is described above by Mortensen (the same notion is also used in e.g. (Estrada-González, 2010; 2015)). As an example of a different definition of a co-topos see e.g. (Angot-Pellissier, 2015), where "cotopos" is considered as 'a closed co-Cartesian category with quotient classifier" (see p.189), which seems to be a different notion, although the author suggests he is considering the same notion and makes reference to unpublished work by James and Mortensen.

As I mentioned in Section 1, my aim in this paper is not a thorough analysis of connections between dual to intuitionistic logic and

toposes, nor a comprehensive study of papers concerning co-toposes. I want to scrutinize only some aspects of the notion of a complement-topos. I argue that the arrow $1 \to \Omega$, distinguished (up to isomorphism) by subobject classifier, may not be arbitrarily interpreted, and thus simply renaming it as false is, at least form the point of view of category theory, inappropriate.

Let me first show some of the consequences of such a definition (cf. e.g. Mortensen, 1995; Estrada-González, 2010), assuming for a moment its validity. Having the complement-classifier *false* : $1 \to \Omega$ (which will be denoted also as $\bot$), we define, by analogy with the standard approach,

$$(1) \qquad\qquad true \equiv \top := \overline{\chi}_{o_1} \, ,$$

where o_1 is the only, and always existing, arrow from the initial object, o, to the terminal object, 1; this arrow is a monomorphism. The logical connectives are also defined by analogy with the standard approach[3], but we have to take into account that now $\overline{\chi}_f$ is the complement-characteristic arrow. We have therefore:

$$\neg := \overline{\chi}_\top \, ,$$
$$\smile := \overline{\chi}_{\langle \bot, \bot \rangle} \, ,$$
$$\frown := \overline{\chi}_{\mathrm{Im}[\langle \bot, \mathrm{id}_\Omega \rangle, \langle \mathrm{id}_\Omega, \bot \rangle]} \, ,$$
$$- := \overline{\chi}_e \, ,$$

where $\langle f, g \rangle$ is the product arrow of f and g (with respect to the projections π_1, π_2 on its first and second factor, respectively), $[f, g]$ is the co-product arrow of f and g (with respect to the standard injections),

---

[3] I assume the Reader's familiarity with the standard definition of logical connectives in a topos, which can be found in e.g. (Goldblatt, 2006, p.139).

Imf is an image of f, i.e. the monic of the epi-monic factorization (which, in a topos, exists for any arrow), and e is the equalizer of $\vee$ and π_1. Therefore, in the process of such a dualization: (i) conjunction and disjunction interchange, and (ii) in place of implication we get the so-called pseudo-difference.

As a result $\big(\mathcal{E}(X, \Omega), \sqsubseteq\,\big)$ changes the order, so it becomes a co-Heyting algebra. $\big(\mathsf{Sub}(X), \subseteq\,\big)$ remains a Heyting algebra, as it depends only on the factorization of the appropriate arrows and thus is independent of the (complement-)classifier. In this way, $\big(\mathsf{Sub}(X), \subseteq\big)$ and $\big(\mathcal{E}(X, \Omega), \sqsubseteq\big)$ are no longer isomorphic Heyting algebras. The arrow $\perp$, the one distinguished (up to isomorphism) by the complement-classifier, is now the lowest element of $\big(\mathcal{E}(1, \Omega), \sqsubseteq\big)$.

If we define (as is the usual way) $\mathcal{E} \models \alpha$ if and only if, for every $\mathcal{E}$-evaluation V, it is $V(\alpha) = \top$ then we get a different set of tautologies. Let me analyze one example, for which I use the subscript "S" for the notation of standard toposes, and no subscript for the present case of co-toposes. Because the definition of co-topos, in comparison with the one of topos, assumes the same properties for certain arrows, but gives only different names (or interpretations) to them, we have that e.g. the arrow "$\perp$" (defined as χ_{0_1}) in a standard topos (i.e. $\perp_S$, in our current notation), is the same as $\top$ (defined in (1)), and thus we have $\perp_S = \top$. The situation is analogous for the other arrows. Now, for any (standard) topos $\mathcal{E}$ we have:

$$\mathcal{E} \models_S \sim (\varphi \wedge \sim \varphi) \ .$$

In IL we have: if $\sim\psi = 1$, then $\psi = 0$ (but not *vice versa*). Thus for all $\mathcal{E}$-evaluations we have trivially (assuming, for convenience, that φ is an atomic sentence)

$$V_S(\varphi \wedge \sim \varphi) = \perp_S$$
$$= \wedge_S \circ \langle V_S(\varphi), \neg_S \circ V_S(\varphi) \rangle$$
$$= \smile \circ \langle V(\varphi), \neg \circ V(\varphi) \rangle$$
$$= V(\varphi \vee \sim \varphi)$$
$$= \top.$$

This means that for any co-topos we would have $\mathcal{E} \models \varphi \vee \sim \varphi$.

3. The subobject classifier

In order to analyze the question of the appropriateness of the definition of a complement-classifier and that of a co-topos, let us first comment on the definition of standard subobject classifier and that of topos. A subobject classifier may be defined in the following way:

Definition 1. If $\mathcal{C}$ is a category with a terminal object $\mathbf{1}$, then the **subobject classifier** for $\mathcal{C}$ is an object Ω together with an arrow $true : \mathbf{1} \to \Omega$ such that for every monic $f : A \rightarrowtail X$ there is a unique arrow $\chi_f : X \to \Omega$ which makes the following diagram a pullback:

$$
\begin{array}{ccc}
a & \xrightarrow{\;f\;} & d \\
{\scriptstyle !}\big\downarrow & & \big\downarrow{\scriptstyle \chi_f} \\
\mathbf{1} & \xrightarrow[\;true\;]{} & \Omega.
\end{array}
$$

The important question for us is whether (in the context of a topos)

(A) the word "true" in the above definition is a meaningful and non-removable part of it,

and, in such case, we should add a comment about what does "true" mean already at this level (without having yet defined any order on $\mathcal{E}(1, \Omega)$); or whether

(B) instead of "...together with an arrow *true* : $1 \to \Omega$ such that ..." we could just have "...together with a certain arrow $\eta : 1 \to \Omega$ such that ..." (where η is just a label added for convenience of referring to it).

Of course, in both cases the distinguished arrow (*true* or η) is defined up to an isomorphism, but the question is whether the interpretation of it as true is an additional feature of the arrow that we are assuming (option (A)), or it is the consequence of the definition (option (B)). In the latter case we can *post factum* add this name in the very definition and obtain the standard textbook definition.

In my opinion, (B) is the only correct option. A subobject classifier is considered mainly in order to define a topos as a Cartesian closed category with a subobject classifier. However, a topos may also be defined without any reference to truth. Namely, from the equivalence of the two definitions of the topos (Mac Lane and Moerdijk, 1994, p.161f and 163), instead of a subobject classifier we can assume an object Ω and for each object X an isomorphism

$$\mathsf{Sub}(X) \cong \mathsf{Hom}(X, \Omega),$$

natural in X. In other words, the functor Sub is required to be representable (the subobject classifier being its representing object).

Representability of Sub

Let us briefly see how representability of the functor Sub is connected with the standard definition of a subobject classifier. Repre-

sentability of Sub means that there is a natural isomorphism between the Sub functor and the contravariant Hom-functor $\mathsf{Hom}(-,\Omega)$ (with Ω being the representing object), which we shall denote as β. The situation can be pictured as

(2)
$$
\mathcal{E}^{op} \underset{\mathsf{Sub}}{\overset{\mathsf{Hom}(-,\Omega)}{\Longrightarrow}} \mathsf{Set} \; .
$$

Now, from the Yoneda lemma we know that

$$
Nat\big(\mathsf{Hom}(-,\Omega), \mathsf{Sub}\big) \cong \mathsf{Sub}(\Omega) \, ,
$$

and every natural transformation α from $\mathsf{Hom}(-,\Omega)$ to Sub is completely determined by an element of $\mathsf{Sub}(\Omega)$ (i.e., by a subobject of Ω), which is $\alpha_\Omega(\mathrm{id}_\Omega)$. Namely, for any object X of $\mathcal{E}$, a component α_X is an arrow between sets $\mathsf{Hom}(X,\Omega)$ and $\mathsf{Sub}(X)$, and its action on any arrow $g : X \to \Omega$ is given by

$$
\alpha_X(g) = \mathsf{Sub}(g)\big(\alpha_\Omega(\mathrm{id}_\Omega)\big) \, ,
$$

where $\mathsf{Sub}(g)$ is an arrow in Set between $\mathsf{Sub}(\Omega)$ and $\mathsf{Sub}(X)$, which takes a monic and by pulling it back along g gives another monic. $\alpha_\Omega(\mathrm{id}_\Omega)$ being a subobject of Ω can be denoted as a monic $\Omega_0 \rightarrowtail \Omega$. Then, the action of α_X can be described as follows: for any $g : X \to \Omega$, $\alpha_X(g)$ gives a subobject of X, let us denote it as $A \rightarrowtail X$, which is given by the following pullback:

$$
\begin{array}{ccc}
A & \rightarrowtail & X \\
\downarrow & & \downarrow{\scriptstyle g} \\
\Omega_0 & \rightarrowtail & \Omega \, .
\end{array}
$$

It can be shown that if α is a natural isomorphism, then the corresponding subobject $\alpha_\Omega(\text{id}_\Omega)$ is not any subobject of Ω, but precisely a global element of it, i.e., an arrow $1 \to \Omega$ (see e.g. Mac Lane and Moerdijk, 1994, part of the proof on p.33f).

From the representability of Sub we know that among all the natural transformations $\{\alpha\}$ there is at least one that is actually a natural isomorphism, which we have denoted as β (see (2)). We now know that such β is completely determined by $\beta_\Omega(\text{id}_\Omega)$, being a global element of Ω, which we shall denote as $\eta : 1 \to \Omega$. In this way, for any object X we have an isomorphism (a bijection) β_X between $\mathsf{Hom}(X, \Omega)$ and $\mathsf{Sub}(X)$ given by a pullback (to avoid confusion with previous notation, we may now denote the arrow from $\mathsf{Sub}(X)$ as f, and the corresponding arrow from $\mathsf{Hom}(X, \Omega)$ as χ_f)

$$\begin{array}{ccc} A & \xrightarrow{\ f\ } & X \\ {\scriptstyle !}\downarrow & & \downarrow{\scriptstyle \chi_f} \\ 1 & \xrightarrow{\ \eta\ } & \Omega \,, \end{array}$$

which is precisely the bijection between characteristic morphisms and subobjects, as assumed in the definition of subobject classifier.

The above considerations suggest that the interpretation of the arrow $\eta : 1 \to \Omega$ as *true* is not an additional feature the arrow has to fulfill, but it follows, as we shall see more clearly in the next section, from the role this arrow plays in the structure of a topos, as option (B) (on page 119) points out.

4. Comments on the notion of a co-topos

As mentioned in Section 1, I assume in this paper that the two definitions of a topos as given in (Mac Lane and Moerdijk, 1994, p.161f

and 163) are equivalent (as stated in this book). In my opinion, this equivalence should be scrutinized, and I plan to examine this in future work, but here I take it for granted. The whole argumentation of my paper hinges on this equivalence, and so, if it is false, the conclusions might also (but do not have to) be false. Nevertheless, I think that a lot of observations in this paper are still valuable (at least as valid conditional reasoning based on some assumptions).

First, a topos may be defined without any reference to the notion of truth, thus making impossible the dualization suggested in the approach under consideration.

Moreover, we have to pay attention to the internal structure of a topos. From the categorical point of view, it is the internal structure that plays a major role, and if the category theory is treated as a foundation of mathematics, then the internal structure is the only one we have. Boolean algebras, which are the Tarski–Lindenbaum algebras of the classical logic, are self-dual in the sense that they preserve the property of being Boolean after changing the order. Such a change of the order swaps the top element (truth) with the bottom (false), and the conjunction (meet) with the disjunction (join) (nota bene, this property of Boolean algebras does not imply that the notions of truth and falsity are utterly interchangeable for, say, the working mathematician). For Heyting algebras, which are the Tarski–Lindenbaum algebras of the intuitionistic logic (or, more precisely, for intermediate logics), the situation is not the same. The dual of a Heyting algebra is a co-Heyting or Brouwer algebra associated with a dual to intuitionistic logic. Now, it is a Heyting algebra structure that is fundamentally[4] and

---

[4] By this I mean the universality of this structure in all toposes. A co-Heyting structure can also be present internally in toposes but only in some of them (e.g. in Boolean toposes, but, as we shall soon see, there is a much larger family of toposes that also exhibit a co-Heyting structure).

internally present in all toposes. We have already mentioned that the poset $\big(\mathsf{Sub}(X), \subseteq\big)$ is a Heyting algebra for any object X in any topos (independently of the (complement-)classifier). This is an example of the external Heyting algebra, as the set $\mathsf{Sub}(X)$ may not be an object of a given topos. There is, however, an internal version of this statement, which I formulate in the form of a proposition without proof, based on (Mac Lane and Moerdijk, 1994, p.201):

Proposition 1. For any object X in any topos $\mathcal{E}$, the power object PX (or equivalently the exponential Ω^X) is an internal Heyting algebra. (In particular, so is the subobject classifier $\Omega = P1$.) For each X in $\mathcal{E}$ the internal structure on Ω makes $\mathsf{Hom}(X, \Omega)$ an external Heyting algebra so that the canonical isomorphism

$$(3) \qquad\qquad \mathsf{Sub}(X) \cong \mathsf{Hom}(X, \Omega)$$

is an isomorphism of external Heyting algebras.

In the proof of the theorem on which this proposition is based, it is shown that $\eta : 1 \to \Omega$ is the top element of the object Ω taken as an internal Heyting algebra.[5] On the 'external level', assuming (3), we may easily see that η is also the top element in the external Heyting algebra $\mathsf{Hom}(1, \Omega)$. In order to achieve this let us note that $\eta : 1 \to \Omega$, having the codomain Ω, is the characteristic arrow for some subobject of its domain, i.e. of 1. If we pull η back along itself, we get that $\eta = \chi_{\mathrm{id}_1}$. Now, because id_1 is the top element of the Heyting algebra $\mathsf{Sub}(1)$, by means of (3) we get our result that η is the top element in

---

[5] The proof may be found on pages 201f of (Mac Lane and Moerdijk, 1994). The authors denote the η arrow as *true* from the very beginning, but, as I have argued in the previous section and as the proof under consideration shows, this notation (or interpretation) is the consequence of its definition in the context of a topos.

the external Heyting algebra $\mathsf{Hom}(1, \Omega)$, which is understood as the algebra of truth-values of a topos and plays a special role in its logical structure.

On the basis of the above considerations, I argue that the distinguished arrow $\eta : 1 \to \Omega$ cannot be interpreted arbitrarily. Therefore, the opinion that: "To dualize, simply rename T with F, and relabel the classifier arrow chi$_\mathsf{f}$ as chi-bar$_\mathsf{f}$" (see Mortensen, 2003, p.259) cannot be considered valid. On the contrary, no matter how we denote the distinguished arrow, by T, F, η or by anything else, if it obeys the required conditions, then it will play the role of the top element of both, the internal Heyting algebra Ω, as well as the external Heyting algebra $\mathsf{Hom}(1, \Omega)$, and thus the role of truth. At the same time, let me remind that the subobject classifier, together with its *true* arrow, is defined (only) up to isomorphism, which is a common (and proper) feature in category theory.

5. Summary and further remarks

In this paper, I have given some reasons why, assuming the equivalence of the two definitions of a topos in (Mac Lane and Moerdijk, 1994, p.161f and 163), I think the way co-toposes are defined in the above-considered works (Mortensen, 1995; 2003; James, 1996; Estrada-González, 2010; 2015) is not appropriate, at least within the conceptual framework of category theory.

This does not mean that there is no place for co-Heyting structures and possibly some manifestations of dual to intuitionistic logic in toposes. On the contrary, apart from the whole class of Boolean toposes (which are in this respect a trivial instance of the presence of a co-Heyting structure), a vast class of toposes is known that

exhibit the co-Heyting structure of subobjects (and thus bi-Heyting, because, as I have mentioned, the Heyting structure is always present). Lawvere (1991) very concisely pointed out that "In any presheaf topos (and more generally any essential subtopos of a presheaf topos), the lattice of all subobjects of any given object is [. . . an] example of a co-Heyting algebra" (Lawvere, 1991, p.280). This line of reasoning was pursued by Reyes and Zolfaghari in (Reyes and Zolfaghari, 1996) where they proved the above Lawvere's assertion and introduced a new approach to the modal operators, based on the existence of the two negations for bi-Heyting structures. The bi-Heyting algebras and their logics were first studied, according to my knowledge, by Cecylia Rauszer (1974a,b), where she uses the name semi-Boolean algebra for bi-Heyting algebra, and Heyting–Brouwer logic, or, in short, H-B logic, for its logic.

It seems that the co-Heyting structure is also connected with a notion of 'boundary', since now the intersection of an element of a co-Heyting algebra with its co-Heyting complement (negation) may not be empty (zero). The 'boundary' defined in this way may also exhibit the Leibniz product rule and perhaps be developed into some further geometric structures (cf. Lawvere, 1991; Majid, 2008, especially pp.123–126).

Bibliography

Angot-Pellissier, R., 2015. The Relation Between Logic, Set Theory and Topos Theory as it Is Used by Alain Badiou. In: A. Koslow and A. Buchsbaum, eds. *The Road to Universal Logic, vol. II*, *Studies in Universal Logic*. Cham: Birkhäuser, Springer, pp.181–200.

Czermak, J., 1977. A remark on Gentzen's calculus of sequents. *Notre Dame Journal of Formal Logic* [Online], 18(3), pp.471–474. Available at: https://doi.org/10.1305/ndjfl/1093888021.

Estrada-González, L., 2010. Complement-topoi and dual intuitionistic logic. *Australasian Journal of Logic* [Online], 9, pp.26–44. Available at: https://doi.org/10.26686/ajl.v9i0.1819.

Estrada-González, L., 2015. The evil twin: the basics of complement-toposes. In: J.-Y. Beziau, M. Chakraborty and S. Dutta, eds. *New Directions in Paraconsistent Logic*. New Delhi: Springer India, pp.375–425.

Goldblatt, R., 2006. *Topoi: The Categorial Analysis of Logic*, *Dover Books on Mathematics*. Mineola, N.Y.: Dover Publications.

Goodman, N.D., 1981. The logic of contradiction. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* [Online], 27(8-10), pp.119–126. Available at: https://doi.org/10.1002/malq.19810270803 [visited on 3 December 2020].

James, W., 1996. *Closed Set Logic in Categories* [Online]. PhD Thesis. Department of Philosophy, The University of Adelaide. Available at: https://digital.library.adelaide.edu.au/dspace/handle/2440/18746 [visited on 3 December 2020].

Kamide, N., 2003. A note on dual-intuitionistic logic. *Mathematical Logic Quarterly* [Online], 49(5), pp.519–524. Available at: https://doi.org/10.1002/malq.200310055 [visited on 3 December 2020].

Lambek, J. and Scott, P., 1994. *Introduction to Higher-Order Categorical Logic*, *Cambridge Studies in Advanced Mathematics*. Cambridge; New York; Melbourne: Cambridge University Press.

Lawvere, F.W., 1991. Intrinsic co-Heyting boundaries and the Leibniz rule in certain toposes. In: A. Carboni, M.C. Pedicchio and G. Rosolini, eds. *Category Theory*. Berlin, Heidelberg: Springer, pp.279–281.

Mac Lane, S. and Moerdijk, I., 1994. *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*, *Universitext*. New York [etc.]: Springer-Verlag.

Majid, S., 2008. Quantum spacetime and physical reality. In: S. Majid, ed. *On Space and Time* [Online]. Cambridge: Cambridge University Press, pp.56–140. Available at: https://doi.org/10.1017/CBO9781139644259. 003.

McKinsey, J.C.C. and Tarski, A., 1944. The algebra of topology. *Annals of Mathematics, Second Series* [Online], 45, pp.141–191. Available at: https://doi.org/10.2307/1969080.

McKinsey, J.C.C. and Tarski, A., 1946. On closed elements in closure algebras. *Annals of Mathematics, Second Series* [Online], 47(1), pp.122–162. Available at: https://doi.org/10.2307/1969038.

McLarty, C., 1990. Book review: John Bell, "Introduction to toposes and local set theory". *Notre Dame Journal of Formal Logic* [Online], 31(1), pp.150–161. Available at: https://doi.org/10.1305/ndjfl/1093635339 [visited on 3 December 2020].

McLarty, C., 1995. *Elementary Categories, Elementary Toposes*. Oxford: Oxford University Press.

Mortensen, C., 1995. *Inconsistent Mathematics*, *Mathematics and Its Applications* vol. 312. Dordrecht [etc.]: Kluwer Academic Publishers.

Mortensen, C., 2003. Closed Set Logic. In: R. Brady, ed. *Relevant logics and their rivals. Vol. 2*, *Western Philosophy Series*. Aldershot; Burlington: Ashgate, pp.254–262.

Rauszer, C., 1974a. A formalization of the propositional calculus of H-B logic. *Studia Logica* [Online], 33(1), pp.23–34. Available at: https://doi. org/10.1007/BF02120864.

Rauszer, C., 1974b. Semi-Boolean algebras and their applications to intuitionistic logic with dual operations. *Fundamenta Mathematicae* [Online], 83, pp.219–249. Available at: https://doi.org/10.4064/fm-83-3-219-249 [visited on 2 December 2020].

Reyes, G.E. and Zolfaghari, H., 1996. Bi-Heyting algebras, toposes and modalities. *Journal of Philosophical Logic* [Online], 25(1), pp.25–43. Available at: https://doi.org/10.1007/BF00357841.

Stone, M.H., 1938. Topological representations of distributive lattices and Brouwerian logics. *Časopis pro Pěstování Matematiky a Fysiky* [Online], 67(1), pp.1–25. Available at: https://doi.org/10.21136/CPMF.1938.124080 [visited on 2 December 2020].

Tarski, A., 1938. Der Aussagenkalkül und die Topologie. *Fundamenta Mathematicae* [Online], 31, pp.103–134. Available at: https://doi.org/10.4064/fm-31-1-103-134 [visited on 2 December 2020].

Urbas, I., 1996. Dual-intuitionistic logic. *Notre Dame Journal of Formal Logic* [Online], 37(3), pp.440–451. Available at: https://doi.org/10.1305/ndjfl/1039886520.

No-signaling in topos formulation and a common ontological basis for classical and non-classical physical theories

Marek Kuś

Warsaw University of Technology,
International Center for Formal Ontology

Abstract

Starting from logical structures of classical and quantum mechanics we reconstruct the logic of so-called no-signaling theories, where the correlations among subsystems of a composite system are restricted only by a simplest form of causality forbidding an instantaneous communication. Although such theories are, as it seems, irrelevant for the description of physical reality, they are helpful in understanding the relevance of quantum mechanics. The logical structure of each theory has an epistemological flavor, as it is based on analysis of possible results of experiments. In this note we emphasize that not only logical structures of classical, quantum and no-signaling theory may be treated on the same ground but it is also possible to give to all of them a common ontological basis by constructing a "phase space" in all cases. In non-classical cases the phase space is not a set, as in classical theory, but a more general object obtained by means of category theory, but conceptually it plays the same role as the phase space in classical physics.

Keywords

no-signalling, topos, quantum logic.

Ontological assumptions of classical and quantum picture of the world differ significantly, despite the fact that both descriptions attempt to grasp the structure of the same reality. The fact that they offer a glance from two different points of view—macroscopic and microscopic, should not interfere with a pragmatic request that, at least to compare both theories, we should have in both of them some common 'elements of reality', e.g., some observables that pertain to the same physical quantities on both levels, like 'position', 'momentum', 'angular momentum' etc. An alternative (let's call it Kuhnian) approach would be to deny connections between concepts used in both theories pointing to (seemingly) the same properties of physical systems, (e.g., Newtonian and Einsteinian mass (Kuhn, 1970)). The latter approach is not particularly attractive from the point of view of a practicing physicist. It leaves no room for many useful and fruitful procedures as e.g., a 'semiclassical/classical approximation'.

Classical systems are described in terms of a phase space, usually a differential manifold with some additional (symplectic/Poisson/metric) structures. Observables, i.e. physical quantities that we can measure, or, in general to which we can ascribe certain numerical values characterizing the observed system, are functions on the phase space. Values of observables, like positions, momenta, energies, angular momenta etc., are some intrinsic properties of physical systems (particles, ensembles of particles, rigid bodies, etc.). They can change in time, but are properties that are possessed by systems alone and do not depend on whether or not they are actually measured at a particular moment. At least in principle, we can measure them without disturbing the system. Consequently, measurements can be performed in an arbitrary order, or even simultaneously, and provide the same results. We can thus pose questions about *exact* values of, say, the position and the momentum of a particle. Usually, however, due to

e.g., inaccuracies of measurements we inquire into the probability that our particle is in a certain subset of the phase space. Such a probability is determined by the volumes of the relevant subsets. Physical states can thus be identified with probability distributions on (measurable) subsets of the phase space. Mean values (results of experiments) can be calculated using these distributions.

In quantum mechanics we do not have a clear notion of a 'phase-space' in the form of a manifold. A backbone structure is provided by a Hilbert space $\mathcal{H}$, observables are identified with self-adjoined operators, and states with non-negative, trace-class operators (density matrices). We may ascribe to each system some properties that pretend to be quantum analogues of classical ones, like positions, momenta, angular momenta, energies, etc. (and some others that seem to be of a purely quantum mechanical nature, like spin, isospin, strangeness, hypercharge etc.). However, they are no longer *intrinsic* in the classical sense. They are not 'carried' by a system during its evolution, rather they are 'brought to life' by an act of measurement.

Hence, it is hard to find a unifying ontological basis for classical and quantum physics. Fundamental elements of physical reality, as positions, momenta, angular momenta, etc. have different ontological status in both theories. For everyday physical practice this does not pose any clear and present danger. Ultimately, physics is an experimental science. It aims at answering experimental questions about outcomes of measurements putting emphasis on the epistemology, at the price of moving apart, or even totally discarding ontological issues.

An attempt to unify classical and quantum physics on common epistemological ground goes back to Birkhoff and von Neumann (1936) in form of the so-called quantum logic. The main idea is to analyze the structure of elementary experimental question/propositions

about a system. In classical physics, elementary propositions can be reduced to statements that values of observed quantities (positions, momenta) belong to a certain subset of the phase space. The logical structure of the set of such propositions, determined by the rules concerning their negations, conjunctions and disjunctions, isomorphically reflects the Boole algebra structure of the set of (measurable) subsets of the phase space. One of the fundamental features of the resulting lattice[1] is the distributivity law, allowing for the distribution of conjunctions over disjunctions and *vice versa*.

In quantum mechanics elementary propositions concern positions of state vectors (characterizing a state of a system) with respect to eigenspaces of observables. As in the classical case we can ask composite questions corresponding to conjunctions and disjunctions. However, the ensuing logical structure is no longer distributive. The logic of a system described by a Hilbert space $\mathcal{H}$ is represented by the orthomodular lattice of closed subspaces in $\mathcal{H}$. The involution sending a subspace to its orthogonal complement represents logical negation, still as in the classical case, satisfies the law of an excluded middle: measuring the spin of an electron will yield either 'up' or 'down'. The resulting lattice is, however, non-distributive: x-spin up does not imply x-spin up and z-spin up or x-spin up and z-spin down. Having the lattice stand for the logic of the system, one derives its probability theory, where states assign probabilities to elements of the lattice, respecting the underlying structure (order and complementation). These states turn out to coincide with the usual density matrices by the Gleason theorem (Gleason, 1957).

---

[1] A partially ordered set in which every two elements have a unique supremum—the set-theoretical sum on the level of subsets and conjunction on the level of proposition, and infimum—the set-theoretical intersection of subsets and disjunction, respectively.

Having two examples of different theories pertaining to the same physical reality, we are tempted to think about other similar constructions. From what we know now, it is hard to construct a successful theory that e.g., supersedes quantum mechanics (Aaronson, 2004). Instead we can identify some common epistemic structures in classical and quantum mechanics encoded in logics of both theories, i.e. the logical structures of the sets of their propositions, and try to construct similar, reasonable theories.

One of such attempts was presented by Popescu nad Rohrlich (1994) in the form of so called no-signaling boxes. They started from a paradigmatic correlation experiment, that can be performed both on classical and quantum level depicted schematically in Fig.1. The model is supposed to describe the most elementary system composed of two separated subsystems. We can think of inputs as observables that we choose to measure, and outputs as the results of measurements. In the simplest case we have two observables, encoded (labeled) by 0 and 1, each of which can take two values 0 and 1. Performing multiple measurements we will obtain a sequence of outcomes allowing us to determine the relative frequency $P(\alpha\beta|ab)$ of getting any pair of outputs $\alpha\beta \in \{-1, 1\} \times \{-1, 1\}$, given any pair of inputs $ab \in \{0, 1\} \times \{0, 1\}$.

In order to have a legitimate interpretation of $P : \{-1, 1\} \times \{-1, 1\} \times \{0, 1\} \times \{0, 1\} \to \mathbb{R}$ in terms of probability, it should fulfill the following requirements,

1. $0 \leq P(\alpha\beta|ab) \leq 1$ (positivity),
2. $\sum_{\alpha\beta} P(\alpha\beta|ab) = 1$ (normalization),
3. $\sum_{\alpha} P(\alpha\beta|ab) = \sum_{\alpha} P(\alpha\beta|cb); \quad \sum_{\beta} P(\alpha\beta|ab) = \sum_{\beta} P(\alpha\beta|ac)$ (no-signaling).
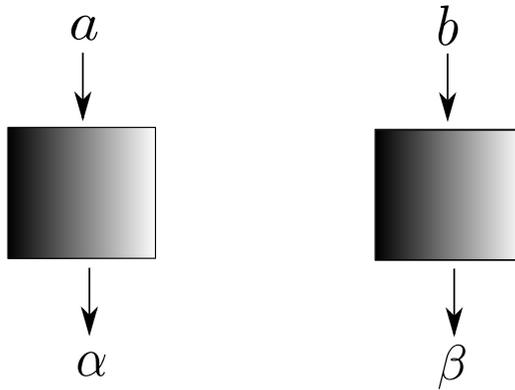
Figure 1: A two-system correlation experiment

The last property, non-signaling, is supposed to encode the principle of relativistic causality, i.e. what happens in one box does not influence the other, obeyed by spatially separated subsystems. We will refer to this particular example of non-signaling boxes as the (2, 2)-box world.

For a particular instance of (2,2)-box world we may postulate concrete values of $P(\alpha\beta|ab)$ fulfilling 1.-3. Popescu and Rohrlich (1994) proposed the following,

$$
(1) \qquad P(\alpha\beta|ab) = \begin{array}{c} \\ -- \\ -+ \\ +- \\ ++ \end{array}\begin{array}{c} 00 \quad 01 \quad 10 \quad 11 \\ \begin{pmatrix} 1/2 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 1/2 \\ 1/2 & 1/2 & 1/2 & 0 \end{pmatrix} \end{array}.
$$

Usually $P(\alpha\beta|ab) \neq P(\alpha|a)P(\beta|b)$, where $P(\alpha|a)$ and $P(\beta|b)$ are one-particle probability distributions of measurements results of a and b. One introduces thus the correlations,

$$(2) \qquad \langle ab \rangle = \sum_{\alpha,\beta \in \{-1,1\}} \alpha\beta P(\alpha\beta|ab).$$

It can be now checked that in the classical and quantum cases the following combination of correlations

$$(3) \qquad S := |\langle a_1 b_1 \rangle + \langle a_2 b_1 \rangle + \langle a_2 b_2 \rangle - \langle a_1 b_2 \rangle|$$

that, in principle, can achieve the maximal value of 4 (each of the term is not larger than 1 in absolute value) is further restricted. Classically $|S| \leq 2$ (Bell (1964) inequality in the CHSH form (Clauser et al., 1969)) whereas in quantum mechanics $|S| \leq 2\sqrt{2}$ (Tsirelson's bound (Cirel'son, 1980)). For the Popescu-Rohlich box (1) we have $S = 4$.

In the quantum logic approach the elementary admissible questions are,

- 'Does our system belong to a (measurable) subset of the phase-space?' (classical mechanics);
- 'Is the result of measuring the projection on a closed subspace of the Hilbert space of the system equal to 1?' (quantum mechanics);
- 'Does measuring a on a subsystem gives an outcome α?' (Popescu-Rohrlich).[2]

One can now combine elementary questions by the conjunction and disjunction and negate them. As already said the resulting lattice is a Boolean algebra in the classical case and so-called orthomodular

---

[2] In fact we should ask about an outcome of a joint measurement of a in the first subsystem and b in the second one.

lattice which is non-distributive[3] (Birkhoff and von Neumann, 1936). For the Popescu-Rohrlich box (1) the appropriate one is that of a so-called orthomodular poset (for details see (Tylec and Kuś, 2015; Tylec and Kuś, 2018)). The resulting structure has some common features with quantum mechanics, e.g., the truth of the alternative $p \vee q$ of statements p and q does not imply that p is true or q is true (the 'Schrödinger cat' paradox). Moreover, as in quantum mechanics measurements are, in general, destructive. Performing a measurement changes irreversibly a state of a system and does not allow its exact reconstruction (Tylec and Kuś, 2015). On the other hand, from some points of view, Popescu-Rohrlich boxes are 'more classical' than 'quantum mechanical', since in both theories the Heisenberg uncertainty relations are not fulfilled (Tylec and Kuś, 2015).

As already emphasized, the quantum logic approach has a rather epistemic flavor. We concentrate on learning system's properties from observations/measurements. Instead, the category theory seems to provide a path to 'restore a common ontology' for all considered theories. One can cast the program into a sequence of goals.

- Find a well behaved phase space for quantum and a no-signaling systems.
- Find a logical structure of the set of propositions.
- Reproduce the probabilistic properties of the theory.

Since the well understood ontological assumptions of classical theory are, as argued above, connected to the notion of the phase-space, it would be desirable to construct phase-spaces for non-classical theories in such a way that they resemble 'as much as possible' the classical one. This opens a possibility to interpret quantum mechanics and no-signaling theories in a realistic sense, like we do it in classical physics,

---

[3] Conjunctions do not distribute over alternatives and *vice versa*.

where we can assign truth values to propositions without reference to measurements. Hence, we can maintain that propositions refer to some 'real objects' or 'real properties' independent of observers and measurements. This is exactly what I call a 'restoration of a common ontology' for all theories considered here.

As it will be clear, we have to be ready to pay some price. The resulting 'logic' of a phase-space need not to be a Boolean one, so the *tertium non datur* principle need not to be fulfilled. Nevertheless the probabilities calculated on such a 'logic' (just as the probabilities in classical physics calculated on the Boolean logic of subsets of the phase space) will reproduce quantum mechanical and no-signaling results. Moreover, what separates logics of non-classical theories from the classical one, i.e. its non-distributivity, responsible for nearly all 'paradoxes' of quantum mechanics (and no-signaling theories) will be avoided.

For quantum mechanics, this can be done within a program presented by Isham and Döring (2008a,b,c,d), in terms of the category theory or, more specifically, the topos theory. In what follows I will employ a slightly different approach (Wolters, 2013), using the same mathematical apparatus of categories and topoi, proposed originally by Heunen, Landsman, and Spitters (2009; 2011) called 'Bohrification'[4]. In (Gutt and Kuś, 2016) we extended the construction to no-signaling boxes. Finding an analog of a classical phase-space is, in a certain sense, the principal goal.

Let me shortly describe main mathematical ingredients of the above outlined approach. A category is a structure consisting of 'objects' connected by 'arrows'. From the definition we may compose the arrows (an arrow from an object to a second one followed by another arrow from the second object to a third one) and the composition is

---

[4] For the explanation of the chosen name consult the cited papers of Heunen *et al.*

associative. A very intuitive example of a category is the category of sets, **Set**, where the objects are sets, and the arrow connecting a set A to a set B is a function from A to B.[5] A topos is a category with some additional properties chosen in a way that results in a certain generalization of **Set**. The appropriate formal definition and all needed technical details can be found in one of numerous books on categories and topoi (e.g., Goldblatt, 2014) and will not be invoked here.[6] What is important is that basic constructions involving sets, like e.g., exponentiation A^B, i.e, the set of all functions from A to B, have their equivalents for topoi, and that (by definition) each topos is equipped with the, so-called, sub-object classifier. The latter is a special object of the considered category, the meaning of which can be understood by taking again as an example a set S and its subsets. We can express the fact that A is a subset of S by considering a characteristic function of A, i.e., the function from S to the set $\{0, 1\}$ which takes the value 1 on $s \in S$ if $a \in A$ and the value 0 otherwise. The two-element set $\{0, 1\}$ is the 'subobject classifier' making **Set** a topos. The fact that the subobject classifier is a two-element set (we can refer to value 1 as 'true' and to 0 as 'false') is clearly strictly connected to the Boolean structure of the algebra of subsets (and to the 'logic' of a classical phase-space described previously). In general topos the subobject-classifier need not be a two-element set, but some more general object in the category. As a consequence the 'logic' of a topos need not be Boolean any longer, but it is a so-called Heyting algebra, which is distributive, but, in general, the principle

---

[5] From the point of view of the category theory **Set** is not so trivial, since the collection of its objects is not a set—such a category is called a 'large category'. As an example of a 'small category' employing sets as objects we can take a category of open subsets of some topological space ('a classical phase-space' in the sense described previously).

[6] For a clear exposition of applications in quantum physics (Flori, 2013) and (Flori, 2018) are, probably, the best choice.

of excluded middle is no longer valid. A nice example is a lattice of open subsets of a topological space. It is partially ordered by the set-theoretic inclusion and, as in the case of subsets of some set, we have here the ordinary set-theoretical algebraic operations of sum and intersection corresponding to alternative and conjunction, but since the set-theoretical completion of an open set is not open. we have to take the interior of the completion to achieve the proper representation of negation. But then the principle of excluded middle is not fulfilled, since the sum of an open set and the interior of its completion is not the whole space, in contrast to the case of subsets of a set S, where a subset A and its completion $S - A$ sum up to the whole S (*tertium non datur*). Obviously the lattice remains distributive, since sums distribute over intersections and *vice-versa*.

Roughly speaking (some technical refinements are needed to end up with a proper result, see the cited papers of Heunen *et al.*), one finds a well-behaved phase space by constructing an 'internal logic' (a Heyting algebra) of some topos and identifying this very topos with the looked-for phase-space. In the construction of Heunen *et al.*, concerning quantum mechanics, the starting point is the set $\mathscr{C}$ of commuting subalgebras ('contexts') of observables on the Hilbert space $\mathcal{H}$ instead of the set of all orthogonal projections acting on it. Each such context has a well defined physical meaning as a set of compatible measurements. The set of contexts partially ordered by inclusion is treated as a category with contexts as objects and arrows as inclusions. The construction of the 'internal logic' goes through several technical steps (again, for details consult (Heunen, Landsman and Spitters, 2009; Heunen, Landsman, Spitters and Wolters, 2011)) ending with a topos which is identified with the 'phase-space' we were looking for. In a natural way one defines also states of a systems,

and what is most important, a method of calculating probabilities of outcomes of experiments (reproducieng the quantum mechanical results).

In (Gutt and Kuś, 2016) the above construction was extended to no-signaling boxes. Here we do not have a natural Hilbert space structure as a playground, consequently thus a natural analogue of commuting algebras is lacking. Nevertheless, one can consistently define contexts (situations where measurements are compatible). As a result it was possible to define an appropriate phase-space Σ, states of the box-world and probabilities on Σ reproducing the correlation structure in it. Hence, in terms of category theory one can 'restore the ontology' (phase-space) also for the Popescu-Rorlich boxes. The resulting phase space is, again, not a set, but a more general object, namely a particular topos. Thus all three theories are put on the same level with phase spaces described by appropriate topoi. This suggests a hypothesis that the approach is, in some sense, universal and applicable also to other, possible 'generalizations of quantum mechanics' of the whole procedure.

For quantum mechanical systems other approaches of generalizing the classical phase-space descriptions were considered in the past (and are still in use). The most popular employs the classical phase-space parameterized by positions and momenta at the price of lack of possibility to define in a consistent way positive-definite probability functions reproducing quantum mechanical results (using instead so called 'quasiprobabilities' like, e.g., the Wigner function). It is not clear how one relates this approach to the topos-theoretic one described above. For no-signaling theories it is even harder, since no starting point (a classical phase-space on which some quasiprobabilities are defined) is easy to identify.

Bibliography

Aaronson, S., 2004. Is quantum mechanics an island in theoryspace? *arXiv:quant-ph/0401062* [Online]. Available at: http://arxiv.org/abs/quant-ph/0401062 [visited on 30 November 2020].

Bell, J.S., 1964. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1, pp.195–200.

Birkhoff, G. and von Neumann, J., 1936. The logic of quantum mechanics. *Annals of Mathematics*, 37, pp.823–843.

Cirel'son, B.S., 1980. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics* [Online], 4(2), pp.93–100. Available at: https://doi.org/10.1007/BF00417500 [visited on 30 November 2020].

Clauser, J.F., Horne, M.A., Shimony, A. and Holt, R.A., 1969. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15), pp.880–884.

Döring, A. and Isham, C.J., 2008a. A topos foundation for theories of physics: I. Formal languages for physics. *Journal of Mathematical Physics* [Online], 49(5), p.053515. Available at: https://doi.org/10.1063/1.2883740 [visited on 30 November 2020].

Döring, A. and Isham, C.J., 2008b. A topos foundation for theories of physics: II. Daseinisation and the liberation of quantum theory. *Journal of Mathematical Physics* [Online], 49(5), p.053516. Available at: https://doi.org/10.1063/1.2883742 [visited on 30 November 2020].

Döring, A. and Isham, C.J., 2008c. A topos foundation for theories of physics: III. The representation of physical quantities with arrows. *Journal of Mathematical Physics* [Online], 49(5), p.053517. Available at: https://doi.org/10.1063/1.2883777 [visited on 30 November 2020].

Döring, A. and Isham, C.J., 2008d. A topos foundation for theories of physics: IV. Categories of systems. *Journal of Mathematical Physics* [Online], 49(5), p.053518. Available at: https://doi.org/10.1063/1.2883826 [visited on 30 November 2020].

Flori, C., 2013. *A First Course in Topos Quantum Theory*, *Lecture Notes in Physics* 868. Berlin; Heidelberg: Springer.

Flori, C., 2018. *A Second Course in Topos Quantum Theory*, *Lecture Notes in Physics* 944. Cham, Switzerland: Springer.

Gleason, A.M., 1957. Measures on the closed subspaces of a Hilbert space. *Journal of Mathematics and Mechanics* [Online], 6(6), pp.885–893. Available at: https://www.jstor.org/stable/24900629 [visited on 25 November 2020].

Goldblatt, R., 2014. *Topoi: the Categorial Analysis of Logic.* Amsterdam: Elsevier Science.

Gutt, J. and Kuś, M., 2016. Non-signalling boxes and Bohrification. *arXiv:1602.04702 [math-ph, physics:quant-ph]* [Online]. Available at: http://arxiv.org/abs/1602.04702 [visited on 30 November 2020].

Heunen, C., Landsman, N.P. and Spitters, B., 2009. A topos for algebraic quantum theory. *Communications in Mathematical Physics* [Online], 291(1), pp.63–110. Available at: https://doi.org/10.1007/s00220-009-0865-6 [visited on 30 November 2020].

Heunen, C., Landsman, N.P., Spitters, B. and Wolters, S.A., 2011. The Gelfand spectrum of a noncommutative C*-algebra: a topos-theoretic approach. *Journal of the Australian Mathematical Society* [Online], 90(1), pp.39–52. Available at: https://doi.org/10.1017/S1446788711001157 [visited on 30 November 2020].

Kuhn, T.S., 1970. *The Structure of Scientific Revolutions*. 2nd ed., enl, *International Encyclopedia of Unified Science* vol. 2, no. 2. Chicago; London: University of Chicago Press.

Popescu, S. and Rohrlich, D., 1994. Quantum nonlocality as an axiom. *Foundations of Physics* [Online], 24(3), pp.379–385. Available at: https://doi.org/10.1007/BF02058098 [visited on 30 November 2020].

Tylec, T.I. and Kuś, M., 2015. Non-signaling boxes and quantum logics. *Journal of Physics A: Mathematical and Theoretical* [Online], 48(50), p.505303. Available at: https://doi.org/10.1088/1751-8113/48/50/505303 [visited on 30 November 2020].

Tylec, T.I. and Kuś, M., 2018. Ignorance is a bliss: Mathematical structure of many-box models. *Journal of Mathematical Physics* [Online], 59(3), p.032202. Available at: https://doi.org/10.1063/1.5027205 [visited on 30 November 2020].

Wolters, S.A., 2013. A comparison of two topos-theoretic approaches to quantum theory. *Communications in Mathematical Physics* [Online], 317(1), pp.3–53. Available at: https://doi.org/10.1007/s00220-012-1652-3 [visited on 30 November 2020].

Quantum contextuality as a topological property, and the ontology of potentiality

Marek Woszczek

Adam Mickiewicz University in Poznań

Abstract

Quantum contextuality and its ontological meaning are very controversial issues, and they relate to other problems concerning the foundations of quantum theory. I address this controversy and stress the fact that contextuality is a universal topological property of quantum processes, which conflicts with the basic metaphysical assumption of the definiteness of being. I discuss the consequences of this fact and argue that generic quantum potentiality as a real physical indefiniteness has nothing in common with the classical notions of possibility and counterfactuality, and that also it reverses, in a way, the classical mirror-like relation between actuality and definite possibility.

Keywords

Quantum contextuality is a generic property of quantum systems, which is at the same time one of the most contentious issues in the discussions concerning the ontology of quantum theory. There are several reasons for the status of contextuality being so puzzling: some more historical, going back to the early days of quantum mechanics and some more or less hazy discussions around Niels Bohr's philo-

sophical arguments; others purely theoretical, related to the general problems and contemporary analyses concerning the interpretation of the theory.

However, despite these controversies, there is a growing abundance of experimental tests of contextuality, both state-dependent (e.g. Liu et al., 2009; Bartosik et al., 2009; Lapkiewicz et al., 2011; Ahrens et al., 2013; Borges et al., 2014; Marques et al., 2014; Singh et al., 2019) and state-independent (e.g. Amselem et al., 2009; Kirchmair et al., 2009; Zu et al., 2012; Huang et al., 2013; Zhang et al., 2013; D'Ambrosio et al., 2013; Cañas et al., 2014; Dogra, Dorai and Arvind, 2016; Leupold et al., 2018; Qu et al., 2020), which have been conducted on many diverse physical systems, including tests using the weak (noninvasive) coupling between systems, as well as recently under the strict no-signaling conditions between successive measurements (Xiao et al., 2018). It has also been proposed that contextuality is one of Nature's fundamental physical resources, which has been shown to be crucial in quantum information processing, also in the multiqubit setting (Raussendorf, 2013; Howard et al., 2014; Bermejo-Vega et al., 2017).

Nevertheless, there is no single mathematical framework for studying contextuality as a fundamental physical property, and there also exists a long tradition of questioning if there is anything nontrivial or mysterious about quantum systems being 'contextual', with some authors doubting whether it is possible to directly test it in a lab at all (see e.g. Hermens, 2011). Such a deep discrepancy between viewpoints is troubling if one seriously acknowledges contextuality to be a generic property of quantum systems and quantum information processing, since it makes issues related to the ontology of physics even more obscure. After all, if the very physical understanding of quantum contextuality were wholly dependent on some freely cho-

sen philosophical presuppositions, it would be impossible to even define it in a simple and uncontroversial way, e.g. while preparing and performing experimental tests, which also have their basic tacit ontology of events, measurements and operations. In short, there seems to be a worrying cleft between recognizing the fundamental nature of contextuality as a physical resource and its uncertain philosophical interpretation, which seems to be a matter of taste. I would like to argue that it is impossible to avoid some free rein in ontological interpretation, but it is not as free as it appears to be, and we should treat contextuality as a central challenge for quantum ontology.

1. Quantum contextuality and ontology: more than just another problem

In general terms, contextuality is the impossibility of a consistent, global assigning of the pre-existing $\{0, 1\}$-values (i.e. binary definite properties) to all possible synchronic physical observables on a quantum system for $dim > 2$, which is a serious deviation from the classical-mechanical case where such assignments are in principle always possible. In a more physical framing, this means that there is an unavoidable contradiction between such a pre-assignment of values and the intuitive expectation that the measurement outcome does not depend on other comeasurable observables. Mathematically, this is expressed by the celebrated (Bell–)Kochen–Specker Theorem (henceforth, BKS) proved in 1966–1967 for finite sets of observables (Kochen and Specker, 1967), which is a particular consequence of a theorem with fundamental importance for quantum mechanics, namely the Gleason Theorem (Gleason, 1957).

A precursor of this result was Specker's short but ingenious paper (Specker, 1960), in which he studied the deep logical structure of the co-decidability of propositions concerning any quantum-mechanical system and their embedding into Boolean lattices while preserving the operations of negation, conjunction and disjunction, and answered by 'an elementary geometrical argument' in negative the hope of having the $\{0, 1\}$-valuations for those propositions. In fact, both Specker's motivation and result had conspicuous metaphysical overtones, since their main, innocent-looking target was the idea of the scholastic or Leibnizian-like 'Absolute Observer', which is just a fancy name for the consistent structure of the Boolean valuations defining the determinate, global actual state of a physical universe characterized by an ordered set of propositions. Thus, Specker's answer is quite plain: such a Leibnizian omniscient spectator of the state of affairs, or, equivalently, such a globally determinate actual state ('reality'), cannot exist without contradiction. Then, one decade after Gleason's formal result, John S. Bell (1966) addressed the problem in the framework of the hidden variable deterministic models, i.e. always possessing the dispersion-free states, and he correctly observed that such models cannot be constructed if they are expected to be entirely noncontextual since the additivity of the expectation values for comeasurable observables cannot hold. The significance of these purely formal results is today quite clear, yet their ontological meaning is not. In what follows, I shall exclusively focus on this purely metaphysical side of the problem, thus ignoring the semantic or epistemological aspects, since the main point of interest here is to gauge the prospects for a realistic ontology of physical (quantum) potentiality *in mundo*.

Before we proceed to the topological and physical meaning of contextuality, let us straighten out why the whole issue is so serious

from the purely philosophical point of view. The question is, after all, what does the 'impossibility of the pre-existing $\{0, 1\}$-values of observables' ('impossibility of binary definite valuations') mean?

There is a deep philosophical controversy lurking behind this question, since the central problem at issue here is, as far as we try to faithfully construe the formal structure of quantum theory, that of *physical definiteness* (henceforth, **D**) and *objective reality* or *pre-existence* (**OR**). In fact, from the ontological point of view, **D** and **OR** are closely interlinked as the core ingredients organizing basic metaphysical arguments and constructions, since the fundamental assumption of the Western realist metaphysics after Parmenides, Plato and Aristotle, also inherited by post-ancient physics, and mechanics in particular, is:

> [**DOR**]
> Being real (Greek ὄν) = being definite (τοδί τι) = being a part of a determinate, consistent whole (τὸ ὅλον) or a logical complemented system (λόγος).

The precondition is that an objective being[1] exists and is comprehensible only in a determinate whole, in which maximal mutual distinguishability and definite exclusion (complementation) relations between all its constituent elements are always guaranteed. **DOR** is the guiding classical ontological intuition concerning, for example, states or elements of some definite world, actual or possible, which is why anything which would deviate from it has been taken to be both nonexisitent and incomprehensible, irrational in the strong sense of the term (alogical, in Greek ἀλόγος).

For example, in his *Republic* Plato is explicit that 'what fully exists is also entirely intelligible, and what does not exist is also

---

[1] Not an image or subjective representation of a being, which may be deficient, uncertain, inaccurate or fuzzy.

unintelligible' (*Rep*. 477a3-4).[2] And in the dialogue *Sophist* it is emphatically stated that intelligibility is possible only in the whole system of propositions, i.e. *logos*, which forms a complete system of definite elements[3] and their exclusion relations, beyond which nothing can be comprehended at all: 'to separate each thing from everything else is a total destruction of the *logos*, for what makes *logos* possible is interdependence between all elements' (*Soph*. 259e4-6). According to this way of thinking, every element in itself has no definite nature in separation, and only when taken as a part of such a complete system of logical relations (*logoi*) can it be regarded as some real being: 'whenever each element becomes a single element, it has a limit in relation to any other element and to the whole, and that whole also has it in relation to all its parts' (*Parm*. 158c7-d2). That is why the metaphysical status of Aristotle's prime matter (πρώτη ὕλη) as an ultimate *potentia* or an utterly formless 'bare stuff' is so questionable and vague, since something which is totally indefinite, i.e. falls outside the system of logical categories of predication, cannot exist *per se* as reality, and cannot be invoked as some efficient potentiality taken as a self-standing substrate of what actually exists (see e.g. Charlton, 1992). In fact, the 'prime matter', in line with **DOR**, is rather non-being.

Most modern philosophical discussions concerning indeterminacy or fuzziness (taken to be equivalent) have been focused on logical, semantic-representational and purely epistemological aspects related to vague predicates and epistemic or pragmatic uncertainty, while the metaphysics of ontic indefiniteness *in mundo*, especially in the

---

[2] Here and thereafter all translations from the sources are mine. For the sake of clarity, I take Plato's εἴδη from *Sophist* 259e and μόρια from *Parmenides* 158c-d as meaning 'elements' (of some complete system).

[3] In the pre-classical Greek, the meaning of *logos* is just a gathering or putting together of some discrete elements.

foundations of physics, has been given, due to the appeal of **DOR** and troublesome maladies such as the Sorites Paradox[4], only little attention. In fact, it is common after Russell to assume that indefiniteness itself and the failure of the principle of the excluded middle both signal merely epistemic vagueness or subjective 'indecision', and cannot be accepted as an essential ingredient of consistent, fully realistic (i.e., proclaiming **OR**) metaphysics and physics. In this sense, **D** makes up a necessary, though not sufficient, condition of classical realism as encapsulated in David Lewis's stringent dictum: 'The only intelligible account of vagueness locates it in our thought and language. [...] Vagueness is semantic indecision' (Lewis, 1986, p.212). Contemporary metaphysical approaches which posit some ontic indeterminacy together with **OR** mostly do not introduce it at the basic level of the constituent identity-laden worldly entities, thus they are compatible with standard mereology (cf. e.g. Morreau, 2002). Even more elaborate theories working in the modal framework (e.g. Akiba, 2004; Barnes and Williams, 2011) admit primitive ontic indeterminacy, but they often uphold bivalent classical logic and semantics as their firm backbone. In such cases, **DOR** operates undisturbed at the higher level of possible worlds being complete (maximal) and fully determinate systems, and it is just objectively indefinite which of them is actualized in time.

Thus, in essence, the Western ontological and physical paradigm based on **DOR** assumes that being something means being a certain determinate entity (e.g. 'this thing' or 'this fact'), which is logically

---

[4] See e.g. (Dummett, 1975) for a classic argument of this kind against fuzzy predicates and indeterminacy as unacceptable and unintelligible, that is, pathological for any coherent system. Assuming a semantic analogue of **DOR** and the classical logic, Dummett famously claimed, in a somewhat Platonic vein, that phenomenal properties as commonly conceived cannot be real. The same could apply to all versions of the Sorites Paradox.

defined in terms of multiple relations, in particular the exclusion negation (complementation)[5], in a total propositional system which does not license any first-order vague predicates corresponding to real indefinite properties, including dispositions. Perhaps the most mature and radical fruit of this tradition in early modern philosophy is Leibniz's masterly binary 'universal algebra of concepts' from his *Generales inquisitiones…* (1686)[6] and the related full-fledged essentialist ontology of possible worlds as global logical systems composed of completely definite, indivisible individuals characterized by intrinsic states (see e.g. Bella, 2005). Here, reality itself is thoroughly definite, and there is an ideal 'Absolute Observer' associated with such a complete system. Leibniz offers a clear picture of every definite universe-system as governed "by laws of the general order (*Loi de l'ordre general*) of this possible universe to which they are appropriate and whose concept they determine, as they do also the concepts of all the individual substances which compose that particular universe" (Leibniz in a letter to A. Arnauld, 14 July 1686; Leibniz, 2009, p.73). The Leibnizian panoptical 'Absolute Observer' of the world-system is exactly the omniscient God from Specker's 1960 paper, who might check in advance by logical calculation the $\{0, 1\}$-value of any proposition about the future contingent value of any possible physical observable. Note that here the ontological re-

---

[5] The ontological nature of negation is at the very centre of the dialogue, cf. *Soph.* 258a-c. It was Proclus in the 5th century AD who in his commentary on Plato took it to extremes, saying that 'negations are causes (αἴτιαι) of affirmations' (*In Parm.*, VI, 1075, 18-19; cf. Proclus, 1992, p.428).

[6] This 'universal calculus' of definite concepts is just Boolean algebra before Boole (Lenzen, 1984; Malink and Vasudevan, 2016), equipped with identity, containment, conjunction and negation.

lation between the possible and the actual is always mirror-like: the systemic rules of construction and the global $\{0, 1\}$-determinacy of worlds are the same all along.

DOR seems to be extremely important for physics, and science in general, since it is a necessary foundation of the Principle of Sufficient Reason (PSR), which states that for every state or fact F of a world, there is a definite reason why F is the case, or, in Spinoza's famous phrasing, '*si nulla detur determinata causa, impossibile est, ut effectus sequatur*'.[7] In other words, whatever contingent physical fact (actual value of some observable) is taken into account, there must be a definite ground for its coming to be, e.g. some other *definite state* and its evolution. PSR can be transferred to the indeterministic setting in ontology and we can accept a generalized probabilistic explanation for any F, e.g. by use of the category of disposition or propensity, but in order to have a consistent picture of reality one needs an agreement concerning both (*i*) what a definite fact F is (*objectivity of actual states*), and *(ii)* what a definite disposition for coming to be of F is (*objectivity of determinate possibility*). In order to have (*ii*) being concordant with **DOR**, it is necessary to have a complete basis of all determinate possible (Boolean) states of affairs, which can then carry the pre-assigned values of the (Boolean-Kolmogorovian) probability measure construed as representing the complete ontic dispositions pertaining to an actual system. Let us call such possible states the classical (or Leibnizian) possibilities. If either (*i*) or (*ii*) does not work, or if neither works, then the whole ontological framework falls through, along with PSR as its foundation. In such cases, strong ontic indefiniteness seems to spoil the very idea of **OR**.

---

[7] 'If there is no determinate cause, it is impossible for any effect to follow' (Spinoza, *Ethics*, I, Ax. III; de Spinoza, 1977, p.6).

Another consequence of assuming **DOR** in physics is also Einstein–Podolsky–Rosen's (1935) famous definition of the 'elements of reality' in terms of what can be predicted (*via* physical laws or computability) with certainty: being real means being actual and fully definite, which always requires, at least in principle, globally consistent $\{0,1\}$-valuations for separate states of affairs occuring at any spatiotemporal location. Einstein called it *So-Sein*, 'being-thus', and made it the essence of his principle of separability. The point is that **DOR** is so prevalent and habitual that EPR physical reasoning seems inescapable under the threat of contradiction: if the world is composed of determinate facts *F*, and is consistent as the reality, there has to be a *global* $\{0,1\}$-valuation on *all possible* synchronic physical observables on all *actual* physical systems, whatever their interactions are. That is in essence what reality is. Niels Bohr's quick answer to the EPR paper (Bohr, 1935) rejected that assumption, hence it also dropped **DOR**, but did so by relying on epistemological considerations and the classical notion of the 'experimental arrangement' and 'eventuality' without any clear ontology, which led to the obscurity of his *quasi*-Kantian phrase which referred to contextuality as 'an influence on the very conditions which define the possible type of predictions regarding the future behaviour of the system' (Bohr, 1935, p.700).

Perhaps this is also the main reason why John S. Bell (1981, pp.58–59) wrote that for him Bohr's reply, which relied on the complete 'freedom of handling the measuring instruments', is incomprehensible, while the EPR clear ontological standpoint concerning the 'nature of reality', which respects **DOR**, is not. Bell was perfectly right when he stressed that what is at stake is not just the old problem of determinism or some particular correlations *per se*, but rather, in

line with Bohr, the commonsense definition of reality[8] and physical systems having some pre-assigned definite properties that are locally detected (measured). And it was Bell himself who constructed the first proof of the state-dependent contextuality for quantum compound systems dissociated in space (Bell, 1964), before it was fully realized that the resulting nonlocality is mathematically just a particular instance of the more general contextuality (e.g. Horodecki et al., 2015; Acín et al., 2015).[9] It seems clear that the latter is indeed, with respect to **DOR** and the impossibility of consistent $\{0, 1\}$-valuations (Bohr, 1935, p.699), a severe ontological problem, of course to the extent that one wants to have some explicit realistic ontology at all.

2. Ontic quantum contextuality as a topological property

Let us define formally what is so curious about physical contextuality and why such a basic and apparently innocuous assumption as **DOR** is questionable precisely on ontological grounds, including the notion of classical possibility. In order to make fully explicit the topological difference between the classical and quantum structures of observables, as well as between the corresponding notions of 'state', we adopt here a theory-dependent formulation of contextuality, in contrast to e.g.

---

[8] Bell (1981, pp.45–46) humorously wrote that Einstein, Podolsky and Rosen 'were with the man in the street in this business'.

[9] The nonexistence of joint probabilites for quantum systems does *not* depend on any distance in spacetime. Spatial separation between subsystems and the nonlocal correlations only make conspicuous the generic nonclassicality of this contextuality. One may call it *general nonseparability*, also in the case of non-composite systems.

the general theory-independent framework proposed by Abramsky & Brandenburger (2011) in terms of sheafs over the space of empirical measurement scenarios.

Let $\mathcal{S}(\boldsymbol{B})$ be a topological Stone space of the Boolean algebra $\boldsymbol{B}$ of propositions about the synchronic physical properties of some classical system O, which means that $\mathcal{S}$ effectively encodes the relevant physics of O (whatever it is: gravitational, electromagnetic, or other). This is quite obvious since any pure mechanical state ψ of O automatically selects an ultrafilter in $\boldsymbol{B}$, and the field of all ultrafilters in $\boldsymbol{B}$ is isomorphic to $\boldsymbol{B}$ as its dual topological Stone space $\mathcal{S}(\boldsymbol{B})$ due to the Stone representation theorem. And let $\boldsymbol{\mathcal{T}}_{\mathcal{S}} = (\mathcal{T}, \pi, \mathcal{S})$ be the constant Boolean sheaf over the state space $\mathcal{S}$, that is, every stalk $\mathcal{T}_S$ is fixed as $\mathcal{T}_S = \pi^{-1}(\{s\}) = \boldsymbol{2} \equiv \{0, 1\}$, $s \in \mathcal{S}$, with the projection $\pi : \mathcal{T} \to \mathcal{S}$ as a local homeomorphism from the étale space $\mathcal{T}$. Local sections of the sheaf $\boldsymbol{\mathcal{T}}_{\mathcal{S}}$ over the proper subsets $S \subset \mathcal{S}$ are continuous maps $\sigma : S \to \boldsymbol{2}$, where $\pi \circ \sigma = id_S$, which gives the values $\sigma(S)$ of the sheaf. By definition, a section is global when $S = \mathcal{S}$, and then every σ is just a restriction of the latter (Mac Lane and Moerdijk, 1992, chap.II; Knoebel, 2012).

The fundamental fact about the classical-mechanical spaces of pure states ψ of any O is that one can always have global sections of $\boldsymbol{\mathcal{T}}_{\mathcal{S}}$, and the full algebra $\boldsymbol{\Gamma}_{\mathcal{S}}(\mathcal{T})$ of those sections is just isomorphic to $\boldsymbol{B}$, which is a constructive backbone of classical mechanics. In the classical world we can always locally glue information into a single coherent picture of the actual properties of O by consistent Boolean valuations, that is, it is possible to consistently extend any local σ for $S = \mathcal{S}$: the global, i.e. unrestricted, sections over $\mathcal{S}(\boldsymbol{B})$ *do always exist* (Fig. 1). One can thus say that the logic of constant $\{0, 1\}$-valuations is 'flat' or well-behaved, or, that the corresponding sheaf
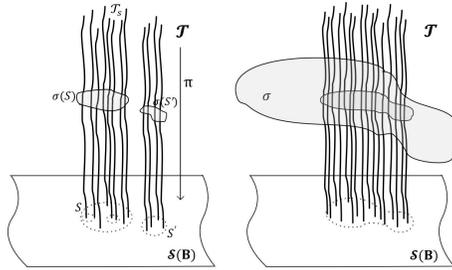
Figure 1

is trivial, which makes the question of what counts as reality quite unproblematic and the whole scaffolding of the physical reference frames to be of purely epistemic character.

In fact, what is of crucial importance is that, due to the existence of ultrafilters, the very notion of physical state as a catalogue of the actual properties of O becomes straightforward; each physical system has its own Einsteinian 'being-thus', and, moreover, there is no difficulty in extending this notion to the whole universe (world-Ψ) taken as a mechanical system of O's. This means that any ontic contextuality of states is absent, and there is only one global 'context' of coexisting properties without restrictions, which is encapsulated by the metaphysical catchword 'Absolute (panoptical) Observer'.

Note that since the ultrafilters on B simultaneously generate the full Boolean algebra of all the global sections of $\mathcal{T}_{\mathcal{S}}$, that is $B \cong \Gamma_{\mathcal{S}}(\mathcal{T})$, one gets the algebra of the classically possible (Leibnizian) mechanical worlds, which is already built from the very beginning into the unitary logical framework. The curious 'mirroring' feature of the latter is that it is *exactly the same logic* which governs the domains of the possible and the actual, hence there seems to be nothing special

about possibility *per se*—in fact, it is difficult to comprehend why the actual would be ontically distinguished as 'fully existing'. That is why thinkers like Leibniz needed a ladder of additional metaphysical constraints as rules of the world selection, or foreknowing God's one-moment choice, while in the classical deterministic cosmology there is always the problem of the extremely improbable boundary conditions and the unpleasant cosmological fine-tuning of the initial (or final) world- Ψ as a bare fact. It looks like the domain of possibility was fully modelled upon the domain of actuality, and the probabilities for the latter could be consistently pre-assigned only after some complete universe of logically possible states of affairs, i.e. the full Boolean algebra $\Gamma_{\mathcal{S}}(\mathcal{T})$, is fixed (that is how the probability of 'improbable boundary conditions' is assessed).

Even if one prefers the indeterministic ontology and objective (noncontextual) dispositions or propensities, like e.g. Karl Popper (1967), it is tacitly assumed that there is no way beyond the Kolmogorovian measure based on Γ. The reason behind this premise is not just PSR and the requirement of concordance between (*i*) and (*ii*), it lies much deeper, in ontological **DOR**: one needs a complete basis of definite (Boolean) elements for the probabilistic measure, and then (*i*) and (*ii*) can be handled in a unitary manner. In the end, there is no way out of the Boolean universe here. The logical state structure $\Gamma_{\mathcal{S}}(\mathcal{T})$ in physics is thus a full realization of the paradigmatic ontological intuition, since being means being definite *and* being a part of some fully determinate, consistent whole, which brings forth, as a sort of a byproduct, the idea of panopticality or a detached observer. One may sum this up by saying that classical ontology favours the actual and the definite, paying the price of making the potentiality a mirror image or an expanded shadow of the actual.

In the quantum regime, the rules of the game change drastically. A lattice of propositions about the properties of a quantum system O (which may also be multipartite) forms an orthomodular σ-orthoposet $\mathfrak{L}(\mathcal{H})$, being isomorphic with the lattice of closed subspaces of the complex Hilbert space $\mathcal{H}$ associated with O, which is complete if it is ordered by inclusion (for simplicity we limit ourselves to cases when $\mathcal{H}$ can be constructed at all, but it can be generalized to all types of W^*-algebras of observables, see (Döring, 2005)). The sheaf-theoretic construction for quantum observables may be achieved in general as follows. Let a topological space $\mathcal{L}_{\mathfrak{L}(\mathcal{H})}$ be a family of all Boolean (commutative) subalgebras L of $\mathfrak{L}(\mathcal{H})$ ordered by inclusion $\subseteq$, and let us have a required Boolean homomorphism (the Kochen–Specker valuation) $h : L \to \mathbf{2}$ for any $L_i \in \mathcal{L}_{\mathfrak{L}(\mathcal{H})}$. We can define the partial ordering $\leq$ for pairs (L_i, h_i), given by: $(L_1, h_1) \leq (L_2, h_2)$ iff $L_1 \subseteq L_2$ and $h_1 = h_2| L_1$, associated with the respective topology of $\mathcal{L}_{\mathfrak{L}(\mathcal{H})}$. Every Boolean subalgebra L_i in $\mathcal{L}_{\mathfrak{L}(\mathcal{H})}$ belongs to some maximal Boolean subalgebra (a maximal quantum observable), which forms a (quasi-) classical *context* $\mathfrak{M}$.

Now the *quantum spectral sheaf* (over the base space of quantum observables or spectral algebras), the closest analogue of the classical $\mathcal{T}_S$ we get, is a triple $\mathbf{\Phi}_{\mathcal{L}} = (\Phi, \pi, \mathcal{L}_{\mathfrak{L}(\mathcal{H})})$, with $\Phi_{L_i} = (L_i, h_i : L_i \in \mathcal{L}_{\mathfrak{L}(\mathcal{H})})$ and a projection as a surjective local homeomorphism (étale map) $\pi : \Phi \to \mathcal{L}_{\mathfrak{L}(\mathcal{H})}$ from the sheaf space Φ. A continuous local section (of the projection π) over some subset $M \in \mathcal{L}_{\mathfrak{L}(\mathcal{H})}$ is a map $\sigma : M \to \Phi$, $\pi \circ \sigma = id_M$, and the images of local sections naturally induce the topology for the whole étale space Φ. That is so because for every $L_i \subset M$ we obviously get $\{0, 1\}$-values of the sheaf, $\sigma(L_i) = (L_i, h_i)$, with the proper restriction condition: $\sigma(L_1) = (L_2, h| L_1)$ for every $L_1 \subseteq L_2$, which always guarantees the local compatibility of sections as a crucial fact about the quantum ob-

servable values. Due to Gleason's no-disturbance principle for $\mathcal{L}_{\mathfrak{L}(\mathcal{H})}$, the probability for obtaining some value cannot depend on a valuation of some other compatible observable, which gives the properly behaving, i.e. additivity satisfying, probabilities for $\mathfrak{M}$'s (Gleason, 1957).[10] Hence the latter define subclassical 'patches' within the full quantum $\mathfrak{L}(\mathcal{H})$, and the essence of the generic quantum 'complementarity' is that it is not possible to have the full $\{0, 1\}$-valuations (i.e. the Kochen–Specker states) over two different $\mathfrak{M}$'s at the same time.[11]

The (Bell–)Kochen–Specker Theorem in the sheaf-theoretic formulation simply says that if a Hilbert space $\mathcal{H}$, representing the physics of a system O, is of $dim > 2$, then the spectral sheaf $\mathbf{\Phi}_{\mathcal{L}}$ on $\mathcal{L}_{\mathfrak{L}(\mathcal{H})}$ has *no global sections*. Thus, we have a strong topological obstruction, which makes the unrestricted extending of local σ impossible if the global consistency of valuations $\sigma(L)$, as in the case of $\sigma(S)$ for $\mathcal{T}_S$, has to be preserved (ultimately, that is what the objectivity of states (*i*) and the EPR intuition about the definite 'elements of reality' do unconditionally demand). A global Boolean 'context' for mechanics does not exist, and only single local contexts $\mathfrak{M}$, as defined by quantum spectral algebras, can be regarded as being definite, hence the whole idea of synchronic $\{0, 1\}$-properties, objectively preexisiting and pertaining to O as 'states', gets into fatal trouble.

In consequence, what is actual, i.e. a definite fact F from (*i*), is only a *random local section* $\sigma(L)$ of π in a particular definite $\mathfrak{M}$, and that physically relativizes the valuation h itself, which is a contingent

---

[10] In the operational framework, one might call this central property the 'noncontextuality of projective measurements'. However, it should be remembered that it is a purely algebraic feature of the whole quantum structure, and not just the statistical characteristics of particular sets of measurement outcomes.

[11] See e.g. (Heunen, Landsman and Spitters, 2012) for a purely operator-algebraic account of quantum complementarity, and (Heunen, Landsman and Spitters, 2012) for a general categorical approach in terms of dagger monoidal kernel categories.

relation of O to some exosystem O' that induces the particular $\mathfrak{M}$.[12] Quantum mechanics simply lacks the constructive backbone of its classical predecessor. Before I discuss what this means for any ontological account of potentiality, it is important to indicate why this topological obstruction has nothing to do with the classical epistemic contextuality and should be physically dealt with as a fully ontic property. These arguments are thus based on strictly physical, including experimental, considerations, being at least relatively independent from contestable metaphysical presuppositions, and they range from quantum-informational to thermodynamical aspects. In essence, $\Phi_{\mathcal{L}}$ induces the totally nonclassical structure of *physical* probability and entropy, and that only in turn effectively restrains ontologies with definite states or dispositions modelled after $\Gamma_{\mathcal{S}}(\mathcal{T})$ (I shall go back to the latter point in the next section).

Firstly, the topological obstruction on $\Phi_{\mathcal{L}}$ gives rise to the entirely novel physics of information processing in Nature, the objective (non-epistemic) quantum regime, which has nothing to do with human labs and disturbance-producing experiments. At the level of the latter, due to the contextuality or algebraic 'locality' of $\sigma(L)$, it is impossible to perfectly distinguish some unknown pure nonorthogonal quantum states, and thus to determine in one shot an unknown state on a single copy, however weak the conceivable measurement is (D'Ariano and Yuen, 1996). One can do that only randomly with the Holevo–Helstrom probability $p_{HH} \leq \frac{1}{2}(1 + \sin\theta)$ of successful guessing, for the equal probability of the preparation of each state

---

[12] Due to the standard minimal interpretation of the situation, quantum mechanics cannot be – and is not – metaphysically realist in the *classical* meaning of the term, which is quite obvious. But the stronger claim that it cannot be realist at all (in any acceptable sense) is premature and overly dependent on the classical metaphysical framework of actuality and its 'mirroring' possibility, that is $B \cong \Gamma_{\mathcal{S}}(\mathcal{T})$. In fact, it rests on the acceptance of **D** as a putative necessary condition of realism *per se*.

and θ being an angle between two vectors in $\mathcal{H}$. However, any purely epistemological explanation of this bound on *discriminability*[13] is misleading. First note that the related basic quantity $\cos^2 \theta$ is a probability that the two corresponding states may be confused, that is, if one of them is 'measured', or, better still, filtered in the basis (or context) of the other, it passes the test, and vice versa. Now, it can be quantitatively estimated for such a quantum nonzero *confusability* of states how much any (quasi-)classical simulation which is maximally epistemic (i.e. assumes the quantum measure to be just a measure of classical ignorance about a distribution of the underlying actual ontic (hidden) $\{0, 1\}$-states) must fail for $dim\ \mathcal{H} > 2$, while the failure grows dramatically with $dim \to \infty$ (Leifer, 2014; Barrett et al., 2014; Branciard, 2014), which has already been tested with success (e.g. Ringbauer et al., 2015; Liao et al., 2016).[14] This means that BKS is sufficient to rule out any maximally epistemic models (including noncontextual preparation models) which try to treat the quantum measure like the quasi-classical measure, which it is not. Apparently, there is no hidden reality with facts satisfying **DOR** – no 'elements of reality' to be mechanically disturbed by our clunky experiments in spacetime, and this seems to be a source of quantum computational gain.

---

[13] If multiple copies of transmitted states are physically available, the empirical probability of successful guessing significantly increases with the number of copies and the specific measurement strategies involved. Note that the probability p_{HH} is also an entropic measure, the so-called min-entropy (Konig, Renner and Schaffner, 2009).

[14] Recently, Schmid and Spekkens (2018) have shown that any noncontextual ontological model must satisfy some nontrivial trade-off relation between the Holevo–Helstrom bound and confusability, which is violated by quantum systems as a clear indicator of ontic contextuality. More precisely, for a given confusability the optimal quantum trade-off allows more success rates than the classical noncontextual one. This no-go result is particularly important since it is theory-independent and can provide a clean operational account of (non)classicality.

But there is, secondly, much more to contextuality and indistinguishability (which in turn have some surprising relation to the causal stability of spacetime). The latter makes it impossible to build perfect quantum cloning or deleting machines due to the quantum no-cloning, no-broadcasting and no-deleting principles. If it were possible to perfectly discriminate unknown nonorthogonal states beyond the Holevo–Helstrom bound, one could freely produce their copies and thus easily violate the no-cloning prohibition (in fact, state discrimination may be regarded as a form of cloning). Although there indeed exist noncontextual ontological models which are able to recover some substantial aspects of the no-cloning phenomenology, what is a purely quantum advantage induced by ontic contextuality is a strictly higher maximum fidelity of cloning as related to the confusability of states (Lostaglio and Senno, 2020). It is again contextuality as a topological property which precludes any fully classical simulation of no-cloning and, presumably, no-deleting.

Now, the physical fact is also that if it was possible e.g. to delete an unknown quantum state available in two copies or exactly clone an unknown state, one could send faster-than-light signals in the Universe and infringe its causal order (Gisin, 1998; Pati and Braunstein, 2003). One can still perform a quantum approximate or just probabilistic cloning, but even in such cases opening a nonlocal signaling channel is ruled out (Pati, 2000; Bruss et al., 2000). It is well known that there is a much larger space of strong nonsignaling (more nonlocal) correlations, which also do not allow free cloning, but are not achievable on quantum systems (Masanes, Acin and Gisin, 2006), however general contextuality itself provides a specific insight even here. Every Kochen–Specker proof can be used to derive the Bell (bipartite) inequalities *maximally* violated by quantum theory and with their values being at the same time *exactly on the boundary* of infringing the

Lorentz invariance, but never actually doing so (Aolita et al., 2012). These special maximal contextuality-based correlations are particularly interesting, since the corresponding probability distributions for actual events have a null set of local models able to simulate them, but nevertheless, in this extremal case, they precisely accord with the relativistic causality. Most philosophical discussions concerning the nonclassical probability distributions have usually been focused on the vexing problem of nonlocality between actual, definite systems in spacetime, but the correlations of this sort reveal that the source of quantum advantage in general resides rather in the *full absence of actual* (spatiotemporal) $\{0, 1\}$-properties, that is, in objective indefiniteness from the Kochen–Specker-type sets. One may surmise that such a sector of tangent concordance between contextuality as a topological constraint (without any reference to the speed of light or the physics of interaction) and the nonsignaling bound (causal structure of classical spacetime)[15] cannot be a mere coincidence and is more readily an ontic feature of the Universe, without any dependence on the messy operations in labs.

Finally, there is one ultimate judge in physics, and this is thermodynamics. Quantum histories (let's call them q-histories) are thermodynamically different from classical histories, since their entropic characteristics deviate extremely from those of classical histories solely due to ontic contextuality. In particular, q-histories of single indivisible systems violate the simplest 5-cyclic entropic noncontextuality inequality in time (Chaves and Fritz, 2012), as well as the 5-cyclic Kurzyński–Ramanathan–Kaszlikowski inequality for condi-

---

[15] Of course, constraints induced solely by the Kochen–Specker contextuality cannot reproduce the full statistical constraints on quantum nonlocality, since the former are more general – they just do not include the complex product structure typical for entanglement in multipartite systems.

tional entropies (Kurzyński, Ramanathan and Kaszlikowski, 2012), which would be impossible if the underlying structure of events were $\Gamma_S(\mathcal{T})$. This is of fundamental significance as far as it needs neither the introduction of the aforementioned correlation functions of single outcomes nor averaging over some hypothetical hidden variables, as in the Bell model. Moreover, one can construct a *purely thermodynamic* formulation of the (Bell–)Kochen–Specker Theorem and show that the quantum entropic contextuality is a very strong obstruction to any putative 'thermodynamic realism' based on the presumed Kolmogorovian measure and hidden dispersionless states (Jia et al., 2018). If the resulting quantum entropic monogamy relations[16] were violated for sequential operations in time, it would also be possible to signal into the past and infringe the causal structure of spacetime, which quantum protocols cannot do. Thus, ontic contextuality induces a more general, consistent thermodynamics, and ontology rather has to come to terms with this nontrivial fact without trying to seek rescue in epistemology.

3. Quantum potentiality is real indefiniteness, not counterfactual possibility

A natural starting point for quantum ontology is the conclusion that the problem is not quantum mechanics itself, in particular BKS, but rather the metaphysical assumption concerning what counts as being

---

[16] Monogamy relations are an important feature of quantum correlations as quantitative constraints on their sets, which limit the shareability of information between distinct parties (see e.g. Dhar et al., 2017). They are closely related to the no-signaling principle and its generalized form, Gleason's no-disturbance principle (for sets of compatible measurements). In the thermodynamic framework studied e.g. in (Jia et al., 2018), the straightforward monogamy relations are constructed for entropic Shannon-type inequalities defined on sets of measurements.

and what is non-being, that is, in short, **DOR**. This is quite interesting since only rarely does physics impose such pressure upon a basic metaphysical intuition which often seems to be a pure *a priori* of physics. Of course, this does not happen without some serious opposition, since there might be doubts about whether the pressure is legitimate or whether physics should be allowed to demand drastic moves at the primitive level of ontology. I fully agree that no physical restrictions can just pick out a particular metaphysics as the right one (cf. e.g. Hawley, 2006; French, 2014), and neither can they solve metaphysical dilemmas like determinism vs. indeterminism, since the constructive inventiveness of general metaphysics is illimitable, but, nevertheless, the naturalistic ontology of physics is not so unconstrained in its dealing with contextuality.

The indicated opposition to the overtly ontological interpretation of contextuality makes quite clear an uneasy tension between physics and theoretical metaphysics, which is discernible in discussions on both sides. In this specific case, it is not so much a methodological problem, but rather a relatively rare clash between genuinely fundamental metaphysical assumptions, spawning a whole class of shared realistic intuitions (concerning e.g. 'state'/'fact', 'disposition' or 'possibility'), and the equally fundamental structure of the highly confirmed, although falsifiable, physical theory, and that indeed makes any choice a delicate issue. On the one hand, it seems rather unsurprising that one tries to avoid abandoning basic intuitions that are a sort of common ground in metaphysics. For example, it is instructive to observe how **D** and counterfactuality still play a guiding constructive role even in novel philosophical conceptualizations of ontic indeterminacy (e.g. Akiba, 2004), which illustrates the previously mentioned 'mirroring feature' of such models. On the other hand, there has been considerable progress in the physics of contextuality, both experimen-

tal, e.g. concerning entropic tests in state-independent scenarios (Qu et al., 2020), and theoretical, related to the theory-independent, purely operational accounts of contextuality (which are beyond the scope of this paper). Both make the pressure on the ontology of physics constantly grow, analogously to the historical case of the theory-independent Bell inequalities. Hence, the clash seems to be more and more conspicuous and hard to ignore.

One of the common answers to contextuality is adopting the view that if ψ is not a catalogue of definite actual properties, one could just accept indeterminism with physical possibilities as an ingredient of ontology, which might solve the riddle. However, the basic lesson from BKS is that merely to keep **DOR** unscathed is simply *not* enough, since possibilistic indeterminism is neither sufficient nor necessary to account for contextuality. For example, the most natural choice for a classical dispositionalist is to assume that an actual system O always possesses a set of definite intrinsic potential properties, which is a consistent catalogue of its possibilities represented by ψ and the full set of probabilities calculated from the complex amplitudes. According to such a view, the actual outcome of interaction with O would be just the indeterministic manifestation of one such pre-existing possibility pertaining to the system.

But BKS itself, as a strong topological obstruction on Φ_{L_i}, makes such a model untenable, that is, taking ψ to be a set of objective, pre-assigned possibilities that reside in O results in physical contradiction. The first simple proofs of this theorem have been offered by Adán Cabello (1999a) who used the GHZ-like construction (without inequalities and probabilities) for three pairs of systems, as well as the Hardy-like proof for two pairs of systems. In fact, every quantum Bell-like scenario with entangled systems is just a particular realization of the contextuality scenario for compound systems, hence it

is a *statistical* illustration of the breakdown of the notion of local intrinsic possibilities pre-assigned to subsystems. One may easily demonstrate that it is in each case related to the noncontextual assignments of properties, even if they are only probabilistic correlations and nothing more (Cabello, 1999b). The maximally nonlocal correlations discussed in the previous section are just an explicit example of this.

Furthermore, one may use the general information-thermodynamic Bell–Kochen–Specker obstruction and the resulting quantum monogamy relations from simple Shannon-like inequalities (Jia et al., 2018) as a direct proof of the impossibility of constructing a model with the physical sources of stochastic behavior intrinsic to systems, whether they are simple or composite. Quantum entropies themselves, not just some particular sets of purely statistical correlations in space and time, rule out pre-assigned internal dispositions that might stochastically manifest in direct response to some external stimuli. This is, as I indicated, because the Kolmogorovian probabilistic measure completely breaks down for $\mathfrak{L}(\mathcal{H})$, the quantum probabilities are not convex combinations of $\{0, 1\}$-states, and this has nothing to do with some local mechanical disturbance of state or the epistemic conditions of access to the properties of some O. The trouble lies in the classical understanding of a (pre-assigned, Boolean) property, whether categorical *or* dispositional, that is, in Einstein's 'being-thus', which is why his principle of separability cannot work. For example, statistical Liouville mechanics with a strong epistemic restriction (i.e. respecting some specific constraints on observables), equivalent to Gaussian quantum mechanics (Bartlett, Rudolph and Spekkens, 2012), is unable to recover quantum contextuality, either state-dependent or state-independent, including the full range of nonlocal correlations for Bell-like scenarios.

Now it is possible to identify the deep metaphysical source of permanent confusion about quantum contextuality as the tacit adoption of **DOR** and its philosophical consequences, such as e.g. the working metaphysics of the Boolean 'elements of reality' (states) collectively fused into some postulated actual world-Ψ, the credence in the definite counterfactual observables (Leibnizian possibilities), as well as the Principle of Sufficient Reason, which encourages mechanical reasoning about producing the outcomes. On the mathematical level, their failure is effected by the total breakdown of the structure $\mathcal{T}_S$ and the corresponding commutative probability theory, but a decision to reject **DOR** has far-reaching consequences, not just for standard philosophical counterfactual analyses, which are of no use here, but for understanding the underlying physics. The problem begins with the very vocabulary of quantum mechanics inherited from the classical theory.

The term 'contextuality', as used in the Bell-like sense, has been employed in the framework of the hidden variable models, that is $\mathcal{T}_S$, and it may still give the impression that it is the mechanical 'context of measurement', or the internal state of the apparatus, which somehow 'influences'[17] the outcome by covertly selecting some subset from among the inferred, physically pre-existing possibilities, even nonlocally in spacetime (for a methodic critique see e.g. de Ronde, 2017). However, the breakdown of $\mathcal{T}_S$, i.e. the absence of ultrafilters on the presumed world-Ψ, is most naturally equivalent to the *global collapse of the valuation definiteness* in the fully ontological sense, which means that quantum potentiality is not any sort of Leibnizian

---

[17] It has been the Bohrian or, more generally, Copenhagen tradition which introduced such the problematic terminology of 'influence', though, to be fair, one must admit that it was construed by Bohr purely epistemically as an 'influence on the conditions of possible experience'. The very idea of 'context' also has a strongly epistemic flavor. Such a quasi-Kantian formulation is unworkable from the point of view of ontology.

possibility and has nothing to do with the classical counterfactuality of measurements. Thus, the ontological source of contextuality from the BKS formulation on $\mathcal{L}_{\mathfrak{L}(\mathcal{H})}$ is rather the purely physical, real indefiniteness (the indeterminacy of being) *and* the associated purely ontic randomness in the quantum regime. Abbott et al. (2012) have proved the stronger version of BKS (without assuming the global Kochen–Specker valuation), in which it is possible to precisely identify the objectively (provably) indefinite values of observables, and hence to produce strongly certified quantum random sequences which are Turing incomputable. In fact, it can be shown that such value indeterminacy is ubiquitous if only one particular observable is locally fixed to be value definite (Abbott, Calude and Svozil, 2014).

In this sense, physical indeterminacy as a quantum resource seems to be almost everywhere in Nature, while definiteness is a local phenomenon, which makes indefinite potentiality physically fundamental, and the definite actuality secondary. It is the former that is *the* main engine of quantum mechanics, and makes possible quantum information processing and nonlocal effects without the slightest infringement of the Lorentz invariance, since what is indefinite and does not exist in spacetime cannot be 'transmitted' or 'sent' in it (as a definite spatio-temporal being can be).

Some misconceptions in construing quantum theory may indeed result from imposing **DOR** as a tacit ontological premise and explicating quantum potentiality through the category of actuality, i.e. taking the latter to be ontologically primary and pervasive. But what we have in the quantum regime is an exact reversal of the classical mirror-like relation of actuality and possibility, the latter being constructed according to the former within the unified logical universe $B \cong \Gamma_{\mathcal{S}}(\mathcal{T})$. As a topological obstruction, BKS contravenes this relation, making the potentiality ontic, primitive and *dissimilar* to

the actuality that cannot simulate or encompass the former within such a Boolean logical universe. It is precisely this dissimilarity and incomputability which indicates that the post-classical terminology of both the counterfactual 'possibility' and 'measurements' of actual states is misleading, for it still assumes the metaphysics of the 'Absolute (panoptical) Observer'.[18] In particular, the breakdown of **DOR** and the Principle of Sufficient Reason would mean that there is no operative rule, law or reason determining the transition from the (generic quantum) potential to the (spatio-temporal) actual, which is 'irrational' or 'alogical' in the fully ontological sense (cf. e.g. Bub, 2014).[19] In this case, the Spinozian axiom would be false: *possibile est, ut effectus sequatur, si nulla detur determinata causa*, that is, the operative cause of some definite state of affairs may be an indefinite (non-Boolean) potency.

The breakdown of determinacy does not imply that ontological realism *per se* is impossible or that objective reality itself vanishes into thin air. In point of fact, the aforementioned reversal itself makes quantum potentiality and randomness an effectively *autonomous* agent and a *physically real* reservoir. For example, entropic contextuality may be utilized in quantum network communication, and indefiniteness in general is critical as a resource for speed-up in universal quantum computation (Howard et al., 2014). Quantum randomness is autonomous or absolute in the sense that it cannot be physically

---

[18] However, the difference between pure quantum *potentiality* and classical *possibility* is not verbal or purely metaphysical, since the issue is not speculation about possible observers but the physical structure of quantum processes as encoded topologically in $\Phi_{\mathcal{L}}$.

[19] As far as I know, it was Wolfgang Pauli who first adopted the epistemological term *Irrationalität* from the early Bohr, emphatically insisted on the metaphysical meaning of that 'irrational actuality' against the currents of the history of Western ontology, and called quantum mechanics the 'theory of becoming' (see e.g. Pauli, 1994).

controlled or influenced by any local agent, it may only be exploited or used as a means for doing physical work, hence it paradoxically resembles the laws of classical mechanics and, from the point of view of any such agent, is maximally independent. In the quantum realm, there is nothing more real than that.

What is more, the only fully noncontextual, objective element of the quantum algorithm, which is explicitly Lorentz-invariant, is the Born rule that fixes the complete probability distribution for each quantum state (algebraically, $\mathfrak{M}$). That noncontextuality is, as the no-disturbance principle, at the centre of Gleason's Theorem and plays a decisive role for defining the objective concept of generic 'quantum state' as a unique measure on $\mathfrak{L}(\mathcal{H})$, which is a situation peculiar to quantum theory.[20] Barnum et al. (2000) have perhaps rightly stressed, from the mathematical point of view, that 'it is hard to imagine a cleaner derivation of the probability rule than this'. Philosophically, it is a hallmark of the quantum reversal, for what is physically real and invariant seems to be a generic potentiality, even if the universal value definiteness is excluded by $\mathfrak{L}(\mathcal{H})$, which cannot support any Boolean measure. If some domains of reality do defy determinacy, one may conclude that the intuition that backs **DOR** is plainly false.

To make it more express and concrete, let us take an example of a test of the state-independent contextuality on a single simple system in time as realized e.g. on single qutrits by Xiao et al. (2018). In such a scenario it is possible to guarantee the statistical no-signaling condition between the consecutive measurements by assuming that any disturbance of the future is indistinguishable within an experi-

---

[20] In classical Liouville mechanics, one can impose the probabilistic measure on a phase space freely as reflecting the contingent ignorance of an agent modelling a system.

mental error from the influence of the future on the past. In perfect agreement with BKS, such a system strongly violates the correlational noncontextuality inequality, which is an empirical manifestation of the topological nonclassicality of the quantum sheaf. The first and fundamental lesson is that there are no *actual* $\{0, 1\}$-properties of it just waiting there to be disclosed by sequential measurements.

But there is much more if one acknowledges the reality of dispositions. The first, conservative, strategy is to try to model them by somehow enforcing **D** as binary valuations on the level of sheer counterfactual possibilities pertaining to that system in each moment. But, as we have seen, it is also structurally excluded on the same grounds as actual $\{0, 1\}$-properties. To say that a single system carries its intrinsic dispositions in time as its own possibilistic 'state' makes no sense, since the respective evolving probability distributions are unable to violate the noncontextuality inequalities. The ontological notions of both intrinsic 'state'/'properties' and intrinsic 'disposition' seem misguided unless one accepts a pervading 'Ptolemaic' hidden fine-tuning in spacetime.[21] Thus, taking the BKS and experimental results seriously, I cannot see any significant gain in invoking world-Ψ and counterfactual (Leibnizian) possibilities as useful metaphysical tools.

Better suited for directly ontological interpretation of the new physical sheaf-structure $\Phi_{\mathcal{L}}$ is the second, more radical, strategy. It takes quantum potentiality to be objective (*context-invariant* and *quantifiable*) pure indefiniteness, thus, in line with Abbott et al. (2014), there are no determinate properties of any system *nor* measures on them before actualization, which might serve as building blocks for spatiotemporal reality and causal interactions, and neither are there

---

[21] For the contextuality-testing scenarios with single systems in time see (Cavalcanti, 2018).

laws or rules governing such actualizations. Instead of trying to abide by **D**, it accepts quantum reversal, that is, rejects the primacy of actuality and binary valuations as a foundation of ontology, hence actual systems in spacetime cannot be its basic entities at all. Here, in stark contrast to classical mechanics, quantum theory is only about potentiality without mechanical states, *not* about any individual properties or events in spacetime and their putative counterfactual variants as logical shadows.

Although it is beyond the scope of the present paper, one may repeat, for the sake of clarity, that since formally nonlocality is a special case of general contextuality, it is precisely the objective non-spatiotemporal indefiniteness which should be seen as a physical source of nonlocal phenomena, not some influence in spacetime or non-Lorentzian (superluminal) action. In this sense, 'nonlocality' might be treated as a sort of misnomer, which results from taking for granted the priority of the spatiotemporal actuality composed of determinate 'states of affairs' in supposed interactions.

4. The quantum world without real potentiality?

It is interesting to observe the systemic cost of some strategies to remove ontic contextuality and fully recover **DOR** (or Specker's 'Absolute Observer') as a tacit metaphysical principle. Even if subtheories such as Gaussian quantum mechanics are unable to do that effectively, there are other approaches whose intent is to reduce both the strong ontological import of BKS and ontic randomness. For example, in the many world interpretation, with one evolving multiworld-Ψ, contextuality loses any significance for there are no unique states of affairs as actualized properties and possibility merges again with actuality,

as in classical $\mathbf{\Gamma}_\mathcal{S}(\mathcal{T})$ but with one crucial difference, namely that these worlds must now be fully included on par, as real elements of physics.[22] One may call this an 'explosivist strategy' which annuls the uniqueness of outcomes as a premise of BKS.

In Bohmian mechanics, which is, as a matter of fact, not just interpretation but rather a new theory, contextuality results from the mechanical 'context of an experiment', i.e. an uncontrolled interaction between O and the hidden degrees of freedom of an exosystem (apparatus), hence it looks as if it were utterly trivial: 'the outcome of an experiment depends on the experiment' (Dürr, Goldstein and Zanghì, 2013, pp.148–153). But since the actual measurement outcomes are strongly relational, it is impossible to say that they are detections of the pre-existing properties of O. In fact, Bohmian measurements in our postulated state of cosmological quantum equilibrium are simply 'false', they do not 'measure' anything (Valentini, 2010, pp.486–488), and properties, if contextual, are in the Bohmian theory simply nonexistent (Dürr, Goldstein and Zanghì, 2013, p.153). In other words, the universal timeless world-Ψ is taken very seriously, it is the only *real* wave function (Dürr, Goldstein and Zanghì, 2013, pp.264–272), but, contrary to the many world interpretation, eigenvalues are deprived of any ontological significance, they do *not* correspond to any real subquantum properties at all. One may call this a double-edged 'superdeterministic-eliminativist strategy' against BKS, which thwarts ontic contextuality but elevates a strongly epistemic variant of it, making it a kind of illusion, due to an assumed 'subquantum heat death' (Valentini, 2010).

---

[22] One of the structural consequences of nullifying the contextuality and invoking counterfactuality are also, not unexpectedly, problems with the many-worlds understanding of the meaning of nonlocal correlations.

A vindication of **DOR** and erasing the indefiniteness *in a single-world ontology* thus have some unpleasant consequences, which are paradigmatically manifest in the Bohmian case. One may evaluate it as the prevailing structural dualism of the latter. In particular, one gets: (*i*) the theoretical postulate of the global world definiteness, i.e. global $\{0, 1\}$-valuation on observables, but an excessively huge amount of information forever hidden, together with the necessary causal fine-tuning for observed correlations (Wood and Spekkens, 2015) which operates on the cosmological scale; (*ii*) an *ad hoc* or 'brutal' segregation of observables into contextual ('false') and non-contextual ones (the hidden real positions of Bohmian 'particles'), where the phenomenality of the former also requires cosmological fine-tuning (Cavalcanti, 2018); (*iii*) a forced divorce between the illusory or epistemic local 'contextuality' and the essential, genuinely ontological form of it—nonlocal action-at-a-distance (Dürr, Goldstein and Zanghì, 2013, pp.154–156); (*iv*) the hidden ultimate, definite actuality of Ψ ('nothing is really potential'), timelessness and the deterministic supercausal order of the postulated subquantum regime, but pseudo-classical possibility, pseudo-randomness and the illusory temporal evolution on the accessible phenomenal level (Dürr, Goldstein and Zanghì, 2013, pp.268–271); and (*v*) absolute space with absolute rest concealed by phenomenal relativistic spacetime.

In fact, the Bohmian theory insists that almost everything *taken to be* real (ontically noncontextual) is perfectly hidden, and what is *empirically* noncontextual, Gleason's no-disturbance and the Born rule (hence no-cloning, no-deleting etc.), is held to be really violated, albeit not empirically (e.g. Valentini, 2010). If one only knew the timeless world-Ψ and the total configuration of Bohmian 'particles', one *would be* an Absolute Observer, but the theory already makes this forbidden for all local agents trapped in the supposed contingent

equilibrium. It looks extremely Ptolemaic as a construction, from the point of view of Gleason's Theorem and BKS, and if the status of quantum potentiality is like reversal of classical ontology, then the Bohmian theory is like a forced re-reversal in order to recover **D**, which produces (*i*)–(*v*).

The fundamental metaphysical convictions such as (in)determinism or the meaning of 'being real' cannot be falsified[23], but some particular interpretations of contextuality have nontrivial physical consequences and are also logically connected to other physical assumptions as well as empirical constraints. These metaphysical and physical assumptions work together, clustered in some consistent metaphysical packages (e.g. Borges et al., 2014), hence the real problem is: how high is the price for the metaphysical package we may prefer? It appears that if one is inclined to have **DOR** intact and ontic indefiniteness eliminated, one must pay a very high price indeed.

Interestingly, that rather specific predicament sheds additional light on the difficult relation between physics and metaphysics. There seem to be substantial historical analogies between the introduction of the quantum sheaf structure $\Phi_{\mathcal{L}}$ into microphysics and the equally revolutionary instillment of non-Euclidean geometry in the physics of spacetime. In the latter case, there have been many attempts to recast the theory of relativity in such a way that the Newtonian ontology of absolute space and time might be recouped, using e.g. the Lorentz–Kelvin ether theory and the ether interpretations of Einstein's equations. It turned out that such metaphysical motivation brings forth not just the recasting of relativity as some interpretation but also different physical theories, which are unable to compete with general

---

[23] Among the founders of quantum theory it was e.g. Schrödinger who stressed that point and agnosticism regarding the (in)determinism, see (Schrödinger, 1935).

relativity.[24] The introduction of the nonclassical $\Phi_{\mathcal{L}}$ under the guise of $\mathfrak{L}(\mathcal{H})$ has also provoked some attempts to recast it so as to regain a sort of pre-quantum ontology based on **D**. Both quantum theory and relativity are, of course, falsifiable, but the recurring point is that the metaphysical packages associated with them should be dealt with seriously, proportionally to the experimental support, since they are not just a marginal speculative excess baggage. We suggest that to the extent that metaphysics is interested in the goings-on in the physical world, it is, from the very outset, the 'twisted' sheaf structure $\Phi_{\mathcal{L}}$ itself that should be treated in all seriousness as a principal guide. In this case, it also seems that it is pure metaphysics itself which may profit from the pressure imparted by physics (even if falsified at some point in the future), since the latter invites to construct new accounts of potentiality which are not one-way dependent on the notion of determinacy.

5. Conclusion

Quantum contextuality is a tough problem for any metaphysical realist, since it puts very restrictive constraints on standard realist strategies which are usually molded after their classical counterparts, for example by sticking to the classical notions of actual state and possibility, like possible spatio-temporal configuration or the intrinsic disposition of a system. As I have stressed, the free inventiveness of metaphysics is illimitable, however the interesting case of BKS is that the constructive costs – both metaphysical and physical – of adopting these

---

[24] For example, they do not possess singularities (due to the trivial topology) or homogenous curved universes, and introduce *ad hoc* some *completely unobservable* scalar fields.

different strategies vary rather markedly, which makes a choice between them not entirely equivalent or just a matter of personal taste. In particular, the more one pushes toward retrieving **DOR** and turning quantum probability into epistemic uncertainty, the more complex and hidden ('Ptolemaic') the empirically inaccessible sector of a model becomes, whether explosivist or superdeterministic with fine-tuning and the violation of relativity.

If quantum potentiality is the real indefiniteness *in mundo*, not *in mente*, and **DOR** collapses as a basic intuition, then quantum probability itself is a completely novel type of ontic probability, the probability *sui generis* of physically indefinite properties[25], which has nothing to do with classical ignorance, determinate counterfactuals, frequencies, or decision making (Barnum et al., 2000). This is a challenge for naturalistic metaphysics, since even philosophical theories of dispositions are constructed with the tacit assumption of **D**, where complete sets of dispositional properties are clumped in objects or regions of spacetime as some definite 'elements of reality' producing causal responses. If one takes quantum potentiality to be real, **OR** might still be perfectly true as a necessary condition of realism *per se*, there *is* indeed a fully objective reality, but **D** is in general false, and Specker's 'Absolute Observer' is an utter fiction, together with the corresponding counterfactual logic of possibility.

As far as pure ontology is concerned, this is intriguing, since the classical framework for possibility also takes the latter to be completely powerless as a faint Leibnizian shadow of actuality. If the relation is reversed, it is the indefinite *potentia* that has full objectiv-

---

[25] That was already recognized by e.g. von Neumann in the 1930s (von Neumann, 1962, pp.196–197) or Pauli (1994) without relying on any metaphysical analysis. Von Neumann was the first who studied deep connection between quantum structure $\mathcal{L}_{\mathfrak{L}(\mathcal{H})}$ and generic quantum probability in the framework of what he called 'probability logics'.

ity and the power for free actualization, while the definite *actual* is derivative and powerless, like a much poorer shadow. It is impossible to derive the former from the latter, as it is impossible to acquire the quantum probabilistic measure and quantum correlations in space and time from the (Boolean-)Kolmogorovian measure and shared classical pseudo-randomness (Bub, 2014). Thus, the question is: in metaphysics could we accept that autonomous cause or causal power as an agential entity would be entirely non-spatiotemporal and indeterminate? Here I propose that indefiniteness *in mundo* should be indeed taken as both objective potentiality-power and the generic resource for quantum advantage.

Bibliography

Abbott, A.A., Calude, C.S. and Svozil, K., 2014. Value-indefinite observables are almost everywhere. *Physical Review A* [Online], 89(3), p.032109. Available at: https://doi.org/10.1103/PhysRevA.89.032109 [visited on 24 November 2020].

Abramsky, S. and Brandenburger, A., 2011. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics* [Online], 13(11), p.113036. Available at: https://doi.org/10.1088/1367-2630/13/11/113036 [visited on 24 November 2020].

Acín, A., Fritz, T., Leverrier, A. and Sainz, A.B., 2015. A combinatorial approach to nonlocality and contextuality. *Communications in Mathematical Physics* [Online], 334(2), pp.533–628. Available at: https://doi.org/10.1007/s00220-014-2260-1 [visited on 24 November 2020].

Ahrens, J., Amselem, E., Cabello, A. and Bourennane, M., 2013. Two fundamental experimental tests of nonclassicality with qutrits. *Scientific Reports* [Online], 3(1), p.2170. Available at: https://doi.org/10.1038/srep02170 [visited on 24 November 2020].

Akiba, K., 2004. Vagueness in the world. *Noûs* [Online], 38(3), pp.407–429. Available at: https://www.jstor.org/stable/3506246 [visited on 24 November 2020].

Amselem, E., Rådmark, M., Bourennane, M. and Cabello, A., 2009. State-independent quantum contextuality with single photons. *Physical Review Letters* [Online], 103(16), p.160405. Available at: https://doi.org/10.1103/PhysRevLett.103.160405 [visited on 24 November 2020].

Aolita, L. et al., 2012. Fully nonlocal quantum correlations. *Physical Review A* [Online], 85(3), p.032107. Available at: https://doi.org/10.1103/PhysRevA.85.032107 [visited on 24 November 2020].

Barnes, E. and Williams, J.R.G., 2011. A theory of metaphysical indeterminacy. In: K. Bennett and D.W. Zimmerman, eds. *Oxford Studies in Metaphysics. Vol. 6* [Online]. Oxford: Oxford University Press, pp.103–148. Available at: https://doi.org/10.1093/acprof:oso/9780199603039.003.0003 [visited on 24 November 2020].

Barnum, H. et al., 2000. Quantum probability from decision theory? *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* [Online], 456(1997), pp.1175–1182. Available at: https://doi.org/10.1098/rspa.2000.0557 [visited on 24 November 2020].

Barrett, J., Cavalcanti, E.G., Lal, R. and Maroney, O.J.E., 2014. No ψ-epistemic model can fully explain the indistinguishability of quantum states. *Physical Review Letters* [Online], 112(25), p.250403. Available at: https://doi.org/10.1103/PhysRevLett.112.250403 [visited on 24 November 2020].

Bartlett, S.D., Rudolph, T. and Spekkens, R.W., 2012. Reconstruction of Gaussian quantum mechanics from Liouville mechanics with an epistemic restriction. *Physical Review A* [Online], 86(1), p.012103. Available at: https://doi.org/10.1103/PhysRevA.86.012103 [visited on 24 November 2020].

Bartosik, H. et al., 2009. Experimental test of quantum contextuality in neutron interferometry. *Physical Review Letters* [Online], 103(4), p.040403. Available at: https://doi.org/10.1103/PhysRevLett.103.040403 [visited on 24 November 2020].

Bell, J.S., 1964. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1, pp.195–200.

Bell, J.S., 1966. On the problem of hidden variables in quantum mechanics. *Reviews of Modern Physics*, 38(3), pp.447–452.

Bell, J.S., 1981. Bertlmann's socks and the nature of reality. *Le Journal de Physique Colloques* [Online], 42(C2), pp.C2–41–C2–62. Available at: https://doi.org/10.1051/jphyscol:1981202 [visited on 24 November 2020].

Bella, S., 2005. *The Science of the Individual: Leibniz's Ontology of Individual Substance* [Online], *Topoi Library*. Dordrecht: Springer. Available at: https://doi.org/10.1007/1-4020-3260-9.

Bermejo-Vega, J. et al., 2017. Contextuality as a resource for models of quantum computation with qubits. *Physical Review Letters* [Online], 119(12), p.120505. Available at: https://doi.org/10.1103/PhysRevLett.119.120505 [visited on 24 November 2020].

Bohr, N., 1935. Can quantum-mechanical description of physical reality be considered complete? *Physical Review* [Online], 48(8), pp.696–702. Available at: https://doi.org/10.1103/PhysRev.48.696 [visited on 24 November 2020].

Borges, G. et al., 2014. Quantum contextuality in a Young-type interference experiment. *Physical Review A* [Online], 89(5), p.052106. Available at: https://doi.org/10.1103/PhysRevA.89.052106 [visited on 24 November 2020].

Branciard, C., 2014. How ψ-epistemic models fail at explaining the indistinguishability of quantum states. *Physical Review Letters* [Online], 113, p.020409. Available at: https://doi.org/10.1103/PhysRevLett.113.020409 [visited on 24 November 2020].

Bruss, D., D'Ariano, G.M., Macchiavello, C. and Sacchi, M.F., 2000. Approximate quantum cloning and the impossibility of superluminal information transfer. *Physical Review A* [Online], 62(6), p.062302. Available at: https://doi.org/10.1103/PhysRevA.62.062302 [visited on 24 November 2020].

Bub, J., 2014. Quantum correlations and the measurement problem. *International Journal of Theoretical Physics* [Online], 53, pp.3346–3369. Available at: https://doi.org/10.1007/s10773-013-1695-z [visited on 24 November 2020].

Cabello, A., 1999a. Quantum correlations are not contained in the initial state. *Physical Review A* [Online], 60(2), pp.877–880. Available at: https://doi.org/10.1103/PhysRevA.60.877 [visited on 24 November 2020].

Cabello, A., 1999b. Quantum correlations are not local elements of reality. *Physical Review A* [Online], 59(1), pp.113–115. Available at: https://doi.org/10.1103/PhysRevA.59.113 [visited on 24 November 2020].

Cañas, G. et al., 2014. Experimental implementation of an eight-dimensional Kochen-Specker set and observation of its connection with the Greenberger-Horne-Zeilinger theorem. *Physical Review A* [Online], 90(1), p.012119. Available at: https://doi.org/10.1103/PhysRevA.90.012119 [visited on 24 November 2020].

Cavalcanti, E.G., 2018. Classical causal models for Bell and Kochen-Specker inequality violations require fine-tuning. *Physical Review X* [Online], 8(2), p.021018. Available at: https://doi.org/10.1103/PhysRevX.8.021018 [visited on 24 November 2020].

Charlton, W., 1992. Appendix: Did Aristotle believe in prime matter? In: Aristotle. *Physics. Books I and II*, *Clarendon Aristotle series*. Transl. with introd., commentary, note on recent work, and rev. bibliogr. by William Charlton. Oxford: Clarendon Press, 1992, pp.129–145.

Chaves, R. and Fritz, T., 2012. Entropic approach to local realism and non-contextuality. *Physical Review A* [Online], 85(3), p.032113. Available at: https://doi.org/10.1103/PhysRevA.85.032113 [visited on 24 November 2020].

D'Ariano, G.M. and Yuen, H.P., 1996. Impossibility of measuring the wave function of a single quantum system. *Physical Review Letters* [Online], 76(16), pp.2832–2835. Available at: https://doi.org/10.1103/PhysRevLett.76.2832 [visited on 24 November 2020].

D'Ambrosio, V. et al., 2013. Experimental implementation of a Kochen-Specker set of quantum tests. *Physical Review X* [Online], 3(1), p.011012. Available at: https://doi.org/10.1103/PhysRevX.3.011012 [visited on 24 November 2020].

Dhar, H.S. et al., 2017. Monogamy of quantum correlations - a review. In: F.F. Fanchini, S. Pinto, D. de Oliveira and G. Adesso, eds. *Lectures on General Quantum Correlations and their Applications* [Online], *Quantum Science and Technology*. Berlin – New York: Springer International Publishing, pp.23–64. Available at: https://doi.org/10.1007/978-3-319-53412-1_3 [visited on 24 November 2020].

Dogra, S., Dorai, K. and Arvind, 2016. Experimental demonstration of quantum contextuality on an NMR qutrit. *Physics Letters A* [Online], 380(22), pp.1941–1946. Available at: https://doi.org/10.1016/j.physleta.2016.04.015 [visited on 24 November 2020].

Döring, A., 2005. Kochen–Specker theorem for von Neumann algebras. *International Journal of Theoretical Physics* [Online], 44(2), pp.139–160. Available at: https://doi.org/10.1007/s10773-005-1490-6 [visited on 24 November 2020].

Dummett, M., 1975. Wang's paradox. *Synthese* [Online], 30(3), pp.301–324. Available at: https://doi.org/10.1007/BF00485048 [visited on 24 November 2020].

Dürr, D., Goldstein, S. and Zanghì, N., 2013. *Quantum Physics without Quantum Philosophy* [Online]. Berlin – Heidelberg: Springer-Verlag. Available at: https://doi.org/10.1007/978-3-642-30690-7 [visited on 24 November 2020].

French, S., 2014. *The Structure of the World: Metaphysics and Representation*. Oxford: Oxford University Press.

Gisin, N., 1998. Quantum cloning without signaling. *Physics Letters A* [Online], 242(1), pp.1–3. Available at: https://doi.org/10.1016/S0375-9601(98)00170-4 [visited on 24 November 2020].

Gleason, A.M., 1957. Measures on the closed subspaces of a Hilbert space. *Journal of Mathematics and Mechanics* [Online], 6(6), pp.885–893. Available at: https://www.jstor.org/stable/24900629 [visited on 25 November 2020].

Hawley, K., 2006. Science as a guide to metaphysics? *Synthese* [Online], 149(3), pp.451–470. Available at: https://doi.org/10.1007/s11229-005-0569-1 [visited on 25 November 2020].

Hermens, R., 2011. The problem of contextuality and the impossibility of experimental metaphysics thereof. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* [Online], 42(4), pp.214–225. Available at: https://doi.org/10.1016/j.shpsb.2011.06.001 [visited on 25 November 2020].

Heunen, C., Landsman, N.P. and Spitters, B., 2012. Bohrification of operator algebras and quantum logic. *Synthese* [Online], 186(3), pp.719–752. Available at: https://doi.org/10.1007/s11229-011-9918-4 [visited on 25 November 2020].

Horodecki, K. et al., 2015. Axiomatic approach to contextuality and non-locality. *Physical Review A* [Online], 92(3), p.032104. Available at: https://doi.org/10.1103/PhysRevA.92.032104 [visited on 25 November 2020].

Howard, M., Wallman, J., Veitch, V. and Emerson, J., 2014. Contextuality supplies the 'magic' for quantum computation. *Nature* [Online], 510(7505), pp.351–355. Available at: https://doi.org/10.1038/nature13460 [visited on 25 November 2020].

Huang, Y.-F. et al., 2013. Experimental test of state-independent quantum contextuality of an indivisible quantum system. *Physical Review A* [Online], 87(5), p.052133. Available at: https://doi.org/10.1103/PhysRevA.87.052133 [visited on 25 November 2020].

Jia, Z.-A. et al., 2018. Entropic no-disturbance as a physical principle. *Physical Review A* [Online], 97(5), p.052128. Available at: https://doi.org/10.1103/PhysRevA.97.052128 [visited on 25 November 2020].

Kirchmair, G. et al., 2009. State-independent experimental test of quantum contextuality. *Nature* [Online], 460(7254), pp.494–497. Available at: https://doi.org/10.1038/nature08172 [visited on 25 November 2020].

Knoebel, A., 2012. *Sheaves of Algebras over Boolean Spaces*. New York: Birkhäuser.

Kochen, S. and Specker, E.P., 1967. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics* [Online], 17(1), pp.59–87. Available at: https://www.jstor.org/stable/24902153 [visited on 25 November 2020].

Konig, R., Renner, R. and Schaffner, C., 2009. The operational meaning of min- and max-entropy. *IEEE Transactions on Information Theory* [Online], 55(9), pp.4337–4347. Available at: https://doi.org/10.1109/TIT.2009.2025545.

Kurzyński, P., Ramanathan, R. and Kaszlikowski, D., 2012. Entropic test of quantum contextuality. *Physical Review Letters* [Online], 109(2), p.020404. Available at: https://doi.org/10.1103/PhysRevLett.109.020404 [visited on 25 November 2020].

Lapkiewicz, R. et al., 2011. Experimental non-classicality of an indivisible quantum system. *Nature* [Online], 474(7352), pp.490–493. Available at: https://doi.org/10.1038/nature10119 [visited on 25 November 2020].

Leibniz, G.W., 2009. *Sämtliche Schriften und Briefe. Zweite Reihe, Band 2: Philosophischer Briefwechsel, 1686-1694* [Online]. Berlin: Akademie Verlag. Available at: https://doi.org/10.26015/ADWDOCS-142 [visited on 25 November 2020].

Leifer, M.S., 2014. Ψ-epistemic models are exponentially bad at explaining the distinguishability of quantum states. *Physical Review Letters* [Online], 112(16), p.160404. Available at: https://doi.org/10.1103/PhysRevLett.112.160404 [visited on 25 November 2020].

Lenzen, W., 1984. Leibniz und die Boolesche Algebra. *Studia Leibnitiana* [Online], 16(2), pp.187–203. Available at: https://www.jstor.org/stable/40694003 [visited on 25 November 2020].

Leupold, F.M. et al., 2018. Sustained state-independent quantum contextual correlations from a single ion. *Physical Review Letters* [Online], 120(18), p.180401. Available at: https://doi.org/10.1103/PhysRevLett.120.180401 [visited on 25 November 2020].

Lewis, D.K., 1986. *On the Plurality of Worlds*. Oxford: Basil Blackwell.

Liao, K.-Y. et al., 2016. Experimental test of the no-go theorem for continuous ψ-epistemic models. *Scientific Reports* [Online], 6(1), p.26519. Available at: https://doi.org/10.1038/srep26519 [visited on 25 November 2020].

Liu, B.H. et al., 2009. Experimental demonstration of quantum contextuality with nonentangled photons. *Physical Review A* [Online], 80(4), p.044101. Available at: https://doi.org/10.1103/PhysRevA.80.044101 [visited on 25 November 2020].

Lostaglio, M. and Senno, G., 2020. Contextual advantage for state-dependent cloning. *Quantum* [Online], 4, p.258. Available at: https://doi.org/10.22331/q-2020-04-27-258 [visited on 25 November 2020].

Mac Lane, S. and Moerdijk, I., 1992. *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*, *Universitext*. New York: Springer-Verlag.

Malink, M. and Vasudevan, A., 2016. The logic of Leibniz's *Generales inquisitiones de analysi notionum et veritatum*. *The Review of Symbolic Logic* [Online], 9(4), pp.686–751. Available at: https://doi.org/10.1017/S1755020316000137 [visited on 25 November 2020].

Marques, B. et al., 2014. Experimental observation of Hardy-like quantum contextuality. *Physical Review Letters* [Online], 113(25), p.250403. Available at: https://doi.org/10.1103/PhysRevLett.113.250403 [visited on 25 November 2020].

Masanes, L., Acin, A. and Gisin, N., 2006. General properties of nonsignaling theories. *Physical Review A* [Online], 73(1), p.012112. Available at: https://doi.org/10.1103/PhysRevA.73.012112 [visited on 25 November 2020].

Morreau, M., 2002. What vague objects are like. *The Journal of Philosophy* [Online], 99(7), pp.333–361. Available at: https://doi.org/10.2307/3655512 [visited on 25 November 2020].

von Neumann, J., 1962. Quantum logics. (Strict- and probability-logics). In: A.H. Taub, ed. *Collected Works. Vol. 4: Continuous Geometry and Other Topics*. Oxford: Pergamon Press, pp.195–197.

Pati, A.K. and Braunstein, S.L., 2003. Quantum deleting and signalling. *Physics Letters A* [Online], 315(3), pp.208–212. Available at: https://doi.org/10.1016/S0375-9601(03)01047-8 [visited on 25 November 2020].

Pati, A.K., 2000. Probabilistic exact cloning and probabilistic no-signalling. *Physics Letters A* [Online], 270(3), pp.103–107. Available at: https://doi.org/10.1016/S0375-9601(00)00281-4 [visited on 25 November 2020].

Pauli, W., 1994. Probability and physics. In: C.P. Enz and K. von Meyenn, eds. *Writings on physics and philosophy* (R. Schlapp, Trans.). Berlin: Springer, pp.43–48.

Popper, K.R., 1967. Quantum mechanics without 'the observer'. In: M.A. Bunge, ed. *Quantum Theory and Reality*, *Studies in the Foundations, Methodology and Philosophy of Science* 2. Berlin: Springer-Verlag, pp.7–43.

Qu, D. et al., 2020. Experimental entropic test of state-independent contextuality via single photons. *Physical Review A* [Online], 101(6), p.060101. Available at: https://doi.org/10.1103/PhysRevA.101.060101 [visited on 25 November 2020].

Raussendorf, R., 2013. Contextuality in measurement-based quantum computation. *Physical Review A* [Online], 88(2), p.022322. Available at: https://doi.org/10.1103/PhysRevA.88.022322 [visited on 25 November 2020].

Ringbauer, M. et al., 2015. Measurements on the reality of the wavefunction. *Nature Physics* [Online], 11(3), pp.249–254. Available at: https://doi.org/10.1038/nphys3233 [visited on 25 November 2020].

de Ronde, C., 2017. Unscrambling the omelette of quantum contextuality (part 1): Preexistent properties or measurement outcomes? *arXiv:1606.03967 [physics, physics:quant-ph]* [Online]. Available at: http://arxiv.org/abs/1606.03967 [visited on 24 November 2020].

Schmid, D. and Spekkens, R.W., 2018. Contextual advantage for state discrimination. *Physical Review X* [Online], 8(1), p.011015. Available at: https://doi.org/10.1103/PhysRevX.8.011015 [visited on 25 November 2020].

Schrödinger, E., 1935. *Science and the Human Temperament*. London: G. Allen & Unwin Limited.

Singh, D., Singh, J., Dorai, K. and Arvind, 2019. Experimental demonstration of fully contextual quantum correlations on an NMR quantum informa-

tion processor. *Physical Review A* [Online], 100(2), p.022109. Available at: https://doi.org/10.1103/PhysRevA.100.022109 [visited on 25 November 2020].

Specker, E., 1960. Die Logik nicht gleichzeitig entscheidbarer Aussagen. *Dialectica* [Online], 14(2-3), pp.239–246. Available at: https://doi.org/https://doi.org/10.1111/j.1746-8361.1960.tb00422.x [visited on 25 November 2020].

de Spinoza, B., 1977. *Die Ethik: Lateinisch und Deutsch*. Revised transl. by Stern, J., *Reclams Universal-Bibliothek* Nr. 851. Stuttgart: Reclam.

Valentini, A., 2010. De Broglie–Bohm pilot-wave theory: Many worlds in denial? In: S. Saunders, J. Barrett, A. Kent and D. Wallace, eds. *Many Worlds?: Everett, Quantum Theory, and Reality* [Online]. Oxford: Oxford University Press, pp.476–509. Available at: https://doi.org/10.1093/acprof:oso/9780199560561.003.0019 [visited on 25 November 2020].

Wood, C.J. and Spekkens, R.W., 2015. The lesson of causal discovery algorithms for quantum correlations: causal explanations of Bell-inequality violations require fine-tuning. *New Journal of Physics* [Online], 17(3), p.033002. Available at: https://doi.org/10.1088/1367-2630/17/3/033002 [visited on 25 November 2020].

Xiao, Y. et al., 2018. Experimental observation of quantum state-independent contextuality under no-signaling conditions. *Optics Express* [Online], 26(1), pp.32–50. Available at: https://doi.org/10.1364/OE.26.000032 [visited on 25 November 2020].

Zhang, X. et al., 2013. State-independent experimental test of quantum contextuality with a single trapped ion. *Physical Review Letters* [Online], 110(7), p.070401. Available at: https://doi.org/10.1103/PhysRevLett.110.070401 [visited on 25 November 2020].

Zu, C. et al., 2012. State-independent experimental test of quantum contextuality in an indivisible system. *Physical Review Letters* [Online], 109(15), p.150401. Available at: https://doi.org/10.1103/PhysRevLett.109.150401 [visited on 25 November 2020].

Quantum geometry, logic and probability

Shahn Majid
School of Mathematics, Queen Mary University of London

Abstract

Quantum geometry on a discrete set means a directed graph with a weight associated to each arrow defining the quantum metric. However, these 'lattice spacing' weights do not have to be independent of the direction of the arrow. We use this greater freedom to give a quantum geometric interpretation of discrete Markov processes with transition probabilities as arrow weights, namely taking the diffusion form $\partial_+ f = (-\Delta_\theta + q - p)f$ for the graph Laplacian Δ_θ, potential functions q, p built from the probabilities, and finite difference ∂_+ in the time direction. Motivated by this new point of view, we introduce a 'discrete Schrödinger process' as $\partial_+ \psi = \imath(-\Delta + V)\psi$ for the Laplacian associated to a bimodule connection such that the discrete evolution is unitary. We solve this explicitly for the 2-state graph, finding a 1-parameter family of such connections and an induced 'generalised Markov process' for $f = |\psi|^2$ in which there is an additional source current built from ψ. We also mention our recent work on the quantum geometry of logic in 'digital' form over the field $\mathbb{F}_2 = \{0, 1\}$, including de Morgan duality and its possible generalisations.

1. Introduction

These are notes growing out of a conference in Kraków with the wonderful question "Is logic physics?". My view is yes, obviously. Indeed, I have long proposed (Majid, 1991) to think of Boolean algebras as the 'simplest theory of physics', in which case we should be able to see how physical and geometric issues in more advanced theories emerge from structures already present there. If so, then we can see the origin of physics in the very nature of language and the structure of mathematical and physical discourse, a philosophy that I have espoused as *relative realism* (Majid, 2012; 1991; 2015; 2014).

One of the things particularly to be explored in this way is the evident role of entropy in the Einstein equations and its deep link with gravity as an ingredient of quantum gravity. My proposal is that while we may not be able to address this in quantum gravity itself, we can try to move quantum gravity ideas back to simplified settings such as finite graphs (Majid, 2019a) and Boolean algebras or digital geometry (Majid and Pachoł, 2020a; Majid, 2020a) and see if it can be addressed there. The present paper is a step in this direction towards entropy and gravity: we see how probability and irreversible processes can be seen naturally as emerging from quantum geometry.

Our main new results are in Section 3 in the context of graphs, working over $\mathbb{C}$ and $\mathbb{R}$. The idea is to explore a curious feature of quantum geometry applied in the discrete case, namely that the lattice 'square length' $x \to y$ need not be same as for $y \to x$. In usual lattice geometry, we would be focussed on the edge-symmetric case where these coincide. I will argue that the asymmetric generalisation is the

natural setting for probability and non-reversible processes. Thus, we consider a Markov process

$$f^{new}(x) = \sum_{y \to x} p_{y \to x} f(y) + (1 - p(x)) f(x),$$

where $p_{x \to y}$ denote the conditional transition probability from x to y and $p(x) = \sum_{y:x \to y} p_{x \to y}$ for each x. If we let ∂_+ be the finite difference on $\mathbb{Z}$ as the 'discrete time' then the Markov process appears as

$$\partial_+ f = -\Delta f + (q - p)f,$$

where $q(y) = \sum_{x:x \to y} p_{x \to y}$ and Δ is the canonical graph Laplacian associated to a (typically degenerate) metric inner product defined by the $p_{x \to y}$, in which context it is natural that the value for $x \to y$ need not be the same as for $y \to x$. In fact, the quantities that we should really think of more geometrically as 'length' are the negative logs $\lambda_{x \to y}$ of the probabilities, then the shortest length path in that sense is indeed the path of maximum probability. Moreover, the shortest length between two points then makes the space into a Lawvere metric space (Fong and Spivak, 2019). Our point of view is that this is not so much a 'generalisation' as what naturally arises in quantum geometry for a discrete space; the real question is why distances are not direction dependent in the continuum limit. Indeed, it could be argued that many processes are irreversible, so already time between events *is* measured in a directed way. In the present context, this is reflected slightly differently in our probabilistic interpretation of 'distance' as conditional transition probabilities.

Finally, our point of view leads us to introduce a 'discrete Schrödinger's process'

$$\partial_+ \psi = \imath(-\Delta + V)\psi$$

and to ask what connections ∇ are 'unitary' in the sense that $f = |\psi|^2$ remains a probability density when we use the associated Laplacian. We find (see Proposition 3.2) that the discreteness results in the evolution of f being a generalised Markov process coupled to a probability current source. Indeed, the discrete Schrödinger process being unitary is reversible, so the irreversibility of the Markov process on f is compensated by this source. We construct such a unitary Schrödinger process explicitly for a 2-state graph $\bullet \longleftrightarrow \bullet$, finding a 1-parameter family of suitable connections. Unitary matrices are the essence of quantum computing and what we have done is to construct a certain family of them quantum geometrically in the spirit of Schrödinger's equation. Another motivation is the recent formulation of 'quantum geodesics' (Beggs, 2020; Beggs and Majid, 2019) which indeed includes the actual Schrödinger's equation of ordinary quantum mechanics. This work is not immediately applicable due to our current use of discrete time, but these ideas should tie up in future work.

We will use and make reference to a modern constructive 'quantum geometry' formalism briefly outlined in Section 2; see (Beggs and Majid, 2020) for details and the extensive literature cited therein. This approach starts with a unital algebra A and choice of differential structure (Ω^1, d) on it and then proceeds to metrics, connections, curvature and so forth. Growing out of many years experience with quantum groups but in no way limited to them, it is very different in flavour from the well-known approach of A. Connes (Connes, 1994) coming out of cyclic cohomology and 'spectral triples' as abstract 'Dirac operators', though with the possibility of useful interaction between the approaches (Beggs and Majid, 2017).

Another aspect of the constructive approach is that it works over any field, which makes possible the 'digital' case by working over the
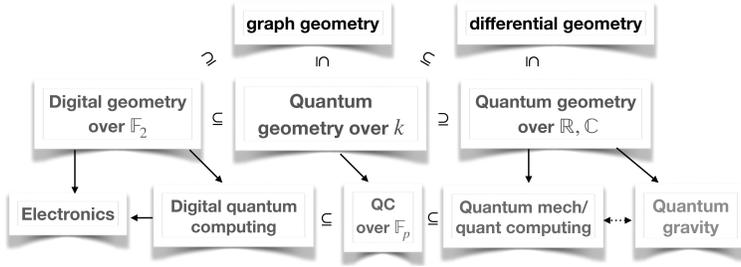
Figure 1: Quantum gravity and quantum mechanics/computing mediated by quantum geometry over $\mathbb{C}$. Digital geometry as a limiting case.

field $\mathbb{F}_2 = \{0, 1\}$ of two elements (Bassett and Majid, 2020; Majid and Pachoł, 2018; 2020a) and allows in principle the transfer of geometric ideas to digital electronics, see Figure 1. A brief overview is in Section 4 with some modest new results for the quantum geometry of $A = M_2(\mathbb{F}_2)$. In the same vein, one could in theory put something like a digital wave operator for a black hole background onto a silicon chip. A motivation here is again from quantum computing. While in quantum computing, a gate is replaced by a unitary operator (which in turn we envisage could be constructed quantum geometrically, e.g. by a Schrödinger process), the essential feature here is the use of vector spaces (the superposition principle) to massively parallelise computations. However, linear algebra works over any field. Hence, if we build our gates quantum geometrically then we could specialise them over other fields, including over $\mathbb{F}_2$ as 'digital quantum computing'. Even if this did not have the speed benefits of actual quantum computing, it would provide conventionally realisable training wheels for the real thing.

2. Outline of quantum geometry

It is well-known that classical geometry can be formulated equivalently in terms of a suitable algebra of functions on the space. The idea in noncommutative or 'quantum' geometry is to allow this to be any algebra A with identity as our starting point (and now no actual space need exist). We replace the notion of differential structure on a space by specifying a bimodule Ω^1 of differential forms over A. A bimodule means we can multiply a '1-form' $\omega \in \Omega^1$ by 'functions' $a, b \in A$ either from the left or the right and the two should associate according to

(1) $(a\omega)b = a(\omega b).$

We also need $\mathrm{d} : A \to \Omega^1$ an 'exterior derivative' obeying reasonable axioms, the most important of which is the Leibniz rule

(2) $\mathrm{d}(ab) = (\mathrm{d}a)b + a(\mathrm{d}b)$

for all $a, b \in A$. We usually require Ω^1 to extend to forms of higher degree to give a graded algebra $\Omega = \oplus \Omega^i$ (where associativity extends the bimodule identity (1) to higher degree). We also require d to extend to $\mathrm{d} : \Omega^i \to \Omega^{i+1}$ obeying a graded-Leibniz rule with respect to the graded product $\wedge$ and $\mathrm{d}^2 = 0$. This 'differential structure' is the first choice we have to make in model building once we fixed the algebra A. We require that Ω is generated by $A, \mathrm{d}A$ as it would be classically. Next, on an algebra with differential, we define a metric as an element $g \in \Omega^1 \otimes_A \Omega^1$ which is invertible in the sense of a map $(\ ,\ ) : \Omega^1 \otimes_A \Omega^1 \to A$ that commutes with the product by A from the left or right and inverts g in the sense

(3) $((\omega,\ ) \otimes_A \mathrm{id})g = \omega = (\mathrm{id} \otimes_A (\ ,\omega))g$
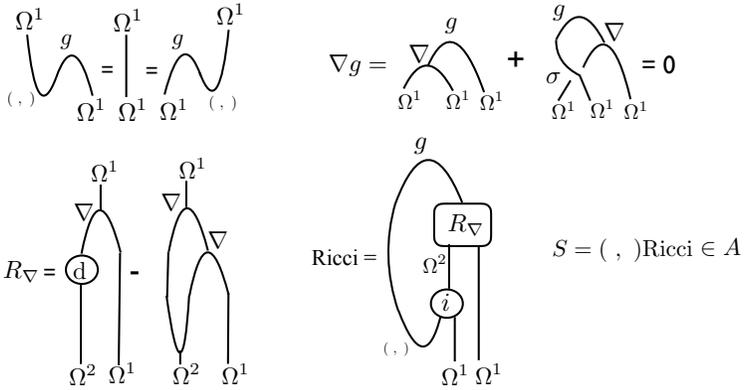
Figure 2: Some axioms of quantum Riemannian geometry. In order: invertibility of a metric, tensor product bimodule connection, Riemann curvature and Ricci curvature tensor and scalar. Diagrams are read down the page as a series of compositions.

for all 1-forms ω. This is shown in Figure 2. In the general theory, one can require quantum symmetry in the form $\wedge(g) = 0$, where we consider the wedge product on 1-forms as a map $\wedge : \Omega^1 \otimes_A \Omega^1 \to A$ and apply this to g. However, we don't need to require this and moreover we can also work with g non-invertible or $(\ ,\ )$ degenerate.

Finally, we need the notion of a connection. A left connection on Ω^1 is a linear map $\nabla : \Omega^1 \to \Omega^1 \otimes_A \Omega^1$ obeying a left-Leibniz rule

(4) $$\nabla(a\omega) = \mathrm{d}a \otimes_A \omega + a\nabla\omega$$

for all $a \in A, \omega \in \Omega^1$. This might seem mysterious but if we think of a map $X : \Omega^1 \to A$ that commutes with the right action by A as a 'vector field' then we can evaluate ∇ as a covariant derivative $\nabla_X = (X \otimes_A \mathrm{id})\nabla : \Omega^1 \to \Omega^1$ which classically is then a usual

covariant derivative on Ω^1. There is a similar notion for a connection on a general 'vector bundle' expressed algebraically. Moreover, when we have both left and right actions of A forming a bimodule, as we do here, we say that a left connection is a *bimodule connection* (Dubois-Violette and Michor, 1996; Beggs and Majid, 2014), if there also exists a bimodule map σ such that

(5) $\sigma : \Omega^1 \otimes_A \Omega^1 \to \Omega^1 \otimes_A \Omega^1, \quad \nabla(\omega a) = (\nabla \omega)a + \sigma(\omega \otimes_A \mathrm{d}a)$

for all $a \in A, \omega \in \Omega^1$. The map σ, if it exists, is unique, so this is not additional data but a property that some connections have. The key thing is that bimodule connections extend automatically to tensor products as

(6) $\nabla(\omega \otimes_A \eta) = \nabla \omega \otimes_A \eta + (\sigma(\omega \otimes_A (\ )) \otimes_A \mathrm{id})\nabla \eta$

for all $\omega, \eta \in \Omega^1$, so that metric compatibility now makes sense as $\nabla g = 0$. This is shown in Figure 2. A connection is called QLC or 'quantum Levi-Civita' if it is metric compatible and the torsion also vanishes, which in our language amounts to $\wedge \nabla = \mathrm{d}$ as equality of maps $\Omega^1 \to \Omega^2$. Given a metric inner product $(\ ,\ )$ and a connection ∇, one has divergence and geometric Laplacian

(7) $\nabla \cdot \omega = (\ ,\ )\nabla \omega, \quad \Delta a = (\ ,\ )\nabla \mathrm{d}a$

for all $\omega \in \Omega^1, a \in A$.

    We also have a Riemannian curvature for any connection,

(8) $R_\nabla = (\mathrm{d} \otimes_A \mathrm{id} - \mathrm{id} \wedge \nabla)\nabla : \Omega^1 \to \Omega^2 \otimes_A \Omega^1,$

where classically one would interior product the first factor against a pair of vector fields to get an operator on 1-forms. Ricci requires more data and the current state of the art (but probably not the ultimate way)

is to introduce a lifting bimodule map $i : \Omega^2 \to \Omega^1 \otimes_A \Omega^1$. Applying this to the left output of R_∇, we are then free to 'contract' by using the metric and inverse metric to define $\mathrm{Ricci} \in \Omega^1 \otimes_A \Omega^1$ (Beggs and Majid, 2017). This is also shown in Figure 2.

Finally, and critical for physics, are unitarity or 'reality' properties. We mainly work over $\mathbb{C}$ and assume that A is a $*$-algebra (real functions, classically, would be the self-adjoint elements). We require this to extend to Ω as a graded-anti-involution (so reversing order with an extra sign when odd degree differential forms are involved) and to commute with d. 'Reality' of the metric and of the connection in the sense of being $*$-preserving are imposed as (Beggs and Majid, 2014; Beggs and Majid, 2017)

$$(9) \quad g^\dagger = g, \quad \nabla \circ * = \sigma \circ \dagger \circ \nabla; \quad (\omega \otimes_A \eta)^\dagger = \eta^* \otimes_A \omega^*,$$

where $\dagger$ is a natural $*$-operation on $\Omega^1 \otimes_A \Omega^1$. These 'reality' conditions in a self-adjoint basis (if one exists) and in the classical case would ensure that the metric and connection coefficients are real.

In the case where there exists $\theta \in \Omega^1$ such that $\mathrm{d}a = [\theta, a]$, one says that the calculus is *inner*. Similarly, with graded-commutator in higher degree. This is never possible in the classical case but is rather typical in the quantum case. One then has that any bimodule connection ∇ has the form (Majid, 2013)

$$\nabla \omega = \theta \otimes \omega - \sigma(\omega \otimes \theta) + \alpha(\omega)$$

for some bimodule maps σ, α. A canonical (but not classical) choice is $\sigma = \alpha = 0$ in which case

$$\nabla_\theta = \theta \otimes, \quad T_{\nabla_\theta} = -(\ ) \wedge \theta, \quad R_{\nabla_\theta} = \theta^2 \wedge, \quad \nabla g = \theta \otimes g$$

so that this connection cannot be Levi-Civita for a nontrivial exterior algebra or nontrivial metric. Nevertheless, it defines a canonical divergence and canonical Laplacian according to (7), namely

$$\nabla_\theta \cdot \omega = (\theta, \omega), \quad \Delta_\theta a = (\theta, \mathrm{d}a) = -(\mathrm{d}a, \theta) + [(\theta, \theta), a]$$

which we will use in Section 3 (note that the latter is $\Delta_\theta/2$ in the conventions of (Majid, 2013; Beggs and Majid, 2020)). Finding actual QLCs is a much more involved problem and usually results in a moduli space of ∇ rather than a unique one.

3. Asymmetric discrete geometry and Markov processes

We are interested in the case $A = k(X)$ of functions on a discrete set X with values in a field k. Here we necessarily have $\Omega^1 = \mathrm{span}_k \omega_{x \to y}$ with basis labelled by the edges of a graph on X. The bimodule structure and exterior derivative are

$$f.\omega_{x \to y} = f(x)\omega_{x \to y}, \quad \omega_{x \to y}.f = f(y)\omega_{x \to y},$$
$$\mathrm{d}f = [\theta, f] = \sum_{x \to y} (f(y) - f(x))\omega_{x \to y},$$

where $\theta = \sum_{x \to y} \omega_{x \to y}$. By definition, we don't include self-arrows in the graph (but it can be useful to *extend* the graph to allow them). We assume that our graph is *bidirected* in the sense that if $x \to y$ is an arrow then so is $y \to x$. In this case, we can define a *metric inner product* $(\, , \,) : \Omega^1 \otimes_A \Omega^1 \to A$ as the bimodule map

$$(\omega_{x' \to y'}, \omega_{y \to x}) = \delta_{x',x}\delta_{y',y}p_{y \to x}\delta_x,$$

for some metric weights $p_{y \to x}$. *Unlike usual quantum geometry*, we now do *not* suppose nondegeneracy in the sense that these are all

nonzero. The arrows $x \to y$ for which $p_{x \to y} \neq 0$ are the more relevant subgraph but it is convenient to use the full bidirected graph for Ω^1 with the price that some of the weights could vanish. Given the bimodule inner product, we use the canonical graph Laplacian (Majid, 2013) as above, which works out as

$$(10) \qquad -\Delta_\theta f = (\mathrm{d}f, \theta) = \sum_{x \to y} (f(y) - f(x)) \delta_x p_{y \to x}.$$

For the moment, we work over $\mathbb{R}$ and define functions

$$p, q \in \mathbb{R}(X), \quad p(x) = \sum_{y: x \to y} p_{x \to y}, \quad q(x) = (\theta, \theta) = \sum_{y: x \to y} p_{y \to x},$$

$\forall x, y \in X$. We will assume that $(\ ,\ )$ is *stochastic* in the sense that

$$p_{x \to y} \geq 0, \quad \forall x \to y, \quad p(x) \leq 1, \quad \forall x \in X.$$

In other words, we do noncommutative geometry but with weights in the Heyting algebra $[0, 1]$. In more conventional terms, we define a Markov transition matrix by $P_{x,y} = p_{x \to y}$ if $x \to y$ and $P_{x,x} = 1 - p(x)$ with other entries zero, which is then right stochastic (all rows sum to 1). One could equally extend the graph to allow self-arrows with $p_{x \to x} = 1 - p(x)$ on the extended graph, defining a generalised calculus $\tilde{\Omega}^1$ in the sense of (Majid and Tao, 2019).

We also consider $f \in \mathbb{R}(X)$ a probability distribution so that $f(x)$ is the probability for each event x, or in vector terms $f = (f(x))$ is a stochastic row vector i.e. its elements are ≥ 0 and sum to 1. A Markov process with transition probabilities $p_{x \to y}$ is an evolution of such distributions, i.e., labelled by a step index i, with

$$(11) \quad f_{i+1}(x) = \sum_y f_i(y) P_{y,x} = \sum_y f_i(y) p_{y \to x} + (1 - p(x)) f_i(x)$$

with the convention that $p_{y \to x} = 0$ if $y \to x$ is not an arrow of the active graph. Here f_{i+1} is again stochastic since $\sum_x f_{i+1}(x) = \sum_y f_i(y)$ so that the normalisation is preserved.

The other ingredient is that the lattice line $\mathbb{Z}$ can be viewed as a graph $\cdots \bullet_i \to \bullet_{i+1} \to \cdots$ and as such there is a 1-dimensional differential calculus $\Omega^1(\mathbb{Z})$ defined by the graph. As it happens, this is a Cayley graph associated to the additive generator 1, which means there is a basic left-invariant 1-form e_+ with bimodule relations $e_+ f = R_+(f)e_+$, where $R_+(f)_i = f_{i+1}$, and exterior derivative $\mathrm{d}f = (\partial_+ f)e_+$ with $(\partial_+ f)_i = f_{i+1} - f_i$ the usual 1-step discrete time derivative. This is not bidirected and not suitable for a $*$-calculus (we would be need Ω^1 to be 2-dimensional as in (Majid, 2019b)) but is more relevant at the moment.

Proposition 3.1. *A discrete Markov process with probability distribution $f_i(x)$ on X at each time i appears naturally as the quantum differential equation*

$$\partial_+ f = -\Delta_\theta f + (q - p)f,$$

where ∂_+ acts on i and on the right we use the graph calculus on X with (,), Δ_θ defined by the transition probabilities.

Proof. We have already done the work and it remains only to write (11) in terms of Δ_θ in (10). $\qquad\square$

Thus, a Markov process is nothing other than the noncommutative diffusion equation with a certain potential term which vanishes in the doubly stochastic case where $p = q$. An example of a right stochastic matrix is shown in Figure 3. The novel idea here is that probabilities play the role of 'quantum metric inner product' and that we obtain
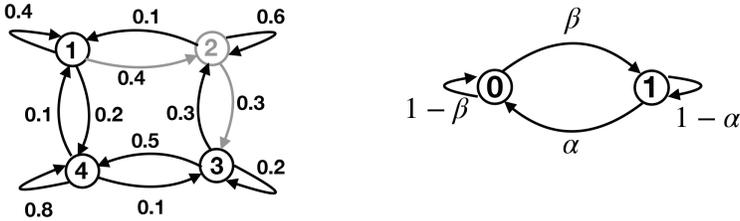
Figure 3: Markov process state diagrams on vertex set X with transition probabilities labelling arrows of an extended graph. Associated functions on the left are $p = (0.6, 0.4, 0.8, 0.2)$, $q = (0.2, 0.7, 0.4, 0.7)$ in vertex order as numbered.

a quantum geometric picture if we use the canonical Laplacian associated to this. The latter can coincide with the Laplacian for other more geometric connections (there are several examples in (Beggs and Majid, 2020) where that happens).

We can go further and extend the geometric meaning of a Markov process by tropicalisation, i.e., by writing

$$p_{x \to y} = e^{-\lambda_{x \to y}}, \quad p(x) = e^{-\lambda_{x \to x}}, \quad \lambda_{x \to y}, \lambda_{x \to x} \geq 0$$

and observe that the transition probability after n-steps is

$$P^n_{x,y} = \sum_{\gamma : x \to \cdots \to y} e^{-\lambda(\gamma)}; \quad \lambda(\gamma) = \sum_{i=0}^{n-1} \lambda_{x_i \to x_{i+1}}$$

where $\gamma = x_0 \to x_1 \cdots \to x_{n-1} \to x_n$ is an n-step path from $x_0 = x$ to $x_n = y$ (in the extended graph where we allow self-steps). If we think of $\lambda_{x \to y}$ as some kind of more geometric 'asymmetric length' associated to the arrow or self-arrow then $\lambda(\gamma)$ is the 'length' of the path γ. Thus we see how a generalised form of Riemannian geometry

emerges as an interpretation of a Markov process, or conversely how Markov processes emerge naturally from generalised Riemannian geometry, with a contribution of maximal probability equated to a path of shortest length. It is generalised in the sense that 'lengths' are direction dependent; there is no assumption that $\lambda_{x \to y} = \lambda_{y \to x}$ or even that one is nonzero when the other is, and we also allow self-arrow lengths $\lambda_{x \to x}$. Figure 3 shows the shortest path in this sense between 1 and 3 as via 2. In fact the generalised geometry here is equivalent to the notion of a Lawvere metric space (Fong and Spivak, 2019) on the vertex set, where $d(x, y)$ is the length of the shortest path. The main difference is that in that context one would have $\lambda_{x,x} = 0$ whereas in our case these are determined by the values on the arrows of the unextended graph. But in both cases, the self-arrow carries no new information and the actual data is the same. Moreover, this difference does not affect the shortest length provided these are all ≥ 0, which *is* a restriction on the unextended graph weights in our case. In other words, we are *not quite* freely labelling the arrows by lengths $\lambda_{x \to y}$ as there is a restriction

$$\sum_{y:x \to y} e^{-\lambda_{x \to y}} \leq 1$$

at every vertex. This has the character of some kind of lower bound on the 'distances' from every vertex except that it is not one minimum e.g. Planck length, but shared between all local directions. Our observations may also be related to path integration defined by stochastic processes, even though our starting point is a fresh one.

Finally, we consider if there is a Schrödinger version of a Markov process. We recall that in usual quantum mechanics, if ψ obeys $\dot{\psi} =$

$(\frac{\imath\hbar}{2m}\Delta + \frac{V}{\imath\hbar})\psi$ and normalisation $||\psi||_{L^2} = 1$ then the probability density $f = |\psi|^2$ obeys the conservation law

$$\dot{f} + \nabla \cdot J = 0, \quad J = -\frac{\imath\hbar}{2m}(\bar{\psi}\nabla\psi - \psi\nabla\bar{\psi}),$$

where J is the probability current. This implies that $\int \dot{f} = 0$ so that $\int f = 1$ holds at all times. Also note the formal similarity of the Schrödinger equation to an 'imaginary time' (Wick rotation) version of the diffusion one.

We first compute in general that if we are on an inner $*$-differential algebra with $V = V^* \in A$, $\theta^* = -\theta$ and $*(,) = (,)\mathrm{flip}(* \otimes *)$ and set $f = \psi^*\psi$ as dependent on an additional continuous time with

$$\dot{\psi} = \imath(-\Delta_\theta + V)\psi = \imath(\mathrm{d}\psi, \theta) + \imath V\psi = -\imath(\theta, \mathrm{d}\psi) + \imath[(\theta, \theta), \psi] + \imath V\psi$$

then

$$\dot{\psi}^* = \imath(\theta, \mathrm{d}\psi^*) - \imath\psi^* V$$

so that

$$\dot{f} = \dot{\psi}^*\psi + \psi^*\dot{\psi} = \imath(\theta, \mathrm{d}\psi^*)\psi - \imath\psi^*(\theta, \mathrm{d}\psi) + \imath\psi^*[(\theta, \theta), \psi].$$

If A is commutative then this reduces to

$$\dot{f} = -(,)(\theta \otimes J) = -\nabla_\theta \cdot J, \quad J = \imath((\mathrm{d}\psi)\psi^* - (\mathrm{d}\psi^*)\psi)$$

as the probability current 1-form, where $\nabla_\theta = \theta\otimes$ is the canonical connection. This remark only applies to continuous time, and moreover, we do not necessarily have $\int : A \to \mathbb{R}$ specified, and even if we did, we may not have $\int \nabla_\theta \cdot J = 0$ for all ψ. In case of an honest quantum metric and QLC, we might expect that $\int$ can be chosen depending on the metric so that integral of a divergence does vanish as in Riemannian geometry (it is not clear how generally we could do

this). However, in our asymmetric setting, this seems unlikely both mathematically and physically; we should not expect actual conserved probability, but rather we have a formal similarity as well as an element of irreversibility. A partial result at our graph theory level is the following.

Proposition 3.2. *Let $\psi_i \in \mathbb{C}(X)$ be complex valued, $V \in \mathbb{R}(X)$ real valued, $f_i = |\psi_i|^2$ and suppose that*

$$\partial_+\psi = \imath(-\Delta_\theta + V)\psi,$$

where ∂_+ acts on the discrete time index i. Then

$$\partial_+ f = V(-\Delta_\theta + V)f + V(\mathrm{d}\bar{\psi}, \mathrm{d}\psi) - \nabla_\theta \cdot J,$$

$$J = \imath((\mathrm{d}\psi)\bar{\psi} - (\mathrm{d}\bar{\psi})\psi) - \frac{1}{2}((\mathrm{d}\psi)\Delta_\theta\bar{\psi} + (\mathrm{d}\bar{\psi})\Delta_\theta\psi)$$

which we can also write as

$$\partial_+ f = V^2 f - \nabla_\theta \cdot J_V,$$

$$J_V = \imath((\mathrm{d}\psi)\bar{\psi} - (\mathrm{d}\bar{\psi})\psi) + (\mathrm{d}\psi)(-\frac{1}{2}\Delta_\theta + V)\bar{\psi} + (\mathrm{d}\bar{\psi})(-\frac{1}{2}\Delta_\theta + V)\psi.$$

Proof. Here $-\Delta_\theta\psi = (\mathrm{d}\psi, \theta)$ etc., so at the next time step,

$$\psi^{new} = \psi + \imath((\mathrm{d}\psi, \theta) + V\psi), \quad \bar{\psi}^{new} = \bar{\psi} - \imath((\mathrm{d}\bar{\psi}, \theta) + V\bar{\psi}),$$

$$
\begin{aligned}
f^{\mathrm{new}} &= (\bar{\psi} - \imath((\mathrm{d}\bar{\psi}, \theta) + V\bar{\psi}))(\psi + \imath((\mathrm{d}\psi, \theta) + V\psi)) \\
&= f + ((\mathrm{d}\bar{\psi}, \theta) + V\bar{\psi})((\mathrm{d}\psi, \theta) + V\psi) + \imath\bar{\psi}(\mathrm{d}\psi, \theta) - \imath(\mathrm{d}\bar{\psi}, \theta)\psi \\
&= (1 + V^2)f + (\mathrm{d}\bar{\psi}, \theta)(\mathrm{d}\psi, \theta) + (\mathrm{d}\psi, \theta)(\imath + V)\bar{\psi} \\
&\qquad + (\mathrm{d}\bar{\psi}, \theta)(-\imath + V)\psi
\end{aligned}
$$

with a cross term $V\bar{\psi}\psi$ cancelling. Next

$$(\mathrm{d}\bar{\psi}, \theta)\psi = (\mathrm{d}\bar{\psi}, \theta\psi) = (\mathrm{d}\bar{\psi}, \mathrm{d}\psi) + (\mathrm{d}\bar{\psi}, \psi\theta) = (\mathrm{d}\bar{\psi}, \mathrm{d}\psi) + ((\mathrm{d}\bar{\psi})\psi, \theta)$$

so this combined with $\bar{\psi}(\mathrm{d}\psi, \theta) = (\bar{\psi}\mathrm{d}\psi, \theta)$ contributes $V(\mathrm{d}f, \theta) + V(\mathrm{d}\bar{\psi}, \mathrm{d}\psi)$ to f^{new}. We also note that

$$-(\theta, (\mathrm{d}\psi)\bar{\psi} - (\mathrm{d}\bar{\psi})\psi) = -(\theta, \mathrm{d}\psi)\bar{\psi} + (\theta, \mathrm{d}\bar{\psi})\psi = (\mathrm{d}\psi, \theta)\bar{\psi} - (\mathrm{d}\bar{\psi}, \theta)\psi$$

which times $\imath$ verifies the first term of J. It remains to note that

$$-(\theta, (\mathrm{d}\bar{\psi})(\mathrm{d}\psi, \theta)) = -(\theta, (\mathrm{d}\psi)(\mathrm{d}\bar{\psi}, \theta)) = (\mathrm{d}\bar{\psi}, \theta)(\mathrm{d}\psi, \theta)$$

which times $1/2$ for each expression as in J, gives the remaining term needed for f^{new}. This gives our first expression for $\partial_+ f = f^{new} - f$. For the second expression, we note that

$$-(\theta, (\mathrm{d}\bar{\psi})V\psi + (\mathrm{d}\psi)V\bar{\psi})$$
$$= V\bar{\psi}(\mathrm{d}\psi, \theta) + V\psi(\mathrm{d}\bar{\psi}, \theta) = V(\mathrm{d}f, \theta) + V(\mathrm{d}\bar{\psi}, \mathrm{d}\psi)$$

as required for the extra terms in J_V to replace the corresponding terms in f^{new}. $\qquad\square$

Writing

$$J = \sum_{x \to y} J_{x \to y}\omega_{x \to y}, \quad \tilde{\psi}(y) = \sum_{y:x \to y} \psi(y)p_{y \to x},$$

explicit formulae for the various terms are

$$(\mathrm{d}\bar{\psi}, \mathrm{d}\psi)(x) = -\sum_{y:x \to y} f(y)p_{y \to x} - f(x)q(x)$$
$$+ \sum_{y:x \to y} (\psi(x)\bar{\psi}(y) + \bar{\psi}(x)\psi(y))p_{y \to x}$$

$$J_{x \to y} = \imath(\psi(y)\bar{\psi}(x) - \bar{\psi}(y)\psi(x)) + \frac{1}{2}\Big(\psi(y)\bar{\tilde{\psi}}(x) + \bar{\psi}(y)\tilde{\psi}(x)$$
$$-\psi(x)\bar{\tilde{\psi}}(x) - \bar{\psi}(x)\tilde{\psi}(x) - (\psi(y)\bar{\psi}(x) + \bar{\psi}(y)\psi(x))q(x)\Big) + f(x)q(x)$$

and the additional terms in $J_{V\,x\to y}$ are

$$V(x)(\psi(y)\bar{\psi}(x) + \bar{\psi}(y)\psi(x) - 2f(x)).$$

A direct calculation from f^{new} in the proof above also gives:

Proposition 3.3. *For the discrete Schrödinger process $\partial_+\psi = \imath(-\Delta_\theta + V)\psi$, we have*

$$\sum_X \partial_+ f = \sum_X \left((V-q)^2 f + (V-q)(\bar{\psi}\tilde{\psi} + \bar{\tilde{\psi}}\psi) + |\tilde{\psi}|^2 + \imath(\bar{\psi}\tilde{\psi} - \bar{\tilde{\psi}}\psi)\right)$$

for $f = |\psi|^2$.

   This is typically not zero, i.e. the discrete Schrödinger process with the ∇_θ connection does not leave the l^2-norm $\sum_x |\psi|^2$ constant, i.e. does not consist of unitary steps. This is visible even for $V = q$, when $\partial_+\psi = \imath\tilde{\psi}$, and is due in part to the 1-sided step $\partial_+ f$ not reflecting a $*$-calculus on $\mathbb{Z}$. It is also due to the connection ∇_θ being a convenient but not necessarily physical choice. Nevertheless, we see from the proposition that $\partial_+ f$ has a certain form which, when $V = 0$, is $\partial_+ f = -\nabla_\theta \cdot J$ as expected in quantum mechanics from a formal point of view, while for other V also contains a Markov-process like element.

3.1. Example of 2-state process

   More generally, we should consider the discrete Schrödinger process with other bimodule connections ∇ on Ω^1 and the associated Δ and $\nabla\cdot$ from (7). We will say that a connection ∇ is *unitary* if there exists a potential V such that $\partial_+\psi = \imath(-\Delta + V)\psi$ indeed has unitary

steps so that the total probability is conserved. In the remainder of this section, we explore this idea for the simplest example of a two state process.

Thus, let $X = \{0, 1\}$ with $\alpha = p_{1\to 0} > 0$ and $\beta = p_{0\to 1} > 0$ define a classical 2-state Markov process as shown on the right in Figure 3. In addition to this process interpreted as in Proposition 3.1 in terms of ∇_θ, we consider the same ideas as in Proposition 3.2 but for Laplacian and divergence given by a general bimodule connection ∇. It is known from (Majid, 2019a, Lemma 2.1) that these take the form

$$\nabla\theta = (1 - b)\theta \otimes \theta, \quad b(0) = s, \quad b(1) = t$$

for two complex parameters s, t. Here, $\theta = \omega_{0\to 1} + \omega_{1\to 0}$ is a basis over the algebra, so it is enough to specify ∇ on this. Its general value deduced from the Leibniz rule

$$\nabla(\psi\theta) = \mathrm{d}\psi \otimes \theta + \psi\nabla\theta; \quad \mathrm{d}\psi = (\psi(1) - \psi(0))(\omega_{0\to 1} - \omega_{1\to 0})$$

comes out as

(12)
$$\nabla\omega_{0\to 1} = \omega_{1\to 0} \otimes \omega_{0\to 1} - s\omega_{0\to 1} \otimes \omega_{1\to 0},$$
$$\nabla\omega_{1\to 0} = \omega_{0\to 1} \otimes \omega_{1\to 0} - t\omega_{1\to 0} \otimes \omega_{0\to 1}.$$

Meanwhile, the metric inner product is

$$(\omega_{0\to 1}, \omega_{1\to 0}) = \alpha\delta_0, \quad (\omega_{1\to 0}, \omega_{0\to 1}) = \beta\delta_1$$

and it is known also from (Majid, 2019a, Lemma 2.1) that ∇ is metric compatible (and hence a QLC for the standard exterior algebra) if and only if $\beta = \pm\alpha$, which in our context means $\beta = \alpha$, and $t = s^{-1}$. Moreover, in this case it is $*$-preserving if and only if s has modulus 1. So there is a circle of $*$-preserving QLCs, including the obvious $s = t = -1$ with $\nabla\theta = 2\theta \otimes \theta$. By contrast, the canonical one we studied above was $\nabla_\theta\theta = \theta \otimes \theta$ at $s = t = 0$.

We proceed with a general bimodule connection (12) for our discussion, remembering that it is a QLC only when $\beta = \alpha$ and $s = t^{-1}$. The metric inner product is

$$(\omega_{0\to1}, \omega_{1\to0}) = \alpha\delta_0, \quad (\omega_{1\to0}, \omega_{0\to1}) = \beta\delta_1$$

so that the Laplacian comes out as as

$$(\Delta\psi)(0) = -\delta_\psi\alpha(1+s), \quad (\Delta\psi)(1) = \delta_\psi\beta(1+t); \quad \delta_\psi := \psi(1) - \psi(0).$$

The discrete Schrödinger process $\partial_+\psi = \imath(-\Delta + V)\psi$ then corresponds to the matrix step on the column vector $\psi(0), \psi(1)$,

$$\psi^{new} = U\psi;$$

$$U = \begin{pmatrix} 1 + \imath V(0) - \imath\alpha(1+s) & \imath\alpha(1+s) \\ \imath\beta(1+t) & 1 + \imath V(1) - \imath\beta(1+t) \end{pmatrix},$$

which is unitary if and only if

$$\imath\alpha(1+s) = -e^{\imath\phi}\bar{z}, \quad |1 + \imath V(1) - z|^2 + |z|^2 = 1,$$
$$1 + \imath V(0) = -e^{\imath\phi}(2\bar{z} - 1 + \imath V(1)),$$

where $z = \imath\beta(1+t)$ and $e^{\imath\phi}$ is some phase. This has solution

(13)
$$V(0) = V(1) = 0, \quad z = \frac{1}{2}(1 - e^{\imath\phi}),$$
$$s = -1 - \frac{\imath}{2\alpha}(1 - e^{\imath\phi}), \quad t = -1 - \frac{\imath}{2\beta}(1 - e^{\imath\phi})$$

for a free angle parameter ϕ, with resulting Schrödinger process step

(14) $\quad U = \begin{pmatrix} 1 - z & z \\ z & 1 - z \end{pmatrix} = e^{\frac{\imath\phi}{2}} \begin{pmatrix} \cos(\frac{\phi}{2}) & -\imath\sin(\frac{\phi}{2}) \\ -\imath\sin(\frac{\phi}{2}) & \cos(\frac{\phi}{2}) \end{pmatrix}.$

Thus, the 2-point graph has a 1-parameter circle of bimodule connec-tions ∇ *which are 'unitary' in the sense that the step operator* U *is unitary.* Some examples are:

(i) $\phi = 0$ or $z = 0$ gives $s = t = -1$ in the QLC family and

$$\nabla \theta = 2\theta \otimes \theta, \quad U = \mathrm{id};$$

(ii) $\phi = \pi$ or $z = 1$ gives $s = -1 - \frac{\imath}{\alpha}$ and $t = -1 - \frac{\imath}{\beta}$ and

$$\nabla \theta = (2 + \frac{\imath}{\gamma})\theta \otimes \theta; \quad \gamma(0) = \alpha, \quad \gamma(1) = \beta; \quad U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

By contrast, our canonical $\nabla \theta = \theta \otimes \theta$ at $s = t = 0$ is not on this circle.

Proposition 3.4. *For the 1-parameter circle of unitary discrete Schrödinger evolutions (13)–(14), and writing* $f = |\psi|^2$ *in vector notation, we find*

$$f^{new} = \begin{pmatrix} \cos^2(\frac{\phi}{2}) & \sin^2(\frac{\phi}{2}) \\ \sin^2(\frac{\phi}{2}) & \cos^2(\frac{\phi}{2}) \end{pmatrix} f$$

$$+ \frac{\imath}{2} \sin(\phi) \left(\bar{\psi}(0)\psi(1) - \bar{\psi}(1)\psi(0) \right) \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

The first term is a general left and right stochastic Markov process on f*. In quantum geometric terms,*

$$\partial_+ \psi = -\imath \Delta \psi, \quad \partial_+ f = \frac{1}{2\imath}(1 - e^{-\imath\phi})\Delta f - \nabla \cdot J,$$

$$J = \frac{\imath}{2}(1 + e^{-\imath\phi}) \left((\mathrm{d}\psi)\bar{\psi} - (\mathrm{d}\bar{\psi})\psi \right).$$

Proof. Here

$$
\begin{aligned}
f^{new}(0) &= |\psi^{new}(0)|^2 = |(1-z)\psi(0) + z\psi(1)|^2 \\
&= |1-z|^2 f(0) + |z|^2 f(1) + (1-\bar{z})z\bar{\psi}(0)\psi(1) \\
&\quad + \bar{z}(1-z)\bar{\psi}(1)\psi(0) \\
&= \cos^2(\frac{\phi}{2})f(0) + \sin^2(\frac{\phi}{2})f(1) - \frac{\imath}{2}\sin(\phi)(\bar{\psi}(0)\psi(1) \\
&\quad - \bar{\psi}(1)\psi(0))
\end{aligned}
$$

and similarly for $f^{new}(1) = |\psi^{new}(1)|^2 = |(1-z)\psi(1) + z\psi(0)|^2$. Note that the entries of f remain positive and summing to 1 even though there is an extra divergence term.

We next consider a general 1-form $J = J_{01}\omega_{0\to 1} + J_{10}\omega_{1\to 0}$ for constants J_{01}, J_{10} and find from (12) and (13) that

$$
\begin{aligned}
\nabla \cdot J &= (\ ,\ )\big(J_{01}(\omega_{1\to 0} \otimes \omega_{0\to 1} - s\omega_{0\to 1} \otimes \omega_{1\to 0}) \\
&\quad + J_{10}(\omega_{0\to 1} \otimes \omega_{1\to 0} - t\omega_{1\to 0} \otimes \omega_{0\to 1})\big) \\
&= (J_{10} - sJ_{01})\alpha\delta_0 + (J_{01} - tJ_{10})\beta\delta_1 \\
&= (J_{01} + J_{10})\gamma + \frac{\imath}{2}(1 - e^{\imath\phi})j
\end{aligned}
$$

where $\gamma(0) = \alpha, \gamma(1) = \beta$ and $j(0) = J_{01}, j(1) = J_{10}$ are functions on X. If we set $J_{01} = -J_{10} = \frac{\imath}{2}(1 + e^{-\imath\phi})(\bar{\psi}(0)\psi(1) - \bar{\psi}(1)\psi(0))$ then $-\nabla \cdot J$ gives the second term of f^{new} as required. This gives J as stated after we observe that

$$
\begin{aligned}
(\mathrm{d}\psi)\bar{\psi} - (\mathrm{d}\bar{\psi})\psi &= (\psi(1) - \psi(0))(\omega_{0\to 1} - \omega_{1\to 0})\bar{\psi} \\
&\quad - (\bar{\psi}(1) - \bar{\psi}(0))(\omega_{0\to 1} - \omega_{1\to 0})\psi \\
&= (\psi(1) - \psi(0))\bar{\psi}(1)\omega_{0\to 1} - (\psi(1) - \psi(0))\bar{\psi}(0)\omega_{1\to 0} \\
&\quad - (\bar{\psi}(1) - \bar{\psi}(0))\psi(1)\omega_{0\to 1} + (\bar{\psi}(1) - \bar{\psi}(0))\psi(0)\omega_{1\to 0} \\
&= (\bar{\psi}(0)\psi(1) - \bar{\psi}(1)\psi(0))(\omega_{0\to 1} - \omega_{1\to 0}).
\end{aligned}
$$

Moreover, $\mathrm{d}f = (f(1) - f(0))(\omega_{0\to1} - \omega_{1\to0})$. Hence, using our computation of $\nabla\cdot$,

$$\Delta f = \nabla \cdot \mathrm{d}f = \frac{\imath}{2}(1 - e^{\imath\phi})(f(1) - f(0))\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$= \frac{\imath}{2}(1 - e^{\imath\phi})\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} f$$

in vector notation for f (which, when applied to ψ, recovers the Schrödinger process step $U = 1 - \imath\Delta$ in (14) as expected). Moreover, the stated first term of $\partial_+ f$ is then

$$\frac{1}{2\imath}(1 - e^{-\imath\phi})\frac{\imath}{2}(1 - e^{\imath\phi})\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} f = \sin^2(\frac{\phi}{2})\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} f$$

as required in $f^{new} - f$. $\qquad\square$

For the l^2-norm, we took the constant measure in summing over X. One could also, in principle, introduce a measure related to the quantum metric $(\ ,\ )$. This would be more in keeping with Riemannian geometry but is not so natural from our point of view on Markov processes (where the usual constant measure is preserved).

4. Digital geometry of 2×2 matrices

Quantum Riemannian geometry and quantum gravity on one of the simplest graphs, a quadrilateral, was recently achieved (Majid, 2019a) (with Lorentzian style negative square-length weights on two of the sides). Here we briefly look at the complementary example of a

noncommutative finite geometry, namely the humble algebra of 2×2 matrices $M_2(\mathbb{C})$. Its quantum Riemannian geometry turns out to be rather rich and is not fully explored.

We take the standard 2D $*$-differential calculus from (Beggs and Majid, 2017; 2020) with a basis of central 1-forms $s, t \in \Omega^1(M_2)$, differential and exterior algebra

$$\mathrm{d}a = [E_{12}, a]s + [E_{21}, a]t,$$

$$s^2 = t^2 = 0, \quad s \wedge t = t \wedge s, \quad \mathrm{d}s = 2\theta \wedge s, \quad \mathrm{d}t = 2\theta \wedge t, \quad s^* = -t,$$

which is inner with $\theta = E_{12}s + E_{21}t$. Here E_{ij} is the matrix with 1 in the (i, j) place and 0 elsewhere.

Next, as the basis is central and the metric has to be central to be invertible in the bimodule sense, any invertible 2×2 complex matrix g_{ij} in our basis can be taken as metric coefficients, with the condition that $g_{12} = -g_{21}$ if we want to impose quantum symmetry, and $g_{22} = \overline{g_{11}}$, g_{12} real if we want g to be 'real' in the required hermitian sense. It is easy to see that

(15)　　　$\nabla s = 2\theta \otimes s, \quad \nabla t = 2\theta \otimes t, \quad \sigma = -\text{flip}, \quad R_{\nabla} = 0$

on the generators is a flat QLC simultaneously for all quantum metrics. But there are typically many more QLCs. The actual moduli of these has only been computed in (Beggs and Majid, 2020, Example 8.13 and Exercise 8.3) for a couple of sample quantum metrics g_1, g_2, with results there as follows.

(i) $g_1 = s \otimes t - t \otimes s$. This has a principal 4-parameter moduli of QLCs of the form

$$
\begin{aligned}
\nabla s =& 2\theta \otimes s - \begin{pmatrix} 0 & \mu\alpha \\ \beta & 0 \end{pmatrix} s \otimes s - \begin{pmatrix} 0 & \alpha \\ \nu\beta & 0 \end{pmatrix} g_1 \\
&+ \begin{pmatrix} 0 & \nu\alpha \\ \nu^2\beta + (\mu\nu - 1)\alpha & 0 \end{pmatrix} t \otimes t,
\end{aligned}
$$

$$
\begin{aligned}
\nabla t =& 2\theta \otimes t + \begin{pmatrix} 0 & \mu^2\alpha + (\mu\nu - 1)\beta \\ \mu\beta & 0 \end{pmatrix} s \otimes s \\
&+ \begin{pmatrix} 0 & \mu\alpha \\ \beta & 0 \end{pmatrix} g_1 - \begin{pmatrix} 0 & \alpha \\ \nu\beta & 0 \end{pmatrix} t \otimes t.
\end{aligned}
$$

The $\alpha = \beta = \mu = \nu = 0$ point is the flat one (15). We rescaled the μ, ν compared to (Beggs and Majid, 2020) here and in the next case.

(ii) $g_2 = s \otimes s + t \otimes t$. This has a principal 3-parameter moduli of QLCs of the form

$$
\begin{aligned}
\nabla s =& 2E_{21} t \otimes s + \begin{pmatrix} 0 & \mu\rho \\ 2\mu - \rho(1 + \mu(\mu + \nu)) & 0 \end{pmatrix} s \otimes s \\
&+ \begin{pmatrix} 0 & -\rho \\ \mu\rho & 0 \end{pmatrix} g_1 + \begin{pmatrix} 0 & \nu\rho \\ \rho & 0 \end{pmatrix} t \otimes t,
\end{aligned}
$$

$$
\begin{aligned}
\nabla t =& 2E_{12} s \otimes t + \begin{pmatrix} 0 & -\rho \\ \mu\rho & 0 \end{pmatrix} s \otimes s - \begin{pmatrix} 0 & \nu\rho \\ \rho & 0 \end{pmatrix} g_1 \\
&+ \begin{pmatrix} 0 & -2\nu + \rho(1 + \nu(\mu + \nu)) \\ \nu\rho & 0 \end{pmatrix} t \otimes t.
\end{aligned}
$$

There are restrictions on the parameters over $\mathbb{C}$ for a $*$-preserving connection. The $\mu = \nu = \rho = 0$ point has curvature

$$R_\nabla s = 2s \wedge t \otimes s, \quad R_\nabla t = 2s \wedge t \otimes t.$$

We also have an obvious 'symmetric lift' $i(s \wedge t) = \frac{1}{2}(s \otimes t + t \otimes s)$ which at the zero parameter point yields

$$\mathrm{Ricci} = s \otimes t + t \otimes s, \quad S = 0,$$

but note that i in the $*$-algebra case is not antihermitian in the required sense so that this Ricci is not hermitian.

The above are principal moduli with nontrivial σ and zero for the bimodule map α in the general construction; there is also a 4 dimensional moduli of QLCs for g_2 with $\sigma = -\mathrm{flip}$. In general, the space of $*$-preserving 'real' pairs (g, ∇) appears to be generically 7 dimensional, although this remains to be determined. Once known, one could define quantum gravity on $M_2(\mathbb{C})$ by integration over this moduli, schematically,

$$\langle \mathcal{O}_1 \mathcal{O}_2 \rangle = \frac{\int \mathcal{D}(g, \nabla) e^{\imath \mathrm{Tr} S[g, \nabla]} \mathcal{O}_1 \mathcal{O}_2}{\int \mathcal{D}(g, \nabla) e^{\imath \mathrm{Tr} S[g, \nabla]}}$$

for the given differential structure and lift i fixed. Here $S[g, \nabla] \in M_2(\mathbb{C})$ is the Ricci scalar, which generically is expected not to vanish. One could also consider summing or integrating over the choice of differential structure up to diffeomorphism.

It is natural to ask how this approach to quantum gravity compares to other, more established, approaches. A direct comparison is not possible due to very different methodologies, but taking the 'coordinate algebra' finite dimensional as we did above is broadly in the spirit of replacing spacetime by a finite geometry, such as a finite lattice or triangulation (Ambjørn, Jurkiewicz and Loll, 2001). A main

difference in our case is that we approach the matter in a coherent way that applies, as we saw in Section 2, to general algebras including finite-dimensional ones at one end of the range and classical $C^\infty(M)$ for a manifold M at the other, without too many ad-hoc steps. A consequence is that one can in principle take a limiting process from one to the other without ever leaving the category of models. Thus, the polygon lattice $\mathbb{Z}_n$ in (Argota-Quiroz and Majid, 2020) with its 2D calculus can be seen in the limit $n \to \infty$ to tend to S^1 with a certain 2D calculus. This turns out to be a central extension of the classical calculus on a circle in the sense of (Majid, 2020b; Beggs and Majid, 2020, Chapter 8.3), where the 'partial-derivative' associated to the extra dimension is a 2$^\text{nd}$ order Laplacian-like operator giving something like an Ito derivative in stochastic calculus. Here, the polygon graph is viewed as bidirected $i \leftrightarrow i + 1$ at vertices numbered in sequence mod n, and this is the reason for the 2D calculus. As we have seen in Section 3, we can also work with 1-way arrows $x \to y$, and this could make contact with causal set models (Dowker, 2013) where $x \to y$ indicates a causal path between spacetime points. So far in quantum Riemannian geometry models, the Lorentzian case has been handled by having 'square-length' weights < 0 on designated timelike edges (Majid, 2019a). In such models, one could think of the causal structure as a choice of orientation of the subgraph of timelike edges (choosing an arrow direction for each such edge).

Other approaches such as loop quantum cosmology (Ashtekar et al., 2007) also drastically cut the degrees of freedom to the point of computability. From such models as well as from experience with 2+1 quantum gravity, it does appear that some quantum gravity effects could be modelled by a noncommutative spacetime, which is support for our underlying framework. What we would hope for here is a kind of 'eigengeometry' where quantum gravity effects render spacetime

noncommutative as a better approximation than classical, then quantisation of gravity on that background results in effects which justify the supposed quantum spacetime as the effective background. This is a circular self-consistency condition, the content of which has yet to be understood. There have also been attempts at discrete quantum gravity in Connes' spectral triple approach to noncommutative geometry, where one integrates over possible Dirac operators (Hale, 2002) on $\mathbb{Z}_2$. Although very different, the results are not unlike the those of (Majid, 2019a; Argota-Quiroz and Majid, 2020) in that there is a constant relative uncertainty. This work also looks at integration over spectral triples on $M_2(\mathbb{C})$.

Next, what about the landscape of *all* finite quantum Riemannian geometries of small algebra dimension? This is a tough classification programme and so far has been achieved (Majid and Pachoł, 2020a) only up to dimension 3 and by specialising to the digital case over the field $\mathbb{F}_2 = \{0, 1\}$ to simplify the problem. Here algebra dimension 2 has no interesting quantum geometries, while for dimension 3 there are 7 algebras labelled A–G but only B,D,F admit regular 2D quantum Riemannian geometries, as summarised in the table in Figure 4. Here B is the Boolean algebra of subsets of a set of three elements with calculus and metric corresponding to an equilateral triangle graph, forming the group $\mathbb{Z}_3$. It is a Hopf algebra and its dual is the group algebra D$= \mathbb{F}_2\mathbb{Z}_3$. Its three quantum metrics are related by an overall element of the algebra (i.e., 'conformally equivalent' in some sense). In each case, the 4 QLCs with 1 flat and 3 nonflat is the opposite of what was found for B, possibly consistent with Hopf algebra duality interchanging high and low curvature (cf. Majid, 1988; 1991). For $\mathbb{F}_8$, the 7 metrics are again related by invertible factors but fall into three groups behaving differently as shown in the table. Details of the connections and curvatures are given in (Majid and Pachoł, 2020a).

For interesting examples with the algebra noncommutative, we therefore need to go to dimension $n = 4$, which was beyond the

3D algebras	metrics	QLCs	Nonflat
$B = \mathbb{F}_2(\mathbb{Z}_3)$	1	4	1
$D = \mathbb{F}_2\mathbb{Z}_3$	3	4	3
		4	3
		4	3
$F = \mathbb{F}_8$	3	-	-
	3	12	7
		12	7
		12	7
	1	4	1
A, C, E, G	-	-	-

Figure 4: Number of quantum geometries over $\mathbb{F}_2$ of algebra dimension 3 and 2D parallelisable calculus Ω. Data from (Majid and Pachoł, 2020a).

computer power available when writing (Majid and Pachoł, 2020a) (where QLCs were found by trying some 2^{24} possible connection values). As a result, the landscape of all quantum geometries for $n = 4$ is largely unexplored, though some flat connections are known for the algebra A_2 in the family of Hopf algebras over $\mathbb{F}_2$ in (Bassett and Majid, 2020). In dimension 4, there are 9 distinct noncommutative algebras (Majid and Pachoł, 2020b) (and 16 commutative ones) over $\mathbb{F}_2$, and one of the former is of course $M_2(\mathbb{F}_2)$.

Here we note that by reducing the above generic solutions for $M_2(\mathbb{C})$ to the $\mathbb{F}_2$ case, we can obtain some, though not necessarily all, of the quantum geometries on $M_2(\mathbb{F}_2)$, at least for the two metrics stated above. Since 2=0 in $\mathbb{F}_2$, we now have $\mathrm{d}s = \mathrm{d}t = 0$, and indeed $t = \mathrm{d}E_{12}$, $s = \mathrm{d}E_{21}$ are exact. Setting the parameters for our generic solutions to all values 0,1 gives us up to 16 and 8 QLCs respectively for our standard metrics, which appear now as

$$g_1 = s \otimes t + t \otimes s, \quad g_2 = s \otimes s + t \otimes t.$$

We broadly identify four cases:

(a) Setting $\alpha = \beta = \rho = 0$ and any μ, ν gives the zero flat QLC $\nabla s = \nabla t = 0$ for both metrics.

(b) Setting $\alpha = \beta = \rho = 1$ and $\mu = \nu = 1$ gives another flat QLC

$$\nabla s = \nabla t = \sigma_1(g_1 + g_2), \quad R_\nabla = 0$$

for both metrics, on noting that $\nabla(s + t) = 0$. Here $\sigma_1 = E_{12} + E_{21}$.

(c) Setting $\alpha = \beta = \rho = 1$ and $\mu = \nu = 0$ gives another flat QLC

$$\nabla s = E_{12}g_1 + E_{21}g_2, \quad \nabla t = E_{21}g_1 + E_{12}g_2, \quad R_\nabla = 0$$

for both metrics, on noting that $\mathrm{d}(E_{12}t) = \mathrm{d}(E_{21}s) = 0$ and $\mathrm{d}(E_{12}s) = \mathrm{d}(E_{21}t) = s \wedge t$.

α	β	μ	ν	∇s	∇t	$R_\nabla s$	$R_\nabla t$
1	1	0	1	$E_{21}s \otimes s + \sigma_1 g_1 + E_{12}t \otimes t$	$E_{12}s \otimes s + E_{21}g_1 + \sigma_1 t \otimes t$	$V \otimes E_{11}s$	$V \otimes (s + E_{22}t)$
1	1	1	0	$\sigma_1 s \otimes s + E_{12}g_1 + E_{21}t \otimes t$		$V \otimes (E_{11}s + t)$	$V \otimes E_{22}t$
$\rho = 1$		0	1	$E_{21}s \otimes s + E_{12}g_1 + \sigma_1 t \otimes t$	$E_{12}s \otimes s + \sigma_1 g_1 + E_{21}t \otimes t$	flat	
$\rho = 1$		1	0		$\sigma_1 s \otimes s + E_{21}g_1 + E_{12}t \otimes t$	flat	
0	1	0	0	$E_{21}s \otimes s$	$E_{12}s \otimes s + E_{21}g_1$	flat	
0	1	0	1		$\sigma_1 s \otimes s + E_{21}(g_1 + g_2)$	$V \otimes (E_{11}s + t)$	$V \otimes (s + E_{22}t)$
0	1	1	0		$\sigma_1 s \otimes s + E_{21}g_1$	flat	
0	1	1	1	$E_{21}(g_1 + g_2)$		$V \otimes (s + t)$	
1	0	0	0	$E_{21}t \otimes t + E_{12}g_1$	$E_{12}t \otimes t$	flat	
1	0	0	1	$E_{12}g_1 + \sigma_1 t \otimes t$		flat	
1	0	1	0	$E_{12}(g_1 + g_2) + \sigma_1 t \otimes t$		$V \otimes (E_{11}s + t)$	$V \otimes (s + E_{22}t)$
1	0	1	1	$E_{12}(g_1 + g_2)$		$V \otimes (s + t)$	

Figure 5: Nonzero QLCs on $M_2(\mathbb{F}_2)$ arising as limits of connections over $\mathbb{C}$ (which may not be all of them). All are for g_1 except the $\rho = 1$ rows for g_2, which are flat. We see 6 with curvature, where $V = s \wedge t$ is the 'volume form'.

(d) For curvature, we therefore need to look to $\alpha \neq \beta$ or $\mu \neq \nu$. The nonzero ones are shown in the table Figure 5, where we see that half of them have Riemann curvature, all QLCs for the metric g_1.

Next, the usual lift i needed for the Ricci curvature, and the Einstein tensor, both present a problem over $\mathbb{F}_2$ in that there is no $1/2$ at our disposal. Our approach in (Majid and Pachoł, 2020a) is to look at all possible i and tentatively to define $\mathrm{Eins} = \mathrm{Ricci} + gS$. Now we explore a different idea which is not systematic but applies when the exterior algebra has a suitable form, namely to look for two halves of the classical 'antisymmetric' lifts' separately, without the factor $1/2$. For the $M_2(\mathbb{F}_2)$ calculus as above, obvious choices would be

$$i_+(s \wedge t) = s \otimes t, \quad i_-(s \wedge t) = t \otimes s.$$

As a result, there are two natural Ricci tensors $\mathrm{Ricci}_\pm$ as defined with $i_\pm$. If $S_+ = S_- = S$, i.e., if the two Ricci scalars are the same then we propose the 'twice-Ricci' and 'twice-Einstein' tensors

$$_2\mathrm{Ricci} = \mathrm{Ricci}_+ + \mathrm{Ricci}_-, \quad _2\mathrm{Eins} = {}_2\mathrm{Ricci} + gS.$$

For the 3-dimensional Boolean algebra B, this gives $_2\mathrm{Ricci} = g$, $S = 1$ and $_2\mathrm{Eins} = 0$ but for the D algebra one has $S_+ \neq S_-$ for the obvious lifts (which is not to say the new approach could not work, for suitable $i_\pm$.)

Proposition 4.1. *On $M_2(\mathbb{F}_2)$ with the metric g_1 as above and its two nonflat joint QLCs in Figure 5 with curvature $R_\nabla s = R_\nabla t = s \wedge t \otimes (s + t)$, we have*

$$\mathrm{Ricci}_+ = t \otimes (s + t), \quad \mathrm{Ricci}_- = s \otimes (s + t), \quad S_\pm = 1;$$
$$_2\mathrm{Ricci} = g_1 + g_2, \quad _2\mathrm{Eins} = g_2.$$

Moreover, ∇ here are also QLCs for g_2 and hence $\nabla(_2\mathrm{Ricci}) = \nabla(_2\mathrm{Eins}) = 0$. The three other types of curvature for QLCs of g_1 in Figure 5 have $S_+ \neq S_-$.

Proof. We use the curvature as shown and take the 'trace' with respect to g_1 as this is the relevant metric. Since the output of R_∇ is the same on s, t,

$$\mathrm{Ricci}_+ = ((s+t, \ )\otimes \mathrm{id})(s\otimes t\otimes(s+t)) = t\otimes(s+t), \quad S_+ = 1,$$

$$\mathrm{Ricci}_- = ((s+t, \ )\otimes \mathrm{id})(t\otimes s\otimes(s+t)) = s\otimes(s+t), \quad S_- = 1.$$

In this case, we also have $_2\mathrm{Ricci} = (s+t)\otimes(s+t) = g_1 + g_2$ on adding these. Hence adding Sg_1 gives us the other metric. We next look more closely at the two connections in the table with this curvature. The one on the bottom line at $\alpha = \mu = \nu = 1$ and $\beta = 0$ has braiding

$$\sigma(\left\{ \begin{matrix} s \\ t \end{matrix}\right. \otimes s) = s\otimes \left\{ \begin{matrix} s \\ t \end{matrix}\right. + g_1 + g_2, \quad \sigma(\left\{ \begin{matrix} s \\ t \end{matrix}\right. \otimes t) = t\otimes \left\{ \begin{matrix} s \\ t \end{matrix}\right. .$$

Then since ∇ has the same value on s, t,

$$\begin{aligned}
\nabla g_2 &= \nabla s\otimes s + \nabla t\otimes t + (\sigma\otimes \mathrm{id})(s\otimes\nabla s + t\otimes\nabla t) \\
&= E_{12}(g_1 + g_2)\otimes(s+t) + E_{12}(\sigma\otimes \mathrm{id})((s+t)\otimes(g_1 + g_2)) \\
&= E_{12}(s+t)^{\otimes 3} + E_{12}(s+t)^{\otimes 3} \\
&\quad + ((\sigma - \mathrm{flip})\otimes \mathrm{id})((s+t)\otimes s\otimes(s+t)) = 0,
\end{aligned}$$

where we apply σ as flip for all cases, plus the additional $\sigma - \mathrm{flip}$ when acting on $(s+t)\otimes s$ in the second term, which contributes zero.

The computations for the three other types of curvature of QLCs for g_1 are similar. For example, for $R_\nabla s = s \wedge t \otimes (E_{11} s + t)$, $R_\nabla t = s \wedge t \otimes (s + E_{22} t)$, we have

$$\mathrm{Ricci}_+ = t \otimes (t + E_{11} s), \quad S_+ = E_{11},$$
$$\mathrm{Ricci}_- = s \otimes (s + E_{22} t), \quad S_- = E_{22}$$

so that $S_+ \neq S_-$. Interestingly, $S_+ + S_- = 1$. These same values of $S_\pm$ are also obtained for the other two cases. Hence for these three curvature types, our definition of $_2\mathrm{Eins}$ does not apply, although for the case shown, we do have $_2\mathrm{Ricci} + S_+ t \otimes s + S_- s \otimes t = g_2$ again. One could still search for other more suitable $i_\pm$ (similarly to the algebra D for $n = 3$) and meanwhile, in all cases, we can still use the tentative proposal in (Majid and Pachoł, 2020a) to define $\mathrm{Eins}_\pm = \mathrm{Ricci}_\pm + S_\pm g_1$. $\qquad\qquad\qquad\square$

The Einstein tensor vanishes automatically for a classical 2-manifold, but this need not be the case in quantum geometry. We see on this sample of connections on $M_2(\mathbb{F}_2)$ that $_2\mathrm{Eins}$ is conserved when it applies but that the general picture for a suitable Einstein tensor remains inconclusive. Stepping away from quantum Riemannian geometry towards other applications relevant to computer science, (Majid and Pachoł, 2020b) classified all quantum groups to dimension 4 over $\mathbb{F}_2$.

5. Beyond de Morgan duality

Quantum geometry also provides a geometric view of de Morgan duality (Majid, 2020a), extending the well-known feature of Boolean algebras that says that the negation of a Boolean expression has the

same form with all elements negated, $\emptyset$ and everything swapped, and $\cup, \cap$ swapped. In propositional logic, this sends $a \Rightarrow b$ to the equivalent statement $\bar{b} \Rightarrow \bar{a}$, while in terms of the power set $P(X)$ of subsets of a set X, the duality sends $a \subseteq X$ to its complement $\bar{a}$ in X.

This was the topic of my conference talk and I refer to (Majid, 2020a) for details. In the present notes, I instead want to recall the philosophical context and discuss what might come next. Therefore, suffice it to say that one can view $P(X)$ as an algebra over $\mathbb{F}_2$ with product $\cap$ and addition $\oplus$ (the 'exclusive or' $a \oplus b = (a \cup b) \cap \overline{a \cap b}$ operation). Next, we fix a directed graph on X as vertex set and let Arr be the set of arrows. Combining the graph calculus as in Section 3 with digital methods as in Section 4, we set $\Omega^1(P(X)) = P(\text{Arr})$ with addition given by 'exclusive-or' of subsets of arrows and the noncommutative bimodule structure (Majid, 2020a)

$$a \cap \omega := \{\text{arrows in } \omega \text{ with tail in } a\},$$

$$\omega \cap a := \{\text{arrows in } \omega \text{ with tip in } a\},$$

$$\mathrm{d}a = \{\text{arrows with one end in } a \text{ and other end in } \bar{a}\},$$

where $a \subseteq X$ and $\bar{a}$ is its complement. One can go on and define the maximal prolongation exterior algebra $\Omega_{max}(P(X))$ as well as two natural quotients.

We also define the dual algebra structure $\bar{P}(X)$ as the same set $P(X)$ but with addition given by 'inclusive and' or 'not-exclusive-or' $a \bar{\oplus} b = (a \cap b) \cup \overline{a \cup b}$ and product $a \cup b$. We define $\Omega^1(\bar{P}(X)) = \bar{P}(\text{Arr})$ with its 'inclusive and' as addition and (Majid, 2020a)

$$a \cup \omega = \{\text{arrows in } \omega \text{ or with tail in } a\},$$

$$\omega \cup a = \{\text{arrows in } \omega \text{ or with tip in } a\},$$

$$\bar{\mathrm{d}}a = \{\text{arrows wholly in } a \text{ or wholly in } \bar{a}\} = \overline{\mathrm{d}a}.$$

This too extends to $\Omega_{max}(\bar{P}(X))$ and its two natural quotients.

Theorem 5.1. *(Majid, 2020a) The algebra map* $^-: P(X) \to \bar{P}(X)$ *is a diffeomorphism, i.e. extends to a map of the corresponding differential exterior algebras.*

That $^-$ is an algebra isomorphism is the usual de Morgan duality of Boolean algebra in our algebraic language of digital geometry, and the claim is that this extends in the right manner to arrows or differentials of the graph calculus. The extension to 1-forms is just complementation of subsets of arrows. The work (Majid, 2020a) also shows how these ideas can be extended to any unital algebra A over $\mathbb{F}_2$ with $\bar{A}$ isomorphic via $\bar{a} = 1 + a$, likewise becoming a diffeomorphism for suitable differential structures.

This is a purely mathematical result but its philosophical motivation is as follows. Indeed, some 30 years ago in (Majid, 1991), I posed the question that if Boolean algebras are the simplest 'theory of physics' then what becomes of de Morgan duality in more advanced theories? I argued that while clearly broken by quantum theory and gravity alone (for example, apples curve space but the presence of not-apples, meaning the absence of apples, does not) such a duality but might re-emerge as a symmetry of quantum gravity. This was and still is meant to be thought-provoking speculation rather than something understood, but the idea was that *we* might say that a region of space is 'as full of apples' as GR allows (forming a black hole and expanding if we put more apples in) while someone else using the dual picture might say that this same region of space was as empty of not-apples as their quantum field theory allows (where in QFT, space is never completely empty in some sense due to vacuum fluctuations). I don't know if this vision is achievable, but what we can do is move quantum gravity down to the digital level of geometry on Boolean algebras and see if de Morgan duality indeed holds there. This is what

is done in (Majid, 2020a) at the Boolean level in so far as GR is a theory of metrics and connections on Ω^1; all of that works in a dual version. We also needed an element of 'quantum' in that our extension of $\cap, \cup$ to arrows was noncommutative, then Theorem 5.1 says that de Morgan duality indeed holds as part of general covariance in some extended sense.

This sense is admittedly a little weird. In usual GR, a diffeomorphism is induced by an underlying set map, but complementation is different and operates directly on subsets of X. For example, if $S \in P(X)$ is the Ricci scalar for a connection then for the same geometry it appears as $\bar{S}$ on the other side of de Morgan duality. So in a region where the curvature characteristic function has value 1 for us, it has value 0 for them. Similarly, the inner generator $\theta \in \Omega^1(P(X))$ is the sum of all arrows, or in some sense the 'maximal density' differential form. Its role on the de Morgan dual side is played by the zero element which is the inner generator for $\Omega^1(\bar{P}(X))$ but from the first point of view is literally the zero differential form. All of this is suggestive of the black-hole/vacuum discussion above but is probably the most we can say at the Boolean level (where it is hard to think about actual quantum theory).

On the other hand, more advanced theories of physics could still have the duality we seek in an increasingly visible form. Thus, speculatively, just as Schrödinger's cat is in a mixed state that is *neither dead or alive*, I have proposed (Majid, 2015) *co-Schrödinger's cat* as a cat falling into a black hole. This is *both dead* in finite proper time *and alive* forever in the frame of the observer at infinity. In other words, while quantum theory is intuitionistic as in a Heyting algebra, where we relax the rule that $a \cup \bar{a} =$ everything, gravity might be expected to be cointuitionistic in character in the de Morgan dual sense, as in a coHeyting algebra, where we relax the rule that $a \cap \bar{a} = \emptyset$.

The latter has also been proposed for other reasons in (Lawvere, 1991) as geometric in nature with $\partial a = a \cap \bar{a}$ a kind of boundary of a. This then requires both effects or quantum-and-gravity for the symmetry to be maintained. This suggests the next level of our duality programme:

> Can we extend quantum differential geometry to Heyting and coHeyting algebras and thereby extend the negation duality to a diffeomorphism?

The problem is that a Heyting algebra has both a 'meet' product $\cap$ and a 'join' $\cup$ but the latter does not in general provide an addition rule and does not play well with negation in order to be able to define a more suitable $\oplus$ over which $\cap$ distributes, i.e. we do not in general have the basics for an actual algebra in the sense of a ring (before even considering a field). I do not know the answer, but I would like to suggest an approach. In fact, both Boolean algebras $P(X)$ and Heyting algebras (and a lot more) are examples of preorders $(P, \leq)$ by which we mean an extended directed graph (adding in all self-arrows) which is closed under composition of arrows (here $a \leq b$ corresponds to an arrow $a \to b$ and in the Boolean case is just subset inclusion). Both cases actually have rather more structure, namely min and max elements $0, 1$ (the empty set and the whole set in the Boolean case), unital tensor product $\otimes$ (the $\cap$) and an 'internal hom' $\underline{\text{hom}}$ leading to a 'dual' $\bar{a} = \underline{\text{hom}}(a, 0)$ (the negation) and making the preorder closed symmetric monoidal in the sense of (Fong and Spivak, 2019) (this is a baby closed symmetric monoidal category in the sense that there is at most one morphism between any two objects). Abstractly, $\underline{\text{hom}}$ is a binary operation characterised by $a \otimes b \leq c$ if and only if $a \leq \underline{\text{hom}}(b, c)$, which you can check is true in the Boolean case with $\underline{\text{hom}}(b, c) = \bar{b} \cup c$ (and its existence is a definition in the Heyting case). In this sense, de Morgan duality has a similar conceptual flavour to

vector space duality in the tensor category of vector spaces, which is in the right direction towards connecting it with observable-state or Hopf algebra duality (Majid, 2014).

Next, instead of working directly with the Boolean or Heyting algebra P as our 'algebra of functions', we work now with functionals, i.e. the algebra $A = \mathbb{C}(P)$ of functions on P. As P being a preorder has a directed graph structure, we have

$$\mathrm{d}\Phi = \sum_{a \leq b} (\Phi(b) - \Phi(a))\omega_{a \leq b}$$

for all functionals $\Phi \in \mathbb{C}(P)$ as a graph calculus on P. The Boolean or Heyting negation as a set map $P \to P$ induces an algebra map on $\mathbb{C}(P)$ and this again is differentiable. We do not directly see the 'manifold' structure on X in the Boolean case where $P = P(X)$ and X itself has a directed graph with arrow set Arr. However, we can similarly look at functionals of the differentials on X understood now as functions on $P(\mathrm{Arr})$. We therefore have the ingredients for some form of variational calculus for functionals of functions on X and their derivatives, in which the graph differentials on X enter, much as quantum field theory on the space of functions on a manifold still captures information about the manifold. When P is now some Heyting algebra or indeed any preorder, we can look for similar structures at the level of $\mathbb{C}(P)$ with a calculus on P handled implicitly. Given the importance of Heyting algebras and topos theory in physics (see e.g. Döring and Isham, 2011) it should be useful to explore their quantum Riemannian geometry, extending what should be achievable in the Boolean case, even if we have to handle it implicitly in terms of functionals. What may emerge is something like working over an algebra but in terms of $\cup$ for addition, or some kind of $\oplus$ but with weakened distributivity. Some ideas as to the latter were previously in (Majid, 2014, Sec. 7).

Heyting algebras provide the link up with probability as follows. We consider, as we did in Section 3, functions on X now with values in $[0,1]$, i.e. probabilities. $[0,1]$ itself is a Heyting algebra and $C(X,[0,1])$ inherits its features pointwise, including a preorder (where $f \leq g$ if $f(x) \leq g(x)$ for all $x \in X$), distributive, commutative and associative meet and join operations $(f \cup g)(x) = \max(f(x), g(x))$ and $(f \cap g)(x) = \min(f(x), g(x))$ as well as an internal hom $\underline{\mathrm{hom}}(f,g)(x) = 1$ wherever $f(x) \leq g(x)$, and $g(x)$ at other points. The negation $\mathrm{hom}(f,0)$ here is 1 at the zeros of f and otherwise 0, which still interchanges $\cup$ and $\cap$ as for de Morgan duality but which loses information and hence is not a proper duality (indeed, double negation is not the identity except on the Boolean subalgebra with values in $\{0,1\}$). By contrast, we observe that there is a different complementation which does square to the identity,

$$(16) \qquad \bar{f}(x) = 1 - f(x), \quad \forall x \in X,$$

and which again still interchanges $\cap, \cup$ as expected for de Morgan duality. It sends high probability to low probability in the same spirit as discussed in the Boolean case (as well as in the spirit of the prophetic work of the English satirist Douglas Adams). Moreover, we can still take the preorder 'field theory' point of view and do differential geometry and perhaps variational calculus on $A = \mathbb{C}(C(X,[0,1]))$ as above, but for the geometry of $C(X,[0,1])$ itself, we again do not have an algebra due to lack of a proper addition.

In fact, $[0,1]$ and $C(X,[0,1])$ have another product which is the more obvious one given by the usual product in $[0,1]$ as real numbers and the ordinary pointwise product fg of functions with values in $[0,1]$. We can then generate its 'de Morgan dual' addition via the complementation (16) as a new operation which we denote $f \oplus g =$

$1-(1-f)(1-g) = f+g-fg$ in the spirit of (Majid, 2020a) for the $\mathbb{F}_2$ case. This is associative and has a zero but one has $f(g \oplus h) \leq fg \oplus fh$ rather than full distributivity, so this pair of operations again does not quite give an algebra. This product in $\mathbb{R}$ again makes both $[0,1]$ and $C(X, [0,1])$ closed symmetric monoidal preorders, but different from our previous ones using $\cap$. The internal hom in $[0,1]$ is $\underline{\mathrm{hom}}(p,q) = 1$ if $p \leq q$ else q/p and similarly for functions $\underline{\mathrm{hom}}(f,g)(x) = 1$ at points where $f(x) \leq g(x)$, and $g(x)/f(x)$ at other points. In fact the negation $\underline{\mathrm{hom}}(\ ,0)$ here is the same as before (but the internal hom's in general are different).

All four binary operations on $C(X, [0,1])$ remain tightly related and we can consider the usual pointwise fg product along with $f \cup g$ given by the pointwise maximum. This has distributivity $f(g \cup h) = (fg) \cup (fh)$ but, as mentioned, $\cup$ is not an addition law. On the other hand, if we now replace probabilities in $[0,1]$ by minus their logarithm as we did in Section 3 in discussing geodesic paths, then we equivalently have functions on X with values in $\mathbb{R}_{\geq 0}$ with pointwise addition and minimum, i.e. valued in the tropical version of $\mathbb{R}$. This point of view has, for example, applications in statistical inference (Pachter and Sturmfels, 2004). So one may possibly garner ideas for differential geometry on X itself from tropical algebraic geometry. From probability functions to positive linear functionals on noncommutative algebras is a further but well-known step. In this way, we have sketched a path from Boolean algebras or logic to probability to full quantum geometry over $\mathbb{C}$, with some kind of generalised de Morgan duality playing a pivotal role.

Finally, it should be mentioned that at the time of (Majid, 1991), such duality ideas motivated the view that quantum gravity needs geometry that is at the same time quantum or noncommutative, with the duality realised slightly differently in concrete 'toy models' (Ma-

jid, 1988) as observer-observed/representation-theoretic Hopf algebra duality. The bicrossproduct quantum groups associated to Lie group factorisations emerging from this, as well as the Drinfeld-Jimbo one q-deforming complex simple Lie groups, contributed to a concrete 'constructive' approach to such quantum Riemannian geometry and included the first convincing model (Majid and Ruegg, 1994) of quantum spacetime with quantum symmetry. These ideas are also tied up with quantum group Fourier transform and in physical terms with 'Born reciprocity' and were at the root of my proposal back in (Majid, 1988; 1991) as well in later works (Majid, 2012; 2015; 2014; 2018). Although now somewhat established, the deepest aspect of this duality – swapping observables and states – remains unexplored and should relate to issues of measurement, probability and logic much as above, for example to the entropic arrow of time (Majid, 1993). How exactly it relates to a duality growing out of de Morgan duality remains a topic for further thought, albeit the notion of bi-Heyting algebra (Reyes and Zolfaghari, 1996) could be a step in that direction.

6. Concluding remarks

The main new results of the paper are our quantum geometric view of Markov processes and the new concept of an underlying 'Schrödinger process' in Section 3. We conclude with some comments on these.

Our first comment is that it could potentially be interesting to extend the graph calculus used here to quivers (where there can be self-arrows *and* multiple arrows between vertices). We have already seen the need for self-arrows but now we can go one step further to multiple arrows. One still has a differential calculus in the sense of part of a DGA but now not all 1-forms need be sums of elements

of the form $a\mathrm{d}b$ (Majid and Tao, 2019). One can still do quantum Riemannian geometry but now a metric is not a number on each edge but a matrix (Majid and Tao, 2019). In the Markov case, we could imagine a completely positive matrix, though this remains to be established. The goal would be to generalise Markov processes to allow different flavours of transition e.g. due to different types of processes between vertices. We also note (Pachter and Sturmfels, 2004) concerning the use of graphs with arrows labelled by linear forms to describe statistical models. Moreover, transitively closed extended graphs (i.e., preorders) are baby versions of categories and one could ask if differential geometry could generalise further to categories in the spirit of Section 5.

Our second comment concerns the notion of geodesics. In non-commutative geometry, one does not have points, so nor does one have geodesics as paths. Instead we have to work directly with functions and our first thought might be probability density functions. However, quantum Riemannian geometry in the constructive form (Beggs and Majid, 2020) is formulated in a linear setting so one is led to 'amplitudes' ψ in a Hilbert space evolving in time and probability density $f = |\psi|^2$. This was the motivation behind (Beggs, 2020; Beggs and Majid, 2019) and indeed quantum mechanics seems to be tightly linked with this point of view. This suggests that the philosophy of quantum mechanics, the measurement problem and so forth might be clearer as geodesic flow in quantum geometry. This, and more generally the role of quantum geometry (which is about extending macroscopic concepts to the quantum level) in the nature of measurement are topics for further study.

Indeed, the concrete result in the present paper is that if one goes further and looks at discrete-time Schrödinger processes in discrete quantum geometry, then coming out of the discreteness is a correction

to the familiar equation $\dot{f} = -\nabla \cdot J$ in usual quantum mechanics, in which there is an extra Markov process induced on f. This has more of a classical flavour and could play a role in measurement collapse. Such finite processes should also be of interest in quantum computing, where finite-dimensional unitary matrices are 'gates' and we have shown how a 1-parameter family of these naturally arises quantum geometrically.

Bibliography

Ambjørn, J., Jurkiewicz, J. and Loll, R., 2001. Dynamically triangulating Lorentzian quantum gravity. *Nuclear Physics B* [Online], 610(1), pp.347–382. Available at: https://doi.org/10.1016/S0550-3213(01)00297-8.

Argota-Quiroz, J.N. and Majid, S., 2020. Quantum gravity on polygons and $\mathbb{R} \times \mathbb{Z}_n$ FLRW model. *Classical and Quantum Gravity* [Online], 37, p.245001. Available at: https://iopscience.iop.org/article/10.1088/1361-6382/abbaa8/pdf.

Ashtekar, A., Pawlowski, T., Singh, P. and Vandersloot, K., 2007. Loop quantum cosmology of k = 1 FRW models. *Physical Review D* [Online], 75(2), p.024035. Available at: https://doi.org/10.1103/PhysRevD.75.024035.

Bassett, M.E. and Majid, S., 2020. Finite Noncommutative Geometries Related to $\mathbb{F}_p[x]$. *Algebras and Representation Theory* [Online], 23(2), pp.251–274. Available at: https://doi.org/10.1007/s10468-018-09846-4.

Beggs, E., 2020. Noncommutative geodesics and the KSGNS construction. *Journal of Geometry and Physics* [Online], 158, p.103851. Available at: https://doi.org/10.1016/j.geomphys.2020.103851.

Beggs, E. and Majid, S., 2017. Spectral triples from bimodule connections and Chern connections. *Journal of Noncommutative Geometry* [Online], 11(2), pp.669–701. Available at: https://doi.org/10.4171/JNCG/11-2-7.

Beggs, E. and Majid, S., 2019. Quantum geodesics in quantum mechanics. *arXiv:1912.13376 [hep-th, physics:math-ph]* [Online]. Available at: http://arxiv.org/abs/1912.13376.

Beggs, E. and Majid, S., 2020. *Quantum Riemannian Geometry* [Online]. Vol. 355, *Grundlehren der mathematischen Wissenschaften*. Springer International Publishing. Available at: https://doi.org/10.1007/978-3-030-30294-8.

Beggs, E.J. and Majid, S., 2014. Gravity induced from quantum spacetime. *Classical and Quantum Gravity* [Online], 31(3), p.035020. Available at: https://doi.org/10.1088/0264-9381/31/3/035020.

Connes, A., 1994. *Noncommutative Geometry* (S.K. Berberian, Trans.). San Diego: Academic Press.

Döring, A. and Isham, C., 2011. "What is a Thing?": Topos Theory in the Foundations of Physics. In: B. Coecke, ed. *New Structures for Physics* [Online]. Vol. 813, *Lecture Notes in Physics*. Berlin, Heidelberg: Springer, pp.753–937. Available at: https://doi.org/10.1007/978-3-642-12821-9_13.

Dowker, F., 2013. Introduction to causal sets and their phenomenology. *General Relativity and Gravitation* [Online], 45(9), pp.1651–1667. Available at: https://doi.org/10.1007/s10714-013-1569-y.

Dubois-Violette, M. and Michor, P.W., 1996. Connections on central bimodules in noncommutative differential geometry. *Journal of Geometry and Physics* [Online], 20(2), pp.218–232. Available at: https://doi.org/10.1016/0393-0440(95)00057-7.

Fong, B. and Spivak, D.I., 2019. *An Invitation to Applied Category Theory: Seven Sketches in Compositionality*. Cambridge: Cambridge University Press.

Hale, M., 2002. Path integral quantisation of finite noncommutative geometries. *Journal of Geometry and Physics* [Online], 44(2), pp.115–128. Available at: https://doi.org/10.1016/S0393-0440(01)00064-X.

Lawvere, F.W., 1991. Intrinsic co-Heyting boundaries and the Leibniz rule in certain toposes. In: A. Carboni, M.C. Pedicchio and G. Rosolini,

eds. *Category Theory* [Online], Lecture Notes in Mathematics. Berlin, Heidelberg: Springer, pp.279–281. Available at: https://doi.org/10.1007/BFb0084226.

Majid, S., 1988. Hopf algebras for physics at the Planck scale. *Classical and Quantum Gravity* [Online], 5(12), p.1587. Available at: https://doi.org/10.1088/0264-9381/5/12/010.

Majid, S., 1991. Principle of Representation-Theoretic Self-Duality. *Physics Essays* [Online], 4(3), pp.395–405. Available at: https://doi.org/10.4006/1.3028923.

Majid, S., 1993. Quantum random walks and time reversal. *International Journal of Modern Physics A* [Online], 08(25), pp.4521–4545. Available at: https://doi.org/10.1142/S0217751X93001818.

Majid, S., ed., 2012. *On Space and Time* [Online], *Canto Classics*. Cambridge: Cambridge University Press. Available at: https://doi.org/10.1017/CBO9781139197069.

Majid, S., 2013. Noncommutative Riemannian geometry on graphs. *Journal of Geometry and Physics* [Online], 69, pp.74–93. Available at: https://doi.org/10.1016/j.geomphys.2013.02.004.

Majid, S., 2014. The self-representing Universe. In: B. Carr, M. Eckstein, M. Heller and S.J. Szybka, eds. *Mathematical structures of the universe*. Kraków: Copernicus Center Press, pp.357–387.

Majid, S., 2015. Algebraic approach to quantum gravity I: relative realism. In: J. Ladyman et al., eds. *Road to reality with Roger Penrose*. Kraków: Copernicus Center Press, pp.117–177.

Majid, S., 2018. On the emergence of the structure of physics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* [Online], 376(2118), p.20170231. Available at: https://doi.org/10.1098/rsta.2017.0231.

Majid, S., 2019a. Quantum gravity on a square graph. *Classical and Quantum Gravity* [Online], 36(24), p.245009. Available at: https://doi.org/10.1088/1361-6382/ab4975.

Majid, S., 2019b. Quantum Riemannian geometry and particle creation on the integer line. *Classical and Quantum Gravity* [Online], 36(13), p.135011. Available at: https://doi.org/10.1088/1361-6382/ab2424.

Majid, S., 2020a. Quantum geometry of Boolean algebras and de Morgan duality. *Journal of Noncommutative Geometry*, in press. Available at: arXiv: 1911.12127 [math.QA].

Majid, S., 2020b. Reconstruction and quantization of Riemannian structures. *Journal of Mathematical Physics* [Online], 61(2), p.022501. Available at: https://doi.org/10.1063/1.5123258.

Majid, S. and Pachoł, A., 2018. Classification of digital affine noncommutative geometries. *Journal of Mathematical Physics* [Online], 59(3), p.033505. Available at: https://doi.org/10.1063/1.5025815.

Majid, S. and Pachoł, A., 2020a. Digital finite quantum Riemannian geometries. *Journal of Physics A: Mathematical and Theoretical* [Online], 53(11), p.115202. Available at: https://doi.org/10.1088/1751-8121/ab1cf2.

Majid, S. and Pachoł, A., 2020b. Digital quantum groups. *Journal of Mathematical Physics* [Online], 61(10), p.103510. Available at: https://doi.org/10.1063/5.0020958.

Majid, S. and Ruegg, H., 1994. Bicrossproduct structure of κ-Poincare group and non-commutative geometry. *Physics Letters B* [Online], 334(3), pp.348–354. Available at: https://doi.org/10.1016/0370-2693(94)90699-8.

Majid, S. and Tao, W.-Q., 2019. Generalised noncommutative geometry on finite groups and Hopf quivers. *Journal of Noncommutative Geometry* [Online], 13(3), pp.1055–1116. Available at: https://doi.org/10.4171/JNCG/345.

Pachter, L. and Sturmfels, B., 2004. Tropical geometry of statistical models. *Proceedings of the National Academy of Sciences* [Online], 101(46), pp.16132–16137. Available at: https://doi.org/10.1073/pnas.0406010101.

Reyes, G.E. and Zolfaghari, H., 1996. Bi-Heyting algebras, toposes and modalities. *Journal of Philosophical Logic* [Online], 25(1), pp.25–43. Available at: https://doi.org/10.1007/BF00357841.

Information and physics

Radosław A. Kycia

Department of Mathematics and Statistics, Masaryk Univeristy
Faculty of Materials Science and Physics, Cracow University of Technology

Agnieszka Niemczynowicz

Faculty of Mathematics and Computer Science,
University of Warmia and Mazury

Abstract

This is an overview article that contains the discussion of the connection between information and physics at the elementary level. We present a derivation of Landauer's bound for heat emission during irreversible logical operation. In this computation the Szilard's version of Maxwell's demon paradox is used as a model to design thermodynamic implementation of a single bit of computer memory. Landauer's principle also motivates the discussion on the practical and emergent nature of the information. Apart from physics, the principle has implications in philosophy.

1. Introduction

Our civilization is based on information and creates it at a rapid rate. The common understanding of 'information' is related to:

- Representation using binary notation (for instance), i.e. two values/bits defined as 0 and 1, associated with 'true' and 'false', which have their own logical operations. In this approach, information represents a sequence of answers for binary questions (Reza, 1994). This approach is adopted in current computer systems.
- Boolean algebras (Reza, 1994). These are mathematical descriptions of logic (and information) mentioned in the previous point.
- Statistical description of information in the context of its effective encoding by the use of Shannon's information (or coding) theory (Reza, 1994). This approach is connected with statistics and a long sequence of bits.

The first question is therefore about the term 'information'. In general, the sequence of bits (e.g. 0101100101) is useless unless the context in which it is produced is provided. In other words, a sequence of bits is a sequence of 'true/false' answers to specific questions. If they are 'not-overlapping', then they increase our knowledge of the nature of the subject being examined. For instance, for a question about a particular object, we can prepare the following sequence: Is it round?, Is it green?, and so forth. Each question that is substantially different from the previous yields an answer that provides us with additional knowledge about the object. This is how we gather 'information' about the world. The representation is a way how we encode this knowledge.

Summarizing, we say that information is knowledge (sequence of answers for substantially different questions) about some objects or phenomena. It is represented by (encoded in) a sequence of letters from some alphabet. Then it can be analyzed and processed in terms of, e.g., statistical analysis or Boolean algebra. Often in literature both

of these notions are merged (Reza, 1994) or the term information is related to its representation and processing without its semantic. Both these concepts are abstract, in the sense that they are not related to any physical phenomenon.

Modern computer systems exclusively use a binary representation of information. However, if one considers arbitrary base $b > 1$ for representation of numbers, then according to Shannon's theory (Reza, 1994), the information per digit is described by $\frac{lnb}{b}$. This function has its maximum at $b = e = 2.7\ldots$, that is, the base of natural logarithms (Napier's or Euler's number). This is an irrational number and its use as a base is difficult to achieve, although not impossible (Hayes, 2001). This fact has also motivated the design of computer systems based on trit ($b = 3$), which is closer to optimal coding base e than to the binary base ($b = 2$). Due to their more complicated construction, such systems are not common. However, this example shows that the encoding we use to collect information is the second-close to optimal.

The next question concerns the nature of the information. Is it an abstract entity, or does some physical quantity represent it? Rolf Landauer (1961) first highlighted the connection between the abstract world of information presented above and its physics representation. It was later popularized by Charles H. Bennett (1973), who also proposed the application of Landauer's principle to resolve the long-lasting paradox of Maxwell's demon (Bennett, 1987).

The principle states that during irreversible (i.e. non-bijective) computational processes, heat is emitted in the memory. The bound for the emitted heat is given by Landauer's bound:

$$(1) \qquad Q_L = k_B T \ln 2,$$

per bit, where T is the temperature at which system operates and $k_B = 1.380649 \times 10^{-23} \frac{J}{K}$ is the Boltzmann constant. Due to the

small value of k_B, this heat does not generate overheating in modern computing systems. This principle states that operations on a physical representation of information produce physical/measurable consequences.

Landauer's principle has recently undergone heavy testing (Bérut et al., 2012), including recent tests in quantum systems (Yan et al., 2018). It has also been theoretically investigated (Ladyman et al., 2007; Sagawa, 2014; Piechocinska, 2000; Still, 2020; Kycia, 2018; Kycia, 2019) from different viewpoints. Moreover, an application to multivalued logic has been proposed (Bormashenko, 2019; Kycia and Niemczynowicz, 2020). This raises questions about the physicality of information: at which level is Landauer's heat expelled and how it is associated with the notion of information?

In this paper we present the ideas behind Landauer's principle and its connection with Maxwell's demon paradox on the elementary level. We also use some of our recent results from (Kycia and Niemczynowicz, 2020) that contain a general method of constructing a higher base thermodynamic analog of memory and its connection with category theory, namely, the Galois connection. Therefore, the paper is a review and not original research. We will treat the subject elementarily, to explain Landauer's principle to the broad community of philosophers, physicists, and a general audience. We will also trace different concepts related to information: an abstract entity including its representation in terms of a sequence of letters of the alphabet, e.g., bits, and its representation in the configuration of a physical system.

The paper is organized as follows. In the next section we describe a simple version of Maxwell's demon paradox attributed to Szilard (1929) and its resolution by the Landauer's principle. We then use this model to construct a single bit of thermodynamic version of computer memory. This model explains a subtle connection between

information and its physical realization. This will be simplified here, with a more pedagogical version of the results presented in (Kycia and Niemczynowicz, 2020). Finally, we comment on the relationship between abstract information and its representation using physical systems.

2. Maxwell's demon

Maxwell's demon is a common paradox of thermodynamics that was solved only recently utilizing Landauer's principle (Landauer, 1961; Bennett, 1973; Feynman and Hey, 2000) which is currently accepted and verified experimentally solution as explained in the Introduction. The physics and philosophy literature on this subject is currently vast, and the good starting point that merges both disciplines is (Ladyman et al., 2007).

In this section we present a simple version of Szilard's experiment (Szilard, 1929; Feynman and Hey, 2000), which uses a single particle of the ideal gas, with the following equation of state (for the particle) (Kycia, 2019; Lychagin, 2019; Schneider, 2019)

$$(2) \qquad pV = k_B T,$$

where p is the pressure and V is the available volume for the particle. In the experiment, all elements are 'idealized' in the sense that there are no energy losses for friction.

Imagine that the particle of the ideal gas is inside the box of temperature T. The particle is in thermal equilibrium with the walls of the box. We construct a thermodynamic cycle that extracts useful work from the system. Maxwell's demon is a new system that ensures the proper course of the cycle.
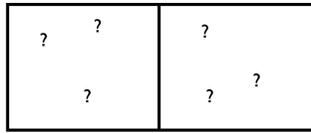
Figure 1: The box split in half, with the unknown location of a particle.
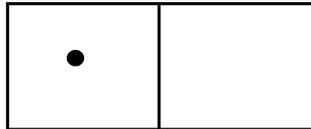

Figure 2: In this case the particle is located in the left chamber of the box.

In the first step of the cycle, insert a wall inside the box, dividing it into two chambers of equal volume $\frac{V}{2}$. At this stage, the demon that guides the machine cycle does not know in which part of the box the particle can be found. The situation is presented in Fig. 1.

To start extracting useful work, the demon must locate the particle, or in other words, gain and store information about particle location. The situation after localizing the particle is presented in Fig. 2.

In the next step, the border between the chambers become movable. Energy is extracted when the gas is decompressed isothermally. This decompression is a reversible thermodynamic process and the work done by the particle is given by:

$$(3) \qquad W = k_B T \int_{V/2}^{V} = k_B T \ln 2.$$

The situation is presented in Fig. 3.

The final part of the cycle is the restart. The demon takes the border out of the box, moves it and then re-insert it again in the
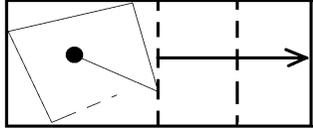
Figure 3: Energy extraction in the isothermal expansion of a single particle of the ideal gas.

middle. We return here to the situation from Fig. 1. At this stage, information about the particle location is not needed, so the demon must delete it to return the whole system (box and demon) to the initial state for another round of the cycle. Therefore, we now have the same state as at the beginning of the previous turn.

This situation can be analyzed in terms of the second law of thermodynamics and entropy S. The second law of thermodynamics says that the change of total entropy (system and environment) must be non-decreasing.

In the case above, in the cycle, according to the conservation of energy (the first law of thermodynamics), the heat from the environment (the thermal bath of the particle) is extracted and used to make work W. Therefore, the deficiency of heat is $Q = W = k_B T \ln(2)$. Given that the balance is negative and the system is held at a constant temperature, the entropy change in the cycle is

$$(4) \qquad \Delta S = -\frac{Q}{T} = -k_B \ln(2) < 0.$$

This contradicts the second law (it must be $\Delta S \geq 0$).

To solve this apparent contradiction, Bennett (1987) has suggested that during the demon's erasure of the bit of information regarding particle location (say 0–the particle in the left chamber, 1–the particle in the right chamber) the heat of value Q_e is expelled. This heat is no

less than Landauer's bound (1), that is $Q_e \geq Q_L$. This assumption corrects the balance of entropy:

$$(5) \qquad \Delta S = \frac{Q_e}{T} - \frac{Q}{T} \geq k_B \ln(2) - k_B \ln(2) = 0.$$

Bennett's explanation solves the long-standing paradox by employing Landauer's idea. It is interesting to note that the paradox remained unresolved until the 1970s. Moreover, its resolution is based on the idea of information, computing, and memory developed in the twentieth century and therefore not known to the fathers of thermodynamics. The solution can almost be regarded as manifest in the works of Szilard (1929) if his experiment is interpreted as a model for computer memory. This reinterpretation of Szilard's results is presented in the next section, which provides a more in-depth insight into the connection between information and its implementation.

3. Memory

Szilard's *gedanken experiment* presented in the previous section can be used to construct a thermodynamic model of computer memory. We consider here the simplest case; a more detailed presentation for general memory for multivalued logic is given in (Kycia and Niemczynowicz, 2020).

The implementation of a single bit thermodynamic memory is presented in Fig. 4.

State a is an inner state of memory that does not represent any logical value. It is used to implement all logical operations (not necessarily those that swap 0 and 1). This assumption is consistent with the previous discussion of Szilard's model and its reverse.
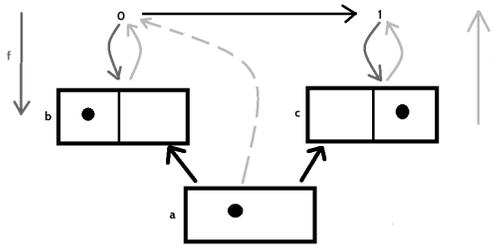
Figure 4: Implementation of memory using the thermodynamic model of the reversed Szilard's model of Maxwell's demon. State a is not associated with logical value: it is an undetermined internal state of the memory, which can be associated with 0 at the logical level. State b represents 0, and state c represents 1. Maps f and g, associating logical values with their physical implementation, are selected in such a way to preserve the order indicated by black lines representing transitions between states or ordering in Boolean algebra, that is, logic. This is a reminiscence of the Galois connection (Fong and Spivak, 2019; Smith, 2018; Kycia, 2018) that usually occurs where there is a system that contains an abstract level and its implementation.

We can now proceed to a detailed description of the types of operations. At the level of logic, we have two types of operations on this memory:

- Reversible: operations that have inverses, i.e., those that are bijections, for instance, a NOT gate whose function is involutive: $0 \leftrightarrow 1$.
- Irreversible: operations that do not have inverse, for instance the assignment of some value that 'forgets' the previous value: $(0 \rightarrow 1, 1 \rightarrow 1)$.

At the level of thermodynamics, we have a similar convention. However, the definition is different:

Possibilities	B reversible	B irreversible
A reversible	YES	YES
A irreversible	NO	YES

Table 1: Systems A is implemented on B.

- Reversible: realized by reversible thermodynamic processes, i.e. processes that can be reversed, or equivalently, the process that combined into a cycle with its inverse has vanishing entropy change $\Delta S = 0$.
- Irreversible: realized by irreversible thermodynamic processes, i.e. if we perform this operation and then return (along any reversible process) to the initial state, then in the cycle $\Delta S > 0$.

As highlighted in (Ladyman et al., 2007), the main idea behind Landauer's principle is the association of operations at the level of logic and their physical implementation in the form presented in Tab. 1; in our case, A is the logical system and B is its memory implementation.

To understand this association, we provide some examples of logical operations and their physical implementations.

To assign some initial value to memory state, let us start from state a and perform isothermal compression to b. This initial preparation expels (stores) the heat $Q = k_B T \ln(2)$ to(in) the environment and sets the initial logical state to 0.

Consider the reversible NOT operation. According to Tab. 1, this can be realized in two inequivalent thermodynamically ways at the physical level:

1. Reversible thermodynamic operation: isothermal expansion $b \to a$ and then isothermal compression $a \to c$. The composed operation is thermodynamically reversible and no heat is expelled to the environment or converted to work.

2. Irreversible thermodynamic operation: we take the middle border out of the box, which generates free adiabatic decompression of the gas (particle) $b \to a$. Next, reversible isothermal compression $a \to c$ is made. When the first step is performed, we lose 'access' to the heat initially stored in 0 state. This heat cannot be regained/reused and therefore must be treated as Landauer's heat.

Consider now the irreversible logical operation of setting 1 starting from 0. According to Tab. 1, this must be realized by an irreversible thermodynamic operation and hence it is the same as the operation from the second point above. Therefore, Landauer's heat is also expelled in this case.

This example is also proof that Landauer's principle is equivalent (at least for these examples) to the association from Tab. 1 between operations on logical and thermodynamic states.

The presentation in this section was elementary, intending to be an introduction to the multivalued logic approach (Kycia and Niemczynowicz, 2020; Bormashenko, 2019), category theory formulation of this principle (Kycia, 2018; Kycia, 2019), mentioned briefly above, and its understanding in more fundamental level (Ladyman et al., 2007). We also emphasize that the principle can be thoroughly understood only at the level of statistical physics (Sagawa, 2014; Piechocinska, 2000), and equilibrium thermodynamics used above is only its limiting case.

In the following conclusion section we briefly discuss the nature of information.

4. Discussion

The thermodynamic realization of a bit of information is a simple model that presents how Landauer's heat is generated for irreversible operations. In this model, energy is initially stored in the environment by setting some starting value (0 or 1) in the memory. An irreversible operation then makes it impossible to reuse/regain this heat to perform work and thus extract energy from the memory to gain the initially 'invested' energy. In this interpretation, a bit labels the state of a system that contains energy that can be accessed and (re)used. Reversible operations are designed to extract this energy, whereas irreversible ones lose it. Therefore, according to the laws of thermodynamics, for irreversible operation, heat is expelled.

Abstracting even more, the information is stored in some sort of energy configuration of the physical system. From the mathematical viewpoint, there is an abstract notion of information (bits, Boolean algebras), but ultimately information must be represented by certain physical states, allowing it to interact with the physical world. This storing process of information forms a connection between its abstract notion and the physical world.

This also has some practical consequences to philosophy. The 'sacrum' (abstraction) cannot be separated from 'profane' (realization). Abstract concepts (encoded as some form of information) must always be stored in some physical configuration of matter/energy. Information is a type of emergent concept or a new complexity built on top of physical constructs and our ability to manipulate energy and matter. This is a highly pragmatic approach.

This perspective shows that information, although immaterial concept, exists in the Universe as its encoding/representation in physical content of matter and energy. Their properties can be abstracted, how-

ever, we always deal and interact with some of its representation. The connection between abstract notion and its multiple (physical) representations is also reflected in the Galois connection (Smith, 2018)—a process of abstraction—which can be seen as a modern allegory of Plato's cave—a projection of mathematical (ideal) entity to the concrete realization (shadow on the cave wall).

5. Conclusion

We presented an overview of the basic facts from the thermodynamics of Maxwell's demon, its model by Szilard and the resolution of this paradox by Landauer. This motivated us to consider the correspondence between information and its physical representation. The aim was to introduce a broad audience to the basics of this subject and provide some references that can be used for more in-depth studies.

Bibliography

Bennett, C.H., 1973. Logical reversibility of computation. *IBM Journal of Research and Development* [Online], 17(6), pp.525–532. Available at: https://doi.org/10.1147/rd.176.0525 [visited on 14 November 2020].

Bennett, C.H., 1987. Demons, engines and the second law. *Scientific American* [Online], 257(5), pp.108–116. Available at: https://doi.org/10.1038/scientificamerican1187-108 [visited on 14 November 2020].

Bérut, A. et al., 2012. Experimental verification of Landauer's principle linking information and thermodynamics. *Nature* [Online], 483(7388), pp.187–189. Available at: https://doi.org/10.1038/nature10872 [visited on 13 November 2020].

Bormashenko, E., 2019. Generalization of the Landauer principle for computing devices based on many-valued logic. *Entropy* [Online], 21(12), p.1150. Available at: https://doi.org/10.3390/e21121150 [visited on 13 November 2020].

Feynman, R.P. and Hey, A.J.G., 2000. *Feynman Lectures on Computation*. 1. paperback print., [Nachdr.], *The advanced book program*. Reading, Mass: Westview.

Fong, B. and Spivak, D.I., 2019. *An Invitation to Applied Category Theory: Seven Sketches in Compositionality* [Online]. Cambridge: Cambridge University Press. Available at: https://doi.org/10.1017/9781108668804 [visited on 14 November 2020].

Hayes, B., 2001. Third base. *American Scientist* [Online], 89, pp.490–494. Available at: http://bit-player.org/wp-content/extras/bph-publications/AmSci-2001-11-Hayes-ternary.pdf [visited on 14 November 2020].

Kycia, R., 2018. Landauer's principle as a special case of Galois connection. *Entropy* [Online], 20(12), p.971. Available at: https://doi.org/10.3390/e20120971 [visited on 13 November 2020].

Kycia, R.A., 2019. Entropy in themodynamics: from foliation to categorization. *arXiv:1908.07583 [cond-mat, physics:math-ph]* [Online]. Available at: http://arxiv.org/abs/1908.07583 [visited on 13 November 2020].

Kycia, R.A. and Niemczynowicz, A., 2020. Simple explanation of Landauer's bound and its ineffectiveness for multivalued logic. *arXiv:2001.00942 [cond-mat, physics:math-ph]* [Online]. Available at: http://arxiv.org/abs/2001.00942 [visited on 13 November 2020].

Ladyman, J., Presnell, S., Short, A.J. and Groisman, B., 2007. The connection between logical and thermodynamic irreversibility. *Studies in History*

and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics [Online], 38(1), pp.58–79. Available at: https://doi.org/ 10.1016/j.shpsb.2006.03.007 [visited on 13 November 2020].

Landauer, R., 1961. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development* [Online], 5, pp.183–191. Available at: http://worrydream.com/refs/Landauer%5C%20-%5C% 20Irreversibility%5C%20and%5C%20Heat%5C%20Generation% 5C%20in%5C%20the%5C%20Computing%5C%20Process.pdf [visited on 14 November 2020].

Lychagin, V.V., 2019. Contact Geometry, Measurement, and Thermodynamics. In: R.A. Kycia, M. Ułan and E. Schneider, eds. *Nonlinear PDEs, Their Geometry, and Applications* [Online]. Cham: Springer International Publishing, pp.3–52. Available at: https://doi.org/10.1007/978-3-030-17031-8_1 [visited on 14 November 2020].

Piechocinska, B., 2000. Information erasure. *Physical Review A* [Online], 61(6), p.062314. Available at: https://doi.org/10.1103/PhysRevA.61. 062314 [visited on 13 November 2020].

Reza, F., 1994. *An Introduction to Information Theory*. New York: Dover Publications, Inc.

Sagawa, T., 2014. Thermodynamic and logical reversibilities revisited. *Journal of Statistical Mechanics: Theory and Experiment* [Online], 2014(3), P03025. Available at: https://doi.org/10.1088/1742-5468/2014/03/ P03025 [visited on 13 November 2020].

Schneider, E., 2019. Differential Invariants in Thermodynamics. In: R.A. Kycia, M. Ułan and E. Schneider, eds. *Nonlinear PDEs, Their Geometry, and Applications* [Online]. Cham: Springer International Publishing, pp.223–232. Available at: https://doi.org/10.1007/978-3-030-17031-8_7 [visited on 14 November 2020].

Smith, P., 2018. *Category Theory: A Gentle Introduction* [Online]. Available at: https://www.logicmatters.net/resources/pdfs/GentleIntro.pdf [visited on 14 November 2020].

Still, S., 2020. Thermodynamic cost and benefit of memory. *Physical Review Letters* [Online], 124(5), p.050601. Available at: https://doi.org/10.1103/ PhysRevLett.124.050601 [visited on 13 November 2020].

Szilard, L., 1929. Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. *Zeitschrift für Physik* [Online], 53(11-12), pp.840–856. Available at: https://doi.org/10.1007/BF01341281 [visited on 13 November 2020].

Yan, L.L. et al., 2018. Single-atom demonstration of the quantum Landauer principle. *Physical Review Letters* [Online], 120(21), p.210601. Available at: https://doi.org/10.1103/PhysRevLett.120.210601 [visited on 13 November 2020].

The homunculus brain and categorical logic

Steve Awodey

Departments of Philosophy and Mathematics, Carnegie Mellon University

Michael Heller

Copernicus Center for Interdisciplinary Studies, Jagiellonian University

Abstract

The interaction between syntax (formal language) and its semantics (meanings of language) is one which has been well studied in categorical logic. The results of this particular study are employed to understand how the brain is able to create meanings. To emphasize the toy character of the proposed model, we prefer to speak of the homunculus brain rather than the brain *per se*. The homunculus brain consists of neurons, each of which is modeled by a category, and axons between neurons, which are modeled by functors between the corresponding neuron-categories. Each neuron (category) has its own program enabling its working, i.e. a theory of this neuron. In analogy to what is known from categorical logic, we postulate the existence of a pair of adjoint functors, called Lang and Syn, from a category, now called BRAIN, of categories, to a category, now called MIND, of theories. Our homunculus is a kind of "mathematical robot", the neuronal architecture of which is not important. Its only aim is to provide us with the opportunity to study how such a simple brain-like structure could "create meanings" and perform abstraction operations out of its purely syntactic program. The pair of adjoint functors Lang and Syn model the mutual dependencies between the syntactical structure of a given theory of MIND and the internal logic of its semantics given

by a category of BRAIN. In this way, a formal language (syntax) and its meanings (semantics) are interwoven with each other in a manner corresponding to the adjointness of the functors Lang and Syn. Higher cognitive functions of abstraction and realization of concepts are also modelled by a corresponding pair of adjoint functors. The categories BRAIN and MIND interact with each other with their entire structures and, at the same time, these very structures are shaped by this interaction.

1. Introduction: On the Computer Screen

We were preparing a paper for publication. A phase portrait was nicely displayed on the computer screen. The network of trajectories represented a class of solutions to the equation we were interested in. At some points, called critical points, certain trajectories crossed each other. These points were important for our analysis. Some of the diagrams we worked with appeared later as figures in our publication (Woszczyna and Heller, 1990). The figures had to be explained, so we decided to attach appropriate labels to some of the critical points. We attached the label "stable saddle" to one of them. No problem. Then we proceeded to attach the label "unstable saddle" to another one. But the label jumped up. We tried to fix it up, but it jumped down. Then we started laughing. After all, it is an unstable point!

Let us try to understand the situation. We were investigating an equation that (virtually) contains in itself its space of solutions (irrespectively of whether we explicitly know them or not). Through the suitable computer program and some "electronic circuits", which are activated by the program, this space of solutions is mapped into the phase portrait displayed on the computer screen. The diagram we see on the screen is certainly something more than just a picture. It does not simply show stable and unstable critical points; it also does what the abstract equation orders its solutions to do (labels jump up and down at instabilities).

Let us go a step forward. In fact, the phase portrait on the screen is a substitute of the world. For suppose that our equation "describes" (or better—models) a mechanical system (e.g., a pendulum or oscillator).[1] Then the unstable critical points of our equation correspond to physical situations in which the considered mechanical system behaves in an unstable way. We thus have, on the one hand, an equation (or a set of equations) or, more broadly, a mathematical theory and, on the other hand, a domain (or an aspect) of the physical world of which the considered mathematical theory is a model. Between the mathematical model and the domain (or aspect) of the physical world there is a mysterious correspondence—a correspondence in the root-meaning of this word: both sides co-respond to each other. It is an active correspondence, and the activity goes both ways: it looks as if the domain of the world informed the theory about its own internal structure, and the theory answered by prescribing what the domain should do. And the domain does it. The equations prescribe what the world should do, and the world executes this. The equations and the world are coupled with each other and act in unison.

---

[1] The equations we considered in our publication referred to a cosmological situation.

And the screen on my computer? It is a part of the world. The program we have constructed reads the structure of the equations and executes what the equations tell it. And because of the coupling between the equations and the world, the computer does, in miniature, what does the world on its own scale. This is the reason why computers are so effective in our reading of the structure of the world.

There is another domain in which a formal structure reveals its effective power and produces real effects. Such processes occur in the brain. The formal structure in question consists of electric signals propagating along nerve fibres between neurons across synapses, and the world of meanings should be regarded as a product of this activity. The interaction seems to go both ways: the "language of neurons" (what happens in the brain) produces the meanings related to this language (in the mind), and the meanings somehow influence the architecture of neurons.

It seems that in both these cases (mathematical laws and their effects in the real world, and the brain–mind interactions) we meet two instances of the same working of logic where syntax (a formal structure), by effectively interacting with its semantics, produces real effects. This kind of interaction, although kept strictly on the level of logic (i.e. with no reference to processes in the real world), is well known in the categorical logic. In the present paper, we attempt to employ these achievements of categorical logic to try to understand the brain–mind interaction.

The traditional terminology of brain and mind (irrespective of current trends in cognitive sciences to get rid of their conceptual load) seems especially well adapted to the present context in which general ideas are more important than structural details. Moreover, to

avoid too hasty associations with the human brain and to emphasize the toy character of the proposed model, we prefer to speak of the "homunculus brain" rather than just of brain.

The action of our argument develops along the following lines. Section 2 is a reminder on formal language, its syntax and semantics. Sections 3 and 4 briefly review those parts of categorical logic that refer to these concepts. Every category, call it C, has its internal logic, and if this logic is sufficiently rich, the category provides semantics for a certain formal theory T. Moreover, there exists a pair of adjoint functors, called Lang and Syn, from a category, called CATEGORIES, of categories belonging to a certain class (for instance, coherent categories) to a category, called THEORIES, of theories and *vice versa*, which describe mutual dependencies between the syntactical structure of T and the internal logic of its semantics given by C. This is described in section 3. In this way, syntax and semantics are interwoven with each other in a manner corresponding to the adjointness of the functors Lang and Syn. This is explored in section 4. In section 5, we consider a deep categorical duality between the syntactic category of a theory and its individual models and suggest a functional interpretation in terms of abstraction and realization of concepts, in anticipation of the cognitive interpretation to be introduced next. In section 6, the category CATEGORIES becomes the category BRAIN. It constitutes a simple model of a homunculus' brain. Objects of this category are categories (belonging to a certain class); every such category models a neuron. Morphisms of this category model signals propagating along nerve fibres between neurons. The category THEORIES becomes the category MIND. Its objects are "theories of neurons"; more precisely, if $C \in$ BRAIN, then its "theory" is Lang(C) in MIND. Morphisms of this category are functors between the corresponding syntactic theories; more precisely, if $T_1, T_2 \in$ MIND, then the morphism between

them is $\mathrm{Syn}(T_1) \to \mathrm{Syn}(T_2)$. The pair of adjoint functors Lang and Syn model the interaction between the syntax of "theories" and their semantics, i.e. the network of neurons. The categories BRAIN and MIND are indeed somehow related to what their names refer to, at least as far as homunculus' brain and mind are concerned.

Following the seminal paper of McCulloch and Pitts, published as early as in 1943, which proposed using classical logic to model neural processes in the brain, there have been so many papers developing and modifying (with various logical systems) this idea, that to quote even a sample of them would be immaterial (for a relatively recent state of art see a short review Koch, 1997). A. Ehresmann claims that it was R. Rosen who was the first to employ category theory to model biological systems. A series of works followed (a non-representative sample: (Gómez and Sanz, 2009; Healy and Caudell, 2006; Mizraji and Lin, 2011; Tsuchiya, Taguchi and Saigo, 2016)) proposing the use of various parts of category theory to model different aspects of the brain activity. In particular, adjoint functors were suggested to model "a range of universal-selectionist mechanisms" (Ellerman, 2015). However, we have not been able to find anything similar to modeling the interaction between brain's language and its meaning anywhere.

2. Syntax and Semantics

In linguistics, syntax and semantics are regarded as parts of semiotics, the study of signs. Syntax studies relations between signs, and semantics relations between signs and what the signs refer to.[2] Syntactic

---

[2] Sometimes one also distinguishes pragmatics which studies relations between signs and their users.

properties are attributed to linguistic expressions entirely with respect to their shape (or form). Semantics, on the other hand, endows them with meaning by referring signs to what they signify. Logic adapts these ideas to its own needs. Since it is a formal science, the signs it considers should be elements of a formal language, and they cannot refer to anything external. Halvorson puts it, "But a formal language is really not a language at all, since nobody reads or writes in a formal language. Indeed, one of the primary features of these so-called formal languages is that the symbols don't have any meaning" (Halvorson, 2016). This is why the meaning should be "artificially" constructed for them. The idea of how this should be done can best be seen in Tarski's prototype of this procedure (Tarski, 1933). If a sentence s, the truth of which we want to define, belongs to a language L then the definition of s should be formulated in a metalanguage M with respect to the language L. And the metalanguage M should contain a copy of s so that anything one can say with the help of s in L, can also be said in M. The definition of "True" should be of the form

$$\text{For all } x, \text{True}(x) \text{ if and only if } \varphi(x)$$

with the condition that "True" does not occur in φ. Here x stands for the copy of the sentence s in the metalanguage L, and $\varphi(x)$ describes, also in M, the state of affairs of which the sentence s in L reports (for more details see Hodges, 2018; Sher, 1999). A metalinguistic copy of s could also be expressed as "s" (taken in quotes). In Tarski's own example:

$$\text{"It snows" is true iff it snows.}$$

For pedagogical reasons, this example is taken from colloquial language, but strictly speaking Tarski's definition refers to formal languages. The formal language L has its own syntax (since it is a formal

language), but is lacking its semantic reference. As we have seen, such a reference had to be constructed for it with the help of the metalanguage M.

Now, the idea is to improve the situation by looking for such a conceptual context in which a semantics for a given theory would arise in a more natural (or even spontaneous) way.

3. Categorical Semantics

To do so we must first define precisely what we mean by language. Since the definition must be precise, let us choose as an example the language of mathematics based on standard first order logic (which is enough for most of the usual mathematics). Many other languages may be formalized in a similar way. In such a language we distinguish:

- constants: $0, 1, 2, \ldots, a, b, c, \ldots,$ and variables: $x, y, z, \ldots,$ which can be combined by primitive operations to give
- terms, for example: $x + y$, $x^3, \ldots$ which, in turn, can be combined, with the help of primitive relations, such as $=, <, \leq, \ldots,$ to produce
- formulae, for example: $x + y = z$, $x \leq y, \ldots$ which, in turn can be combined, with the help of the usual logical connectives and quantifiers, into
- more complicated formulae.

To make the language more flexible and more adapted for concrete applications, we diversify its expressions into various types (called also sorts). In mathematics, we might use different letters for natural

and real numbers, or different symbols for vectors an scalars. We say that, in both cases, we are using a two-typed language. There may be languages with as many types as is needed.

What we need is not so much a language, but rather a theory. In mathematical logic theory is almost the same as language; it is a formal language aimed at axiomatizing a certain class of sentences. The concept of theory, as it is functioning in modern physics can, in principle, be regarded as the special case of the logical concept of theory, although in scientific practice theories are rarely formulated with the full logical rigor.

Let then T be a theory expressed in a multi-type language. Such a theory is defined as consisting of the following data:

1. A set of types $\{X_1, X_2, \ldots, X, Y, \ldots\}$.
2. A set of variables $\{x, y, z, \ldots, x_1, x_2, x_3, \ldots\}$ with a type assigned to each variable.
3. A set of function symbols with a type assigned to each domain and codomain of every function symbol; for instance, to the term $x_1 + x_2$, with the variable x_1 of type X_1 and the variable x_2 of type X_2, there corresponds the function symbol $f : X_1 \times X_2 \to Y$, and the term $f(x_1, x_2) = x_1 + x_2$ is of type Y.
4. A set of relation symbols with a type assigned to each argument of every relation symbol; for instance, to the formula $x + y = z$, with the variable x of type X_1, the variable y of type X_2 and the variable z of type X_3, there corresponds the relation symbol $R \subseteq X_1 \times X_2 \times X_3$, and $R(x, y, z)$ is an atomic formula.
5. A set of logical symbols.
6. A set of axioms for a given theory built up from terms and relation symbols with the help of logical connectives and quantifiers, respecting types of all terms.

This is, in fact, a purely syntactic definition of theory (for details see Borceux, 1994, pp.344–348; Mac Lane and Moerdijk, 1992, pp.527–530). Now, we want to create a semantics, i.e. a model, for a theory T. This is done by constructing a category C_T which will serve us as such a model. The construction is almost obvious:

1. each type of T is an object of C_T,
2. for each function symbol f in T with types A and B as its domain and codomain, correspondingly, f is a morphism from the object A to the object B in C_T,[3]
3. variables are identity morphisms in C_T,
4. for each relation symbol R in T, its counterpart in C_T is a subobject in C_T. Suppose ϕ is a subobject of an object A in C_T then, by analogy with the usual theory of sets, ϕ can be thought of as a collection of all things of type A that verify ϕ.

This definition must be supplemented with all of the (first order) logic which is used to express axioms in T (for details see nLab, 2017). Roughly speaking, since formulae correspond to subobjects, and all subobjects of a given object are partially ordered by inclusions (they form a poset), the axioms can be expressed in terms of the order relation on the subobject poset in the category C_T. The category, defined in this way, is appropriately called the categorical semantics for a theory T.

We have thus created (almost automatically!) a domain (the category C_T) the theory T refers to. The internal architecture of the category C_T exactly matches the logic involved in the theory T.

Let us also mention that, *vice versa*, having a (sufficiently rich) category C', we can construct the formal theory T' the logic of which

---

[3] Since f is now regarded as being in C_T rather than in T, it should formally be denoted by a different symbol such as $[f]$, but we omit such formalities for present purposes.

matches the internal architecture of the category C'. This can be done by reading the above definition of the categorical semantics "backwards", i.e. we regard objects of C' as types of T', identity morphisms of C' as variables in T', etc. The theory T', reconstructed in this way from the category C', is called internal logic of C. This entire process can be regarded as a functor, called Lang, from a category of categories, call it CATEGORIES, to a category of theories, call it THEORIES,

$$\text{Lang: CATEGORIES} \rightarrow \text{THEORIES.}$$

For the time being this definition remains informal since neither CATEGORIES nor THEORIES have been properly defined, but it will be done below.

Let us start with a formal theory T. We now want to organize it into a category $\mathrm{Syn}(T)$, called the syntactic category of T. It is done in the following way.

Let Γ be a collection of type assertions, i.e. a collection of rules assigning a type to each term of a given theory, and Φ a collection of all well-defined formulae of T. The pair (Γ, Φ) is called a context. It is a formalization of what in ordinary language one means by this term.

If T is a type theory, its syntactic category, $\mathrm{Syn}(T)$, is defined as follows. Its objects are contexts (Γ, Φ) and its morphisms $(\Gamma, \Phi) \rightarrow (\Delta, \Psi)$ are interpretations (or substitutions) of variables. The latter means that for each type, prescribed by Δ, we must construct an expression of this type out of data contained in Γ. In general, this is done by substituting terms from Γ for variables in Δ. We must also present, for each assumption required by Δ (if there are any), a proof of this assumption from the assumptions contained in Γ (for details see Fu, 2019; nLab, 2020c).

The category $\mathrm{Syn}(T)$, constructed in this way, is also called a category of contexts (for details see Fu, 2019; nLab, 2020b)).

Since from a theory T we have constructed the category $\mathrm{Syn}(T)$, we can have a functor,

$$\mathrm{Syn: THEORIES} \rightarrow \mathrm{CATEGORIES}$$

provided we define the categories THEORIES and CATEGORIES. We do this in the next section.

4. Syntax–Semantics Interaction

Let us start with objects for both of these categories. It is obvious that they will be categories and theories, respectively. To have workable categories, one must restrict the class of theories as candidates of being objects in THEORIES (and analogously for CATEGORIES). The criterion one follows is the kind of logic that underlines a given theory. It could be what logicians call: finite product logic, regular logic, coherent logic, geometric logic, etc.[4] For our further analysis it is irrelevant which one will be chosen. However, for the sake of concreteness we may think about coherent logic. Roughly speaking, this is a fragment of the first order logic which uses only the connectives $\wedge$ and $\vee$, and the existential quantifier. Large parts of mathematics can be formalised with the help of this logic. To this logic there correspond coherent theories and coherent categories. They will constitute objects of THEORIES and CATEGORIES, respectively. Morphisms for CATEGORIES are obviously functors between corresponding categories; for instance coherent functors for coherent categories (nLab,

---

[4] As it could be expected, the internal logic of the corresponding semantic category will be of the corresponding kind, i.e. finite product logic, regular logic, etc. (nLab, 2017).

2011). Let now T_1 and T_2 be objects in THEORIES. A morphism $T_1 \to T_2$ is a functor between their corresponding syntactic theories $\mathrm{Syn}(T_1) \to \mathrm{Syn}(T_2)$. Roughly speaking, this means that it is possible to express (to interpret) T_1 in terms of T_2 (for details and discussion see Halvorson and Tsementzis, 2017).[5]

As a side remark let us notice that by studying the category THEORIES, we could learn "how individual theories sit within it, and how theories are related to each other" (Halvorson and Tsementzis, 2017, p.413). This is nicely consonant with a newer trend in the philosophy of science to investigate the so-called inter-theory relations (Batterman, 2016; Rosaler, 2018).

A truly remarkable fact is that the functors Lang and Syn constitute a pair of adjoint functors. Let us explain precisely what this means.

Let us consider any pair of objects: $\mathcal{C}$ of CATEGORIES and T of THEORIES. Adjoint functors serve to compare them. However, they cannot be compared directly since they live in different categories. Adjoint functors serve to move each of them to the correct category so as to enable the comparison. Let us follow this process step by step (Simmons, 2011, pp.148-153).

Let us first consider the object T which lives in THEORIES. We want to compare it with the object $\mathcal{C}$ which lives in CATEGORIES. We thus move $\mathcal{C}$ to THEORIES with the help of the functor Lang to obtain the object $\mathrm{Lang}(\mathcal{C})$. We now make the comparison with the help of a suitable morphism,

$$f : \mathrm{Lang}(\mathcal{C}) \to T$$

---

[5] Strictly speaking, CATEGORIES is a 2-category (since its objects are categories and morphisms are functors), and THEORIES is a 2-category, in this case, called also a doctrine (nLab, 2020a).

in THEORIES. We do the same starting with $\mathcal{C}$ in CATEGORIES and T in THEORIES, and compare $\mathcal{C}$ with $\mathrm{Syn}(T)$,

$$g : \mathcal{C} \to \mathrm{Syn}(T)$$

in CATEGORIES. To complete the definition of adjunction we demand that morphisms f and g should constitute a pair of bijections which is natural both in $\mathcal{C}$ and T (see below).

The above definition can be put into a concise form

(1) $\mathrm{THEORIES}(\mathrm{Lang}(\mathcal{C}), T) \cong \mathrm{CATEGORIES}(\mathcal{C}, \mathrm{Syn}(T)),$

expressing an isomorphism between the right and left hand sides of this formula that is natural in $\mathcal{C}$ and T. The latter condition says that when $\mathcal{C}$ varies in CATEGORIES and T varies in THEORIES, the isomorphism between morphisms $\mathrm{Lang}(\mathcal{C}) \to T$ in THEORIES and $\mathcal{C} \to \mathrm{Syn}(T)$ in CATEGORIES vary in a way that is compatible with the composition of morphisms in CATEGORIES and THEORIES, correspondingly, and with the actions of Lang and Syn on both these categories (see Awodey and Forssell, 2013; Leinster, 2014, pp.50-51).[6]

We should notice that in the above definition, in fact, we not only compare objects of two different categories, but rather categories themselves (objects $\mathcal{C}$ and T are any pair of objects). Moreover, comparing two categories we are not so much interested in their objects, but rather in morphisms between objects. This is clear from the fact that at the end, we have identified those morphisms of two categories that are pairwise naturally isomorphic among themselves.

---

[6] For a full definition of adjoint functors see any textbook on category theory.

As we can see, categorical logic does not simply create a semantics for a given language, but shows that dependencies between them go both ways: in a sense, syntax and semantics create each other. More precisely, they condition each other through the adjointness relation.

5. Realization and Abstraction

There is another aspect of categorical logic that we shall make use of, and it may be seen as a mathematical description of the processes of abstraction and realization of concepts. The category $\mathrm{Syn}(T)$ representing a theory T may be regarded as presenting a general *concept*, of which the theory T is a particular syntactic description. For example, there is a theory T_{Group} consisting of a single basic type X, and function symbols $* : X \times X \to X$ and $(-)^{-1} : X \to X$ and a constant $u : X$, together with the usual equations for groups as its axioms:

$$x * (y * z) = (x * y) * z$$
$$x * u = x$$
$$u * x = x$$
$$x * x^{-1} = u$$
$$x^{-1} * x = u$$

The syntactic category $\mathrm{Syn}(T_{Group})$ then represents the general concept of a group. This concept can also be represented by another

theory T'_{Group} with a different choice of basic equations, or even a different choice of operations,[7] as long as the resulting categories $\mathrm{Syn}(T_{Group})$ and $\mathrm{Syn}(T'_{Group})$ are equivalent.

A general concept may have many individual *instances*; an instance of the concept of a group is, of course, just a particular group: a set G of elements, equipped with functions interpreting the operations of multiplication and inverse, and satisfying the group equations. A logician would call such an instance a *model of the theory of groups*, but we shall avoid this over-worked term and refer to it instead as a *realization* of the theory of groups. A realization of a theory T in any category $\mathcal{C}$ is essentially the same thing as a functor $\mathrm{Syn}(T) \to \mathcal{C}$ that preserves the relevant structure of the theory—in the case of groups, the finite products $X \times X$. (This is in fact the defining universal property of the syntactic category $\mathrm{Syn}(T)$.) The realizations in the category SET, consisting of all sets and functions, are thus exactly what we called the *instances* of the general concept of a group, namely groups.

The standard category GROUP of all groups and their homomorphisms, as usually defined in abstract algebra, is then essentially the same as the category of all such instances, that is, the category $\mathrm{REAL}(\mathrm{Syn}(T_{Group}), \mathrm{SET})$ of all SET realizations, i.e. (structure-preserving) functors, where the morphisms are just natural transformations of such functors (that these correspond exactly to group homomorphisms is not trivial). In this way, for any general concept $\mathrm{Syn}(T)$ corresponding to a theory T we can define the category of its SET realizations,

$$\mathrm{REAL}(T) =_{df} \mathrm{REAL}(\mathrm{Syn}(T), \mathrm{SET}),$$

---

[7] For example there is an axiomatization of groups using a single ternary operation in place of the two operations $x * y$ and x^{-1}.

which may be viewed as the category of instances of the concept $\mathrm{Syn}(T)$.

Now an amazing and mathematically deep fact emerges, which can only be seen using the tools of categorical logic: from the category $\mathrm{REAL}(T)$ of all instances of the concept presented by T, one can actually recover the general concept $\mathrm{Syn}(T)$. Indeed, for any structured category $\mathcal{R}$ of the same kind as $\mathrm{REAL}(T)$ (we will say a bit more about the condition "of the same kind" below), one can consider all of the *continuous* functors $f : \mathcal{R} \to \mathrm{SET}$; these may be regarded as "images" or "abstractions" of the (generalized) realizations in $\mathcal{R}$. The category of all such abstractions $\mathrm{ABSTRACT}(\mathcal{R}, \mathrm{SET})$ (again, with natural transformations as morphisms) may be called the *abstract* of $\mathcal{R}$, and written simply

$$\mathrm{ABSTRACT}(\mathcal{R}) =_{df} \mathrm{ABSTRACT}(\mathcal{R}, \mathrm{SET}).$$

A similar construction that the reader may know is the ring $\mathcal{C}(X) = \mathcal{C}(X, \mathbb{R})$ of continuous, real-valued functions on a space X. The noteworthy fact that we mentioned above is this: if for $\mathcal{R}$ we take a category $\mathrm{REAL}(T)$ of realizations of a theory T, then the abstract of $\mathrm{REAL}(T)$, consisting of all "abstractions" $\mathrm{REAL}(T) \to \mathrm{SET}$, will be the associated concept $\mathrm{Syn}(T)$.[8] Thus *the abstraction of the realizations of a concept is the concept itself*. We can even summarize this briefly by saying that *All concepts are abstract*, since every concept is the abstraction of its realizations. More generally, for any suitable category $\mathcal{R}$, the category $\mathrm{ABSTRACT}(\mathcal{R})$ of all continuous functors $f : \mathcal{R} \to \mathrm{SET}$ (the "abstractions" of $\mathcal{R}$) is a general concept, of which $\mathcal{R}$ is either the category of realizations, or an approximation thereof.

---

[8] Under suitable assumptions, and up to the relevant notion of equivalence, of course; see (Awodey, 2021) for the general theory.

The general correspondence is given by a (contravariant!) adjunction between the *functors* of Realization and Abstraction which relate these operations; schematically,

$$\text{CONCEPTS} \underset{\text{Abstraction}}{\overset{\text{Realization}}{\rightleftarrows}} \text{INSTANCES}^{op}$$
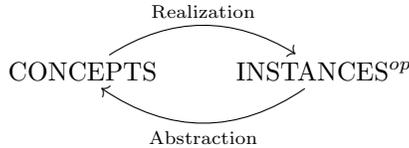
Here CONCEPTS is the category consisting of all "conceptual" categories $\mathrm{Syn}(T)$ and their (relative) "realizations", i.e. functors $\mathrm{Syn}(T) \to \mathrm{Syn}(T')$, and the functor of Realization is defined by taking realizations in SET,

$$\mathrm{Realization}(\mathrm{Syn}(T)) = \mathrm{REAL}(\mathrm{Syn}(T), \mathrm{SET}),$$

which we also called the category of "instances" of the concept.

And INSTANCES is the category consisting of all (generalizaed) categories of instances $\mathcal{R}$ (such as the categories GROUP, RING, etc.) with their "continuous" functors $\mathcal{R} \to \mathcal{R}'$, and the functor of Abstraction is defined by taking continuous functors into SET,

$$\mathrm{Abstraction}(\mathcal{R}) = \mathrm{ABSTRACT}(\mathcal{R}, \mathrm{SET}),$$

which we called "abstractions" of the category $\mathcal{R}$.

Let us consider a simple example! Propositional logic consists of basic propositional variables $x, y, z, \ldots$ and constants $\top, \bot$, which can be made into formulae using the usual propositional connectives $\neg z, x \wedge y, x \vee y, x \Rightarrow z$, and which are assumed to satisfy the usual logical laws, such as $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$, $\neg\neg x = x$, etc. A *theory* T in this simplified case is just a set of propositional letters $V = \{p_1, p_2, \ldots, p_n\}$ (regarded as 0-ary relation symbols), and a list

of propositional formulae $A = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ built up from these letters, as the axioms of the theory. There are no types, typed variables, or function symbols (or rather, there is a single, implicit type 1), and the logical symbols are just the propositional connectives.

The *syntactic category* $\mathrm{Syn}(T)$, representing the "concept", is then the Boolean algebra $F(V)/A$ obtained as the free Boolean algebra $F(V)$ on the variables V as generators, quotiented by the filter generated by the axioms A. This "concept" associated to the propositional theory $T = (V, A)$ is independent of the particular syntactic presentation (V, A). A *realization* of T is then a boolean homomorphism $F(V)/A \to 2$, where $2 = \{0, 1\}$ is the Boolean algebra of truth values. Thus such a realization is just a truth-value assignment to the variables in V, in such a way that the "conditions" in A are all satisfied, i.e. the elements $a \in A$ are all taken to the value "true" (in other words, a "model" of the propositional theory T). For instance, if the theory T is $V = \{x, y\}$ and $A = \{x \vee y, \neg(x \wedge y)\}$, then a realization would be an assignment of x to an actual sentence p, and y to one q, such that only one of p and q is true (or more formally, a direct assignement of such truth values, by-passing the actual sentences). Under our description above, such a realization is an instance of the general concept $F(V)/A$.

Now, such realizations are exactly the points of the *Stone space* $\mathrm{Stone}(F(V)/A)$, the topological space associated to the Boolean algebra $F(V)/A$ under the celebrated Stone duality theorem (Johnstone, 1982)—which is in fact the "propositional logic" case of the categorical logical duality that we are considering here. Formally, the points of $\mathrm{Stone}(F(V)/A)$ are prime filters in $F(V)/A$, and the topology has basic open sets determined by the elements of $F(V)/A$. The Boolean algebra $F(V)/A$ can be recovered from this space $\mathrm{Stone}(F(V)/A)$ as the algebra of continuous functions

$\mathrm{Stone}(F(V)/A) \to 2$ into the discrete space 2, with the pointwise Boolean operations. These *abstractions* of $\mathrm{Stone}(F(V)/A)$ form a Boolean algebra $\mathrm{Bool}(\mathrm{Stone}(F(V)/A))$ which, by Stone duality, is isomorphic to $F(V)/A$,

$$\mathrm{Bool}(\mathrm{Stone}(F(V)/A)) \; \cong \; F(V)/A \, .$$

Indeed, for any (not necessarily Stone) space X, we can form the Boolean algebra $\mathrm{Bool}(X)$ of continuous functions $X \to 2$, and the original space X will then map canonically to $\mathrm{Stone}(\mathrm{Bool}(X))$, giving the "best approximation" of X by a Stone space.

In the general case, in categorical logic we consider many other fragments of logic—propositional, equational, coherent, first-order—and for each such subsystem there is an associated Realization–Abstraction adjunction between theories, and the concepts they represent, on the one hand, and their realizations by instances of these concepts, on the other. The propositional theories just considered give rise to Stone duality (Johnstone, 1982); equational theories (like groups) give rise to Lawvere duality (Adámek, Lawvere and Rosický, 2003); coherent and first-order logic are treated by analogous duality theories developed by Makkai and others (Makkai, 1987; Awodey and Forssell, 2013). In each diffferent case, the associated notion of structured category, structure-preserving functor, continuous functor, etc., is suitably adapted to the respective situation. Many of these logical dualities are discussed from the standpoint of categorical logic in the paper (Awodey, 2021).

6. Categories BRAIN and MIND

So far everything that has been said has merely been a reminder of standard and well-known things. From now on, everything will be hypothetical and highly simplified. The bold and maximally simplified hypothesis is that neurons in the brain can be modeled as categories, the internal logic of which is sufficiently complex (yet manageable). Of course, our inspiring motive is the human brain, and in constructing our model we shall try to imitate what is going in it; however, being conscious of our simplified and highly idealized assumptions, we prefer to speak about a homunculus brain. Our homunculus is a kind of "mathematical robot", the aim of which is to provide us with the opportunity to study how such a simple brain-like structure could "create meanings" out of its purely syntactic program. Our other drastically simplifying assumption consists in systematically ignoring all of the brain's functions and processes that are not directly related to the proposed syntax–semantics relationship.

As it is well-known, neurons communicate through signals transmitted via: presynaptic (source) neuron – axon – synapse – dendrite – postsynaptic (target) neuron, and this *via* is unidirectional. In our homunculus model, these transmission processes will be regarded as functors between categories (neurons).

Let us consider the category CATEGORIES, which we now aptly call BRAIN. Its objects are categories modeling neurons, and morphisms are functors between these categories.

We thus assume that each neuron in the homunculus brain is represented by a category (belonging to a certain class of categories; in the following we shall simply say that a neuron is a category). At the moment, we are not interested which biological mechanisms implement this assumption. Everything that counts in this model is

the assumption that neurons consist of collections of objects and morphisms satisfying conditions from the category definition. We should have in mind that these simple conditions might lead to highly complicated structures.

Morphisms (arrows) in the category CATEGORIES are functors between object-categories, that is to say axons through which neurons communicate with each other. The crucial thing is that they must satisfy the usual conditions for morphisms: composition of morphisms, its associativity, the existence of identity morphisms. With the latter there is no problem: no output from a neuron counts as its identity morphism. To check whether two other conditions are verified in the human brain would require going deeper into the neural structure of our brain. In the case of the homunculus brain, this is not necessary. Since the homunculus is of our construction, we simply assume that synapses in its brain well-compose and do so in the associative way.

The next step seems obvious. Each neuron (modeled as a category $C \in$ BRAIN) has its own program enabling its working, i.e. an internal logic underlying this program. We thus can define a counterpart of Lang(C) which is a "theory" of this neuron. It is reasonable to claim that it is an object of the category THEORIES which we now call MIND, and the functor Lang: BRAIN $\rightarrow$ MIND is defined in analogy to that between CATEGORIES and THEORIES.

What about the morphisms between such objects? We proceed in strict analogy with what has been done in THEORIES. Let now T_1 and T_2 be objects in MIND, a morphism between them, $T_1, \rightarrow T_2$, is a functor between their corresponding syntactic theories, i.e. $\mathrm{Syn}(T_1) \rightarrow \mathrm{Syn}(T_2)$, where the functor Syn: MIND $\rightarrow$ BRAIN is defined in analogy to that between THEORIES and CATEGORIES.

The analogy is only apparently straightforward. In fact, it is based on a huge extrapolation, and as such highly hypothetical, but it is

worth exploring it since the problem at stake deserves even a higher risk. By pursuing this analogy we could claim that also in this case the functors Lang and Syn are adjoint functors. If so, we have a very interesting conjunction between brain and mind; it is interesting even if brain and mind are modeled by such a naive construction.

Neurons, their interactions and programs underlying their working are, in contrast with abstract categories like CATEGORIES and THEORIES, real things, at least in the homunculus world, and we are entitled to suppose that the functors Lang and Syn between Brain and Mind really do what they formally signify (like our phase portrait on the computer screen really did what the program told it to do).

Roughly speaking the functor Lang provides a collection of theories (mind) for a collection of neurons (brain), and the functor Syn transfers the syntax of these theories to the network of neurons. The action of these two functors is adjoint; consequently it determines a strict interaction between BRAIN and MIND. Let C be any object (a neuron) in BRAIN and T any object (the theory of this neuron) in MIND, then equation (1) assumes the form

$$(2) \qquad \mathrm{MIND}(\mathrm{Lang}(\mathcal{C}), T) \cong \mathrm{BRAIN}(\mathcal{C}, \mathrm{Syn}(T)).$$

The natural isomorphism $\cong$ appearing in this equation is crucial. It states that when we go from neuron to neuron as objects in BRAIN, and their corresponding theories vary in THEORIES, then the isomorphism between morphisms $\mathrm{Lang}(\mathcal{C}) \to T$ in MIND and $\mathcal{C} \to \mathrm{Syn}(T)$ in BRAIN varies in a way that is compatible with the composition of morphisms in BRAIN and MIND, correspondingly, and with the actions of the functors Lang and Syn (see Awodey and Forssell, 2013; Leinster, 2014).[9] Finally, the "higher" cognitive functions of abstraction

---

[9] For a full discussion of the role of the naturality condition in the definition of adjoint functors see any textbook on category theory.

and realization of concepts are modelled by a corresponding adjunction between the associated functors $\mathrm{Abstraction}$ and $\mathrm{Realization}$ relating these categories BRAIN and MIND. We could summarise the situation by saying that the categories BRAIN and MIND interact with each other with their entire structures and, at the same time, these very structures are shaped by this interaction.

7. A Comment

The interactions between syntax and semantics are omnipresent both in our everyday conversations and in various forms of practicing science. The world around us is full of meanings and our attempts to decipher them. Science could be regarded as a machine to produce signs, through experimentation and critical reasoning, and extracting from combinations of them information about the structure of the world. Logicians put a lot of effort to make the syntax–semantics interaction precise. As we have seen in section 2, despite the fact that formal languages are lacking any external references, it was possible to create semantical references for them by cleverly exploiting the relation between language and its metalanguage. In categorical logic the situation has improved. Any formal theory T generates via the functor Syn the category $\mathrm{Syn}(C) = C_T$ of which it is a theory, i.e. C_T provides a "natural" semantics for T. And *vice versa*, any (sufficiently rich) category C', via the functor Lang, generates its own theory $\mathrm{Lang}(C') = T'_{C'}$ which constitutes the internal logic of C'. It is interesting to notice that T_{C_T} does not coincide with T, they are only Morita equivalent. Here, we shall not go into technical details; it is enough to say that two Morita equivalent theories could be regarded as two interpretations of the same theory (Halvorson, 2016).

The fact that T_{C_T} does not coincide with T is a consequence of the fact that the functors Lan and Syn are not mutually inverse functors but constitute a pair of adjoint functors. This in turn implies that in categorical logic the interaction between syntax and semantics is skillfully complex, with creative influences coming both ways.

All the above discussed properties of the syntax–semantics interaction can be presumed to be preserved if applied to the categories BRAIN and MIND. There is only one big difference: now "neurons and their theories" are real things (although in a highly idealised, toy version in the homunculus world). Nevertheless, the situation is not so different from the one which we can observe in many empirical sciences, in which some abstract mathematical structures model some real processes (always more or less idealised). We should not be surprised that the method of mathematical modeling works when applied to our cognitive processes, but rather that mathematical structures not only describe the real world (whether it is our brain or the world of physics), but that they are also effectively acting in it (like in the little arrow on the computer screen).

Bibliography

Adámek, J., Lawvere, F. and Rosický, J., 2003. On the duality between varieties and algebraic theories. *Algebra Universalis* [Online], 49(1), pp.35–49. Available at: https://doi.org/10.1007/s000120300002 [visited on 3 November 2020].

Awodey, S., 2021. Sheaf Representations and Duality in Logic. In: C. Casadio and P.J. Scott, eds. *Joachim Lambek: The Interplay of Mathematics, Logic, and Linguistics*. Springer International Publishing.

Awodey, S. and Forssell, H., 2013. First-order logical duality. *Annals of Pure and Applied Logic* [Online], 164(3), pp.319–348. Available at: https://doi.org/10.1016/j.apal.2012.10.016 [visited on 3 November 2020].

Batterman, R., 2016. Intertheory Relations in Physics. In: E.N. Zalta, ed. *The Stanford Encyclopedia of Philosophy (Fall 2016 Edition)* [Online]. Stanford: Metaphysics Research Lab, Stanford University. Available at: https://plato.stanford.edu/archives/fall2016/entries/physics-interrelate/ [visited on 3 November 2020].

Borceux, F., 1994. *Handbook of Categorical Algebra. 3: Categories of Sheaves*, *Encyclopedia of Mathematics and its Applications* vol. 52. Cambridge: Cambridge University Press.

Ehresmann, A., 2017. Applications of Categories to Biology and Cognition. In: E. Landry, ed. *Categories for the Working Philosopher*. Oxford, New York: Oxford University Press, pp.358–380.

Ellerman, D., 2015. On Adjoint and Brain Functors. *arXiv:1508.04036 [math]* [Online]. Available at: http://arxiv.org/abs/1508.04036 [visited on 3 November 2020].

Fu, Y., 2019. *Category Theory, Topos and Logic: A Quick Glance* [Online]. Available at: http:%20//charlesfu.me/repository/topos.pdf [visited on 28 February 2019].

Gómez, J. and Sanz, R., 2009. *Modeling Cognitive Systems with Category Theory: Towards Rigor in Cognitive Sciences* [Online]. Autonomous Systems Laboratory, Universidad Politécnica de Madrid. Available at: http://www.aslab.org/documents/controlled/ASLAB-A-2009-014.pdf [visited on 4 November 2020].

Halvorson, H., 2016. Scientific Theories. In: P. Humphreys, ed. *Oxford Handbook of Philosophy of Science* [Online]. Vol. 1. Oxford: Oxford University Press, pp.402–429. Available at: https://doi.org/10.1093/oxfordhb/9780199368815.013.33 [visited on 3 November 2020].

Halvorson, H. and Tsementzis, D., 2017. Categories of Scientific Theories. In: E. Landry, ed. *Categories for the Working Philosopher*. Oxford, New York: Oxford University Press, pp.402–429.

Healy, M.J. and Caudell, T.P., 2006. Ontologies and Worlds in Category Theory: Implications for Neural Systems. *Axiomathes* [Online], 16(1), pp.165–214. Available at: https://doi.org/10.1007/s10516-005-5474-1 [visited on 3 November 2020].

Hodges, W., 2018. Tarski's Truth Definitions. In: E.N. Zalta, ed. *The Stanford Encyclopedia of Philosophy (Fall 20218 Edition)* [Online]. Stanford: Metaphysics Research Lab, Stanford University. Available at: https://plato.stanford.edu/%20archives/fall2018/entries/tarski-truth/ [visited on 3 November 2020].

Johnstone, P.T., 1982. *Stone Spaces*, *Cambridge studies in advanced mathematics* 3. Cambridge: Cambridge University Press.

Koch, C., 1997. Computation and the single neuron. *Nature* [Online], 385(6613), pp.207–210. Available at: https://doi.org/10.1038/385207a0 [visited on 3 November 2020].

Leinster, T., 2014. *Basic Category Theory*, *Cambridge Studies in Advanced Mathematics* 143. Cambridge: Cambridge University Press.

Mac Lane, S. and Moerdijk, I., 1992. *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*, *Universitext*. New York [etc.]: Springer-Verlag.

Makkai, M., 1987. Stone duality for first order logic. *Advances in Mathematics*, 65, pp.97–170.

McCulloch, W.S. and Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics* [Online], 5(4), pp.115–133. Available at: https://doi.org/10.1007/BF02478259 [visited on 3 November 2020].

Mizraji, E. and Lin, J., 2011. Logic in a Dynamic Brain. *Bulletin of Mathematical Biology* [Online], 73(2), pp.373–397. Available at: https://doi.org/10.1007/s11538-010-9561-0 [visited on 3 November 2020].

nLab, 2011. *Coherent Functor*. Available at: https://ncatlab.org/nlab/show/coherent+functor [visited on 3 November 2020].

nLab, 2017. *Internal Logic* [Online]. Available at: https://ncatlab.org/nlab/show/internal+logic [visited on 3 November 2020].

nLab, 2020a. *Doctrine* [Online]. Available at: https://ncatlab.org/nlab/show/coherent+functor [visited on 3 November 2020].

nLab, 2020b. *Substitution* [Online]. Available at: https://ncatlab.org/nlab/show/substitution [visited on 3 November 2020].

nLab, 2020c. *Syntactic Category* [Online]. Available at: https://ncatlab.org/nlab/show/syntactic+category [visited on 3 November 2020].

Rosaler, J., 2018. Inter-Theory Relations in Physics: Case Studies from Quantum Mechanics and Quantum Field Theory. *arXiv:1802.09350 [quant-ph]* [Online]. Available at: http://arxiv.org/abs/1802.09350 [visited on 3 November 2020].

Rosen, R., 1985. Organisms as Causal Systems Which Are Not Mechanisms: An Essay into the Nature of Complexity. *Theoretical Biology and Complexity: Three Essays on the Natural Philosophy of Complex Systems* [Online]. New York [etc.]: Academic Press, pp.165–203. Available at: https://doi.org/10.1016/B978-0-12-597280-2.50008-8 [visited on 3 November 2020].

Sher, G., 1999. What is Tarski's Theory of Truth? *Topoi* [Online], 18(2), pp.149–166. Available at: https://doi.org/10.1023/A:1006246322119 [visited on 3 November 2020].

Simmons, H., 2011. *An Introduction to Category Theory*. Cambridge [etc.]: Cambridge University Press.

Tarski, A., 1933. *Pojęcie prawdy w językach nauk dedukcyjnych*, *Prace Towarzystwa Naukowego Warszawskiego. Wydział III: Nauk Matematyczno-Fizycznych* 34. Warszawa: nakładem Towarzystwa Naukowego Warszawskiego, z zasiłku Ministerstwa Wyznań Religijnych i Oświecenia Publicznego.

Tsuchiya, N., Taguchi, S. and Saigo, H., 2016. Using category theory to assess the relationship between consciousness and integrated information theory. *Neuroscience Research* [Online], 107, pp.1–7. Available at: https://doi.org/10.1016/j.neures.2015.12.007 [visited on 3 November 2020].

Woszczyna, A. and Heller, M., 1990. Is a horizon-free cosmology possible? *General Relativity and Gravitation* [Online], 22(12), pp.1367–1386. Available at: https://doi.org/10.1007/BF00756836 [visited on 3 November 2020].

Book reviews

Category Theory in the hands of physicists, mathematicians, and philosophers

Category Theory in Physics, Mathematics, and Philosophy, Kuś M., Skowron B. (eds.), Springer Proc. Phys. 235, 2019, pp.xii+134.

The book under review is a result of the conference under the same title which was held in Warsaw, Poland in November 2017. As far as I know, it was the first conference in Poland's history on the subject of category theory (CT) and its applications. Warsaw is an excellent place for such an event if only because of its traditions of the Lvov–Warsaw School, but also being the place where Samuel Eilenberg (one of the founding fathers of CT) was born and defended his Ph.D. thesis. The majority of the organizers and speakers were Polish. Nevertheless the significance and impact of this conference, mainly thanks to the proceedings, is definitely international. The present volume of ZFN, which itself is devoted to the conference proceedings of the next great event of this kind, shows that the tradition of Polish conferences on the subject of CT and its applications emerges. It could only be wished it would thrive in the future.

The publication consists of 10 chapters, which are all separate articles, the majority of which are papers presented at the conference. The subjects of the papers are very diverse, which itself evinces how vast are the applications of CT to different areas of knowledge.[1] The reader not familiar with at least the basics of CT will be much hindered in the study of them.[2] However, not all the papers hinge on the technicalities of CT (see the first four chapters). Others also require advanced knowledge (mainly of physics, for chapters 7 to 9).

The book begins with a few pages of introduction, which presents the work's context and succinctly discusses each chapter. The first article, *Why Categories?*, written by M. Kuś, B. Skowron, and

---

[1] The reader interested in the subject of broad applications of CT may benefit also from e.g. the great book (Awodey, 2010) (ZFN published its Polish review by Michael Heller (2018)).

[2] Textbooks in CT that may be consulted are e.g. (Awodey, 2010; Smith, 2018; McLarty, 1995).

K. Wójtowicz, is also a kind of intro-duction to the whole book, but from a different perspective. It poses a question (and attempts to answer it, at least partially) about reasons for such a significant increase in pop-ularity of CT and its applications in recent years. The great diversity of the following chapters is a good, though still only partial, example of this phenomenon. This paper presents some aspects of the origins of CT, its relations with, and possi-ble applications to philosophy and physics. Within applications to phi-losophy, it discusses structuralism in the philosophy of mathematics, foundations of mathematics, unity of mathematics, and metaphysics. In connection with physics, it ex-amines mainly quantum mechan-ics and its ontological foundations. Nonetheless, the text is more than a review. The authors, when dis-cussing the above mentioned dif-ferent applications of CT, also give examples of their main claim: "cat-egory theory is a formal ontology that captures the relational aspects of the given domain in question" (p.1).[3] These examples show how the formal-ontological shift brought by CT, meant as the shift from the "standard and natural attitude to-wards the objects, as if they were individual subjects of properties [. . .] [to] a form of pure relational-ity" (p.18), sheds some new light on the old problems.

The next chapter, *Category The-ory and Philosophy*, by Z. Król, deals with relations between CT and philosophy. One of the topics addressed in this article, although not the main one, is how set the-ory (ST) and CT are similar and dif-ferent. The author considers them more broadly than just the formal theories: "CT and ST are not singu-lar formal theories, but rather open domains accompanied by the rel-evant methods and styles of con-sideration, together with some ba-sic concepts which can be investi-gated within many different formal theories" (p.22). The main part of this article is a case study of cer-tain classical problems in philoso-phy and the question of the useful-ness of CT in dealing with them. The study is rather general without deeper analysis, albeit the author gives some details and references to literature. It starts with remarks on the great influence of mathematics—in general—upon philosophy in its

---

[3] References with no information except the page number refer to the book under review.

history. The studied cases concern some issues in ontology (such as monism vs pluralism, or mathematical Platonism) and epistemology.

K. Wójtowicz in his paper, *Are There Category-Theoretical Explanations of Physical Phenomena?* (chapter 3), addresses a question of mathematical explanation in science. He asks especially whether CT may give explanations of physical phenomena. To answer this question he first analyzes the general problem of the explanatory role of mathematics in physics, assuming (as a working hypothesis) that "there are genuine mathematical explanations in science" (p.37). He admits that "it cannot be denied that CT contributes to our understanding of physics" (p.33) and gives some illustrations for this, mainly within the Topos Quantum Theory. However, after noting first that "if a model has no predictive potential at all, it is doubtful whether it really has explanatory character" (p.38), his central claim is that "in this sense there are no category-theoretical explanations of physical phenomena, in spite of there being mathematical explanations" (p.38). The author considers the contribution of CT to be "on the metatheoretical rather than the theoretical level" (p.40). He eventually writes: "CT offers 'meta-abstract explanations'"

(p.42). I have the feeling that this article underestimates rather than overestimates even the hitherto relevance of CT to other disciplines. Besides, maybe the future will show more.

Chapter 4, entitled *The Application of Category Theory to Epistemic and Poietic Processes*, by J. Lubacz, tries to explore the possibility of applying CT and its notional framework to "the analysis and monitoring of progress in the unfolding of [. . .] processes," (p.45) such as acquiring knowledge and some activity that results in the creation of artefacts. These are called epistemic and poietic processes, respectively. The style of this paper is definitely philosophical. After the introduction, the author presents some considerations about the processes in question in the broader context of philosophy as well as presenting his own proposition of the possible "conceptual structure of the pattern and dynamicity of epistemic and poietic processes" (p.48). Next, the author ponders the potential application of CT to epistemic and poietic processes. He rightly notes that "it must be clearly stated that the apparatus of CT can only be employed for those conceptual components of epistemic and poietic processes which are expressible in some for-

mal language and form" (p.50). He then develops his own proposal of the possible use of CT in this context, which is, however, rather general and preliminary in nature.

In his paper, *Asymmetry of Cantorian Mathematics from a Categorial Standpoint: Is It Related to the Direction of Time?* (chapter 5), Z. Semadeni addresses an interesting feature of the *Cantorian Mathematics*. In his approach, *Cantorian Mathematics* is referring to "basic mathematical structures of algebra, topology, functional analysis etc. expressed in terms of set theory, as they were conceived prior to the emergence of category theory, i.e., by the middle of the 20th century" (p.55f). The author notes and explains briefly that CT, in a certain sense (with respect to the reversal of arrows), is symmetric. However, when we consider some categories with objects being structures taken from *Cantorian Mathematics*, then the products and coproducts of these objects (separately in each category) each have a different 'style.' Namely, while almost all examples considered (and the author gives quite a few of them) of products turn out to be "the cartesian product endowed with a suitable structure" (p.57) (and the remaining two

examples can be in some way cured or revised alike), coproducts, however, "may be markedly different from each other" (p.57). For some categories coproducts "are based on the same construction, namely on the disjoint union" (p.57), but "in other categories coproducts may differ basically" (p.57). As an example, one of many the author gives, let me note the category **Grp** (of groups and their homomorphisms) in which the coproduct is the free product of groups, whereas in its full subcategory of abelian groups the coproduct is the (external) direct sum of groups. The author notes that similar asymmetry concerns also equalizers and coequalizers ("albeit in a much milder form" (p.59)), and other limits and colimits. At this point, he poses a philosophical question: "What features of *Cantorian Mathematics* lie behind this asymmetry?" (p.60), and suggests that it follows "from the asymmetry of many-to-one relationship in the notion of a function $f : X \rightarrow Y$" (p.60). In the last two paragraphs, we find an interesting discussion showing that the asymmetry between products and coproducts changes, so to speak, direction, when instead of considering examples of an "algebraic" nature

one considers structures of "coalgebraic" nature connected with one-to-many relation ("mostly stimulated by Computer Science" (p.61)).

The next article, entitled *Extending List's Levels*, by N. Dewar, S.C. Fletcher, and L. Hudetz is an interesting extension and modification of the unified framework for modeling different types of levels (descriptive, explanatory and ontological) proposed by C. List (2019). The paper is well written—particular notions are clearly introduced and commented on and occasionally important and helpful examples are given. First, the authors succinctly review List's approach and correct a minor defect. Next, they analyze the relationship between supervenience and reduction in this setting. In general, supervenience does not entail reduction (and the reader is familiarized through a simple example), but the authors have shown that if the levels and supervenience maps fulfill certain additional conditions (they have to be compatible and jointly characterizable, the notions being defined and commented in the text), then supervenience does entail reduction. They note, moreover, that "in many cases of supervenience between scientific levels of description, this [fulfilling the

above-mentioned condition(s)] can be expected. So it is quite plausible that in many cases of interest, supervenience and reduction of levels go hand in hand" (p.79). After these considerations, the authors propose two extensions of List's framework: from supervenience maps treated as (total) functions to partial maps, and from surjective maps to non-surjective ones. These generalizations open new possibilities, described in the text. Subsequently, the authors move on to the most important, in my opinion, part of their work, namely to the modification of List's framework which involves considering levels not only as elements of a certain poset (or objects of a posetal category) but as categories of structures (more precisely, they suggest "to represent a level of description, L, as a pair $\langle \mathbf{L}, \Omega \rangle$ consisting of a description language, $\mathbf{L}$, and a category, Ω, of $\mathbf{L}$-structures" (p.73)). This means that in order to specify a level of description "one does not only specify its structures but also the morphisms (admissible transformations) between these structures" (p.73). The choice of morphisms "reflects which expressions of $\mathbf{L}$ are taken to be meaningful within the level L" (p.73). Moreover, in this setting superve-

nience relations between the levels are viewed as functors. Another extension of the original framework is made by allowing "all sorts of functors to be included in a system of levels of description" (p.77), rather than just the supervenience functors. The authors give also some examples involving these generalizations and note that the extended framework better serves for a philosophy of science in general. This last generalization (of treating levels as categories) is far-reaching, as "taking levels as categories themselves demands a more robust use of categorial ideas [such as natural transformations, adjunctions, and others] that could also prove to be more fruitful" (p.79).

The next three chapters deal with applications of CT to physics. The first of them, written by K. Bielas and J. Król, titled *From Quantum-Mechanical Lattice of Projections to Smooth Structure of* $\mathbb{R}^4$, uses CT to relate quantum algebra structure with the smooth structure of spacetime. The paper is a work in progress and is connected with some earlier article by the authors and another scholar, namely Król et al. (2017). After the introduction, in which the authors already bring in some key concepts,

they outline some quantum mechanical preliminaries, i.a. introducing a complete orthomodular lattice of projections on a Hilbert space associated with the initial quantum system. This lattice (denoted as $\mathbb{L}$) is an algebraic basis of the so-called quantum logic. As is well known, generally (whenever dim $\mathscr{H} > 2$) $\mathbb{L}$ is not Boolean, which means that logic of a quantum system defined in this way is not classical. In order to get the connection with the classical world the authors, in the next section, look at the subalgebras of $\mathbb{L}$ which are Boolean. Each such Boolean subalgebra, they argue, is in a certain way "to be considered as a local, classical frame of reference for a quantum system" (p.87). Subsequently, they show a way to construct an orthomodular lattice from its Boolean subalgebras as a suitable colimit in the category of so-called partial Boolean algebras and appropriate homomorphisms (one has to do it in this larger category, which extends the category of orthomodular lattices and lattice homomorphisms). Then it is shown, in the next section, how one can arrive at the smooth manifold by means of some categorical constructions, namely the authors conclude that "given a smooth manifold, it is al-

ways the colimit of its atlas" (p.89).[4] Finally, the authors try to relate the quantum structure with the differential (smooth) structure of the manifold. However, it is still an open problem if certain correspondence has a functorial character. In the *Discussion* section some further considerations are addressed, i.a. the cardinality of the smooth atlas of exotic R^4.

Differential structure of a manifold, or rather its enrichment in terms of a so-called formal manifold is a key tool of the next paper, *Beyond the Space-Time Boundary*, by M. Heller and J. Król. The article is quite technical, preparatory for further research, but also fundamental and opening a truly intriguing path for new explorations.[5] In order to "cross the boundary" of a singular spacetime, they use the so-called Synthetic Differential Geometry (SDG), which is a categorical version of standard differential geometry and is based on intuition-

istic logic. After the introduction, the authors offer many new notions (known in the literature) needed for further considerations and their model in the next sections. Among others, one considers various kinds of infinitesimals, which enrich the standard structure of $\mathbb{R}$ (or a manifold in general) and make differentiation a purely algebraic operation. The authors give a nice image: "We may imagine that they [infinitesimals] constitute the entire world inside every point of $\mathbb{R}$, a sort of a fiber over $x \in \mathbb{R}$" (p.96). The reader not familiar with such infinitesimals may find the definition such as "$D = \{x \in R | x^2 = 0\}$" (p.97) astonishing. Let me only note that the intuitionistic logic plays one of the key roles here. In SDG such an object D obviously does not reduce to $\{0\}$. It comprises so-called nilsquare infinitesimals, which are different, but at the same time indistinguishable, from 0.[6] The whole collection of (countably many) differ-

---

[4] The reader might find it helpful to note that there is a certain confusion with the notation: on p.88, V_i is a subset of $\mathbb{R}^n$, whereas in the Corollary 2, on p.89, V_i is a subset of the manifold. It would be more natural to denote on p.89, in accordance with the notation from p.88, V_i as U_i, and W_i as V_i.

[5] The interested reader may consult also other papers in this subject by these authors, e.g. see (Heller and Król, 2016; 2017).

[6] In the intuitionistic logic, double negation does not, in general, imply identity (double negation elimination is not a theorem), so not being different from 0 (not being not equal 0) does not imply being identical to zero or, in other words, indistinguishability does not, in general, imply identity.

ent kinds of infinitesimals serves to distinguish various kinds of neighbourhoods, which in turn allow us to define various kinds of the so-called monads. In the proposed model, for example when thinking about the evolution of the universe back in time, these monads, which are ordered (relative to each other) and represent various kinds of differentiability properties, provide suitable structures to describe the evolution after crossing the singular boundary (which from the physical point of view might possibly be identified with Planck's threshold). The authors note that the model "does not pretend to describe the actual evolution of the universe. At its present stage of development, it is nothing more than a toy model" (p.96). The toy models, however, can play a surprisingly significant role in the development of more realistic physical models! Two appendices give more details, for example about the way infinitesimal spaces may appear in the model.

The methods of SDG are also the main mathematical tools used in the next paper in application to the pursuit of quantum gravity (QG). J. Król in his article *Aspects of Perturbative Quantum Gravity on Synthetic Spacetimes* gives his own con-

tribution to this huge endeavor. In the introduction, the author presents some of the results in QG obtained so far. He follows the known perturbative approach to QG (which involves an expansion of the metric tensor around the flat Minkowski spacetime and an attempt to quantize the fluctuations around it) but tries to use new methods in the context of SDG. As he notes, "The particularly well-defined area of applicability of synthetic methods will appear to be perturbative quantum gravity" (p.107). Some definitions and notions are common with the previous chapter (which J. Król coauthored). The reader has to be familiar with some background knowledge and notations from the literature in order to fully follow the contents of the paper. I will not bring up the details of the article, which comprises much more considerations than are mentioned here. I only note that the author considers both the case of pure gravity (no matter fields) and the case with matter. Synthetic methods are applied in the context of the so-called BCFW procedure (from the names of the authors of (Britto et al., 2005), also see Appendix B of the paper by Król) and lead, *inter alia*, to the following results: (i) "The pure perturbative

covariant QG on synthetic space-time is entirely finite theory [...] [and] $g_{\mu\nu}$ can now be considered as fully quantized field" (p.112); (ii) "supergravity theory [for $\mathcal{N} = 8$ and with matter fields] formulated on synthetic spacetime is again finite and $g_{\mu\nu}$ quantized" (p.112). In the last section the author also reflects on the question, "how (if at all) it can be that a perturbative QG is considered a fundamental theory" (p.113).

The last chapter, titled *Category Theory as a Foundation for the Concept Analysis of Complex Systems and Time Series*, by G.N. Nop, A.B. Romanowska and J.D.H. Smith, deals with the applications of CT to so-called concept analysis. The paper extends and generalizes the existing applications. Concept analysis studies relationships between some items and properties they possess. After the short introduction, the authors give a summary of the notions used in this context. Although they are adopted in various fields of study (and may have different names), they "provide essentially equivalent tools for the analysis of a static system functioning at a single level" (p.120). To get some taste of the paper let me only mention one special notion, that of

a concept of a certain context. A context is a triple made of a set of items (Ω), a set of properties (Π), and a relation between them attributing properties to items (being a relation it is a subset of $\Omega \times \Pi$). A concept of such a context is an ordered pair (A, B) (denoted in the text as $(A|B)$), where A is a subset of items and B is a subset of properties, such that the set of properties common to all items in A is exactly B and at the same time the set of items attributed to all properties in B is exactly A. I think one may say that in this way we get a certain characterization of some items in terms of properties (and vice versa). In general, the power sets of Ω and Π (treated as poset categories) are connected with certain adjoint functors known as a Galois connection. In the paper, the reader is familiarized with many other notions and some simple examples as well. Then the authors proceed to "present the foundations for an extension of the ideas of concept analysis to changing environments, evolving complex systems, and time series analyses of successive stages of a given system" (p.119), which is the main goal of the paper. The extension to complex systems is done, generally speaking, in terms of certain functors from a

semilattice (which serves as a kind of index of distinct levels or, if it is a chain, of time points) to appropriate categories, though the procedure is more sophisticated.

In sum, the book under review is a testimony to the breadth of applications of CT in physics, mathematics, and philosophy. The book is worth reading, in general, although the chapters are so diverse (in the subject matter and sometimes also in the scientific value) that I suppose the reader that is not interested in all of the interdisciplinary applications should focus on the part (s)he is interested in; e.g. chapters 7–9 may be truly engaging for readers familiar with general relativity and quantum mechanics but may be completely incomprehensible for non-physicists. The papers contained in the book are not review articles nor popular science, but are mostly conference papers, which makes the reading much more demanding (recommended for graduate-level readers and beyond) though at the same time much more fascinating.

Mariusz Stopa

Awodey, S., 2010. *Category Theory*. 2nd ed., *Oxford Logic Guides* 52. Oxford; New York: Oxford University Press.

Britto, R., Cachazo, F., Feng, B. and Witten, E., 2005. Direct proof of tree-level recursion relation in Yang-Mills theory. *Phys. Rev. Lett.* [Online], 94, p.181602. Available at: arXiv: hep - th / 0501052. Available at: https:// doi.org/10.1103/PhysRevLett. 94.181602.

Heller, M. and Król, J., 2016. How Logic Interacts with Geometry: Infinitesimal Curvature of Categorical Spaces. Available at: arXiv: 1605.03099 [math.DG].

Heller, M. and Król, J., 2017. Infinitesimal Structure of Singularities. *Universe* [Online], 3(1), p.16. Available at: https://doi.org/10.3390/universe3010016.

Heller, M., 2018. Filozoficznie prowokująca teoria kategorii. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (65), pp.232–241. Available at: https://zfn.edu.pl/index.php/zfn/article/view/451 [visited on 24 July 2020].

Król, J., Asselmeyer-Maluga, T., Bielas, K. and Klimasara, P., 2017. From Quantum to Cosmological Regime. The Role of Forcing and Exotic 4-Smoothness. *Universe* [Online], 3(2), p.31. Available at: https://doi.org/10.3390/universe3020031.

List, C., 2019. Levels: Descriptive, Explanatory, and Ontological.

Noûs [Online], 53(4), pp.852–883. Available at: https://doi.org/10.1111/nous.12241.

McLarty, C., 1995. *Elementary Categories, Elementary Toposes*. Oxford: Oxford University Press.

Smith, P., 2018. *Category Theory: A Gentle Introduction* [Online]. University of Cambridge, lecture notes. Available at: http://www.logicmatters.net/resources/pdfs/GentleIntro.pdf [visited on 24 July 2020].

Contemporary Polish ontology. Where it is and where it is going

The *Contemporary Polish Ontology* is a collection of recent works by Polish ontologists. The volume includes thirteen papers and introduction as well as discussion sections. The authors published in the volume are associated with the International Center for Formal Ontology (ICFO) at the Faculty of Administration and Social Sciences, Warsaw University of Technology, and they are affiliated with several research centers across Poland. The articles in the book therefore reflect, to some degree, the state of current ontological research in Poland.

The volume is not intended to be read from cover to cover, as it includes a diverse collection of topics united only by a common ontological vantage point. Few people, I suspect, would be interested in reading each chapter. So, how should we approach this book, and why might someone want to read it?

It seems there may be two reasons for doing this and two ways to do it. For example, someone may want a primer on the state of ontological research in Poland, such as who is doing what, what topics are being investigated, and what questions are being posed. Alternatively, someone may be interested in a specific topic and what has been done in that area. If you fall into the first group, you should read the introduction written by Bartłomiej Skowron and the final chapter, *An Assessment of Contemporary Polish Ontology* (Skowron, 2020, pp.271–294), which has been coauthored by several researchers, but particularly the section written by Skowron. This section provides a succinct yet comprehensive review of the state of Polish ontology, together with some added historical background. If you want to then go further, you may find specific topics that pique your interest by reading the introduction before diving deeper into the volume. If you fall into the second group, namely being interested in just a specific topic, or are concerned with specific ontological questions, you should start with the introduction and then establish which chapters are of most interest to you.

So, what topics are presented in the volume? Tomasz Bigaj dis-

cusses the concept of symmetry in structures, proposing its redefinition and indicating its possible impact on Quantum Theory. Mariusz Grygianiec, meanwhile, analyzes the possibility of integrating the concept of personal identity into animalism and that of the Simple View. Next, Filip Kobiela analyzes the concept of the present and Ingarden's take on this, thus formulating the "outline of the ontological theory of relativity of the duration of the present." Zbigniew Król and Józef Lubacz then take over by exploring the epistemic conditions for the knowledge of existence, suggesting how this type of knowledge may be obtained in the conscious subject. Andrzej Biłat then attempts to formulate the classical conception of philosophy, which can be understood as the philosophy of Plato and Aristotle, and shows its compatibility with contemporary logic and science. Urszula Wybraniec-Skardowska, meanwhile, tries to formulate the ontology of language, where language has a particular ontological status. Krzysztof Śleziński explores Bornstein's concepts of general ontology and shows how his ideas live on in the current research into spatial logic. Next, Janusz Kaczmarek explores the concepts of topological ontology and

their relation to Leibniz's *Monadology*. Coming back to Krzysztof Wójtowicz, he discusses the necessary conditions for the existence of mathematical objects and demonstrates the import of this discussion to the realism–antirealism debate in the philosophy of mathematics. Rafał Urbaniak, meanwhile, reviews approaches to the formulation of mathematical theories, focusing on the status of neologicism. Jacek Paśniczek explores the application of an algebraic framework based on the De Morgan lattice for the representation of the ontological features of possible worlds and situations. Next, Marek Magdziak presents a logical study on the concepts pertaining to the notion of action, while Michał Głowala endeavors to demonstrate how some conceptual tools of scholastic philosophy can be helpful in resolving the current debates about the ontologies of intentionality and powers.

All in all, the *Contemporary Polish Ontology* is certainly worthy of attention. It provides a panorama of the various ontological topics that are currently being pursued in Poland. Anyone with an interest in ontological questions will likely find an engaging paper. For those interested in the state of Polish phi-

losophy, meanwhile, the discussion in the final chapter, *An Assessment of Contemporary Polish Ontology*, is worthy of particular attention. This discussion—which has been coauthored by Bartłomiej Skowron, Tomasz Bigaj, Arkadiusz Chrudzimski, Michał Głowala, Zbigniew Król, Marek Kuś, Józef Lubacz, and Rafał Urbaniak—lists several challenges that are being faced by Polish ontology and philosophy in general. The presented facts are rather well known but rarely explicated, and the remedy for these shortcomings is less obvious and seemingly out of reach for now. Some of the contributing authors to this chapter question whether we are doing good research (which we seem to be), why are we ranking so low in the world market for ontological ideas, and why is what we do is rather unknown (albeit with notable exceptions) outside the limited circle of Polish universities. Our recent history and the relative obscurity of the Polish language clearly offer some excuses here, but these do not entirely explain everything away. The final section of this chapter, Skowron's excellent essay, takes a more positive note and is essential to retaining a balanced, yet critical, perspective on Polish Ontology. Unfortunately, the essay is rather short.

All the papers in the volume are clearly written and well organized, but a few additions could have improved the value of the collection. For example, the papers could have placed the presented research within the context of similar discussions outside Poland. In other words, the authors could have provided a well-documented background to the topics. If Polish philosophy is to come out into open and avoid the potential accusations of navel-gazing, it should do so by relating its findings to the dialogue among the international ontological community about current problems. Naturally, this should be done while preserving our unique and original perspective, although I concede this may be a challenge. With the notable exception of a few papers, this larger context is not really emphasized in the *Contemporary Polish Ontology*.[1]

What is also a little concerning is the limited references that are sup-

[1] To gain a wider perspective on ontology, we may look for comparison at the selection of topics presented during the latest Joint Ontology Workshops' (JOWO) meeting and see where *Contemporary Polish Ontology* fits in. The program of the JOWO 2020 Episode VI: *The Bolzano Summer of Knowledge* is available at <https://www.iaoa.org/jowo/2020/>.

plied in some articles, again with the notable exception of several papers. There is also a notable absence of some hotly pursued topics in current ontology, such as the ontologies of computing objects, sciences (e.g., biology, genetics, medicine, engineering), and system design, as well as object ontology. What is more, as we have indicated, what would potentially improve the book's reception is a short chapter that would place the book (assuming that the intention is to give an overview of Polish ontology) in the context of ontological research outside Poland. This information, as we have said, maybe found in the introduction to several essays, but unfortunately it is dispersed and not very detailed and systematically exposed. Moreover, the collection would provide a more rounded image of contemporary Polish ontology if it could avoid the potential critique of assuming a limited perspective, such as if the volume mentioned work originating from other places not associated with the ICFO. (In his review, Skowron does provide a more comprehensive and inclusive list of Polish philosophers dealing with ontological themes.)[2]

To summarize this review, we may say that *Contemporary Polish Ontology* does serve several detailed ontological studies from leading Polish research centers. The range of topics is rather broad, and anyone interested in ontology will certainly find something of interest. In hindsight, however, one may question why certain topics that are prevalent in the current ontological debates are absent from the volume. Maybe the selection of topics reflects the specificity of the Polish school, however. The clearly written introduction guides the reader to specific topics and therefore removes the need to read all of the abstracts. The final chapter, in contrast, is very much addressed to the Polish audience: One may call it a manifesto of sorts, a "What to do?" in the world of Polish ontology and Polish philosophy in general. From the book, one may come to know the key Polish researchers (their email addresses are provided) working in ontology. Such information is valuable, and the volume also

---

[2] For example, the volume could mention the work of Michal Heller on existence in physics (Heller, 2018) or Edward Malec's work on the existence of black holes (Malec, 2018). The volume from which these two works are cited is entitled *On what exists in physics*.

serves the secondary purpose of being a "Who's who" in Polish ontology.

Roman M. Krzanowski

Heller, M., 2018. What does it mean 'to exist' in physics? *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (65), pp.9–22. Available at: http://zfn.edu.pl/index. php/zfn/article/view/436 [visited on 28 June 2019].

Malec, E., 2018. Black holes: do they exist? *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (65), pp.47–59. Available at: https:// zfn . edu . pl / index . php / zfn / article / view / 438 [visited on 16 July 2020].

Skowron, B., ed., 2020. *Contemporary Polish Ontology* [Online]. Vol. 82, *Philosophical Analysis* 2627-227X. Berlin; Boston: De Gruyter. Available at: https://doi. org/10.1515/9783110669411.