Philosophical Problems in Science

ZFN

No 78
2025

Philosophical Problems in Science

Zagadnienia Filozoficzne w Nauce

e-ISSN 2451-0602 (electronic format)
ISSN 0867-8286 (print format, discontinued)

Editorial Office

Philosophical Problems in Science
(Zagadnienia Filozoficzne w Nauce)
e-mail: info@zfn.edu.pl
www.zfn.edu.pl

Copernicus Center

Publisher: Copernicus Center Foundation
Pl. Szczepański 8, 31-011 Kraków POLAND
tel. (+48) 12 422 44 59
e-mail: info@copernicuscenter.edu.pl
www.copernicuscenter.edu.pl

Philosophical Problems in Science

Zagadnienia Filozoficzne w Nauce

78 — 2025

Articles

Philosophical Problems in Science (2025), **78**, 3–11
(Zagadnienia Filozoficzne w Nauce)
doi:10.59203/zfn.78.717

Philosophical Problems in Science

ARTICLE

The synthetic *a priori*. Rejoinder to Linsbichler

Walter Block

Loyola University New Orleans
Email: wblock@loyno.edu

Abstract

Linsbichler is a gifted economist and philosopher. He delves mightily and thoroughly into the difficult thickets of basic praxeology, the methodology of the Austrian school of economics. Not content with merely probing the meaning and importance of the synthetic *a priori*, he pushes further into the very justification of this foundational element of Austrianism. I learned a lot from reading this splendid essay, most of with which I enthusiastically agree. However, there are a few divergences between the two of us, and the present paper is devoted to exploring them. My procedure in this response is one of mentioning numerous quotations from the author, interspersed with my own comments and reactions. That is because I greatly appreciate this gifted economist's contribution to Austrian methodology.

What are the specifics? One error is that he does accurately distinguish between analytic, empirical, and synthetic *a priori*. Another is his misunderstanding of "human action". My debating partner opines that Mises (1998) merely suggests that human being act. In actual point of fact in the of this latter author, that is the very title of his most famous book, and contains the very core of praxeological economics. Then Linsbichler fails to accurately distinguish between mere behavior and human action. Further there is a dispute over the source of ideas. My learned colleague maintains it is empirical. I demur. He is of the opinion that empirical criticisms can be relevant to praxeological findings. I attempt to correct him on this matter. He takes the position can something can be "mildly aprioristic". I demonstrate that this concept is an all or none phenomenon; no gradations.

Keywords: tautology, empiricism, synthetic a priori, purposefulness, economic methodology, Austrian economics, praxe

As an Austrian economist, I am grateful to Linsbichler (Linsbichler 2024). In my view, this should apply to all of my fellow praxeologists.[1] Before reading this very important contribution to our discipline, if I had thought about this issue, which I had not, I would have thought that there was nothing more to say about the synthetic *a priori*. Thanks to him, I have now been disabused of this erroneous notion of mine. I think this brilliant author make some very telling and important points, many of which I go along with, and from which I have learned. However, I find I cannot go along with him 100% of the time. The present paper explores some of the divergences between him and myself.

As I was taught at my mentor's [Murray N. Rothbard] knee, Austrian economics is predicated upon the synthetic *a priori* insight. There are certainly claims that are not tautologies, or analytic, but, rather, synthetic, in that they pertain to the real world and are thus not mere definitions. Rather, they are necessarily true. They are synthetic *a priori* statements.

Hoppe (Hoppe 1995) offers a few examples:

> Whenever two people A and B engage in a voluntary exchange, they must both expect to profit from it. And they must have reverse preference orders for the goods and services exchanged so that A values what he receives from B more highly than what he gives to him, and B must evaluate the same things the other way around.
>
> Or consider this: Whenever an exchange is not voluntary but coerced, one party profits at the expense of the other.
>
> Or the law of marginal utility: Whenever the supply of a good increases by one additional unit, provided each unit is regarded as of equal serviceability by a person, the value attached to this unit must decrease. For this additional unit can only be employed as a means for the attainment of a goal that is considered less valuable than the least valued goal satisfied by a unit of such good if the supply were one unit shorter.
>
> Or take the Ricardian law of association: Of two producers, if A is more productive in the production of two types of goods than is B, they can still engage in a mutually beneficial division of labor. This is because overall physical productivity is higher if A specializes in producing one good which he can produce most efficiently, rather than both A and B producing both goods separately and autonomously.
>
> Or as another example: Whenever minimum wage laws are enforced that require wages to be higher than existing market wages, involuntary unemployment will result.
>
> Or as a final example: Whenever the quantity of money is increased while the demand for money to be held as cash reserve on hand is unchanged, the purchasing power of money will fall (Hoppe 1995, pp.14–15).

What, then, are we to make of Linsbichler statement to the effect that: "conventionalism[2] avoids the charge of extreme apriorism by construing the fundamental axiom of praxeology as analytic instead of synthetic" (Linsbichler 2024, p.43).

First of all, there is no need to "avoid [. . .] the charge of extreme apriorism" the view of Murray Rothbard. Rather, all praxeologists may embrace it. Second, the fundamental axiom of praxeology is neither analytic *a priori* nor synthetic *a posteriori*. The former is a tautology of the sort "Bachelors are unmarried men". This tells us nothing about the real world, whether anyone is married or not. It only indicates how we define the words "married" and "bachelor". It is necessarily true that no bachelor is married. This cannot be tested or falsified. Second, "synthetic" is an empirical claim, such as "it is now raining". This tells us about the real world, alright, but is not necessarily true. It can be tested and falsified, when it is a sunny day. Austrian economics, or praxeology, consists of a series of synthetic *a priori* states of the sort offered by Hoppe which both pertain to reality, and, are necessarily true.

Next consider this statement of Linsbichler's:

> Explications of the overly short "man acts" identify its content along the following lines: human individuals and only human individuals (as opposed to viruses, planets, or social classes) at least sometimes behave purposefully, i.e. they choose goals and apply means they subjectively consider expedient to attain these goals on the basis of their beliefs. Strictly speaking, the way Mises and other Austrian economists apply the fundamental axiom only suggests that human individuals act and none of the other known types of objects act. In case we encounter intelligent aliens, praxeologists might reconsider the "and only human individuals" clause(Linsbichler 2024, pp.44–45).

I maintain, in contrast, that the way Mises (Mises 1998) and Rothbard (Rothbard 1962) and other Austrians "apply the fundamental axiom" does far more than "suggest that human individuals act". This is an integral part and parcel of the entire enterprise. Nor is the claim made that *only* people can engage in action, as opposed to mere behavior. Certainly, when and if intelligent aliens come visit us in their spaceships, no Austrian will have any difficulty in ascribing action[3] to them.

Saith our author: "Strictly speaking, the fundamental axiom is of no help in ascertaining whether a certain human behavior is merely behavior or an action either" (Linsbichler 2024, p.45). True enough. The fundamental axiom, there is such a thing as human action, cannot itself make this determination. Needed is some aspect of *Verstehen*, or human intelligence to decide upon this matter. But the same goes for the law of non-contradiction, excluded middle, or any of the other basis building blocks, foundations, of logic. They, alone, cannot decide whether or not something is coherent. They can be the basis upon which such a decision is made, but the "decider" has to be a reasonably intelligent person.

Our author maintains that "experience is not the source or origin of ideas for theories" (Linsbichler 2024, p.47). I say, in contrast, "different strokes for different folks. Or, to be more technical about this matter: *De gustibus non disputandum*". Also, in a similar manner, the source or origin of ideas is very subjective. Some people get new ideas while they are asleep, in the form of a dream; most of us, while awake. Some while being a couch potato, others while engaging in physical activity. I don't think it is reasonable to ascribe any one "source or origin" of ideas for all people.

Linsbichler's next foray is this: "[...] extreme apriorism which immunizes large parts of theory from empirical criticisms, has become highly suspect in the development of philosophy of science and, with some time lag, also among economists [...]" (Linsbichler 2024, p.47). I regard this as an understatement of the year. "Highly suspect," nothing. Far worse, Austrian economists are denigrated as "cultists" for claiming that basic praxeological truths cannot be tested any more than can the Pythagorean Theorem. For example, states Krugman (Krugman 2013): "Austrian economics very much has the psychology of a cult. Its devotees believe that they have access to a truth that generations of mainstream economists have somehow failed to discern; they go wild at any suggestion that maybe they're the ones who have an intellectual blind spot. And as with all cults, the failure of prophecy — in this case, the prophecy of soaring inflation from deficits and monetary expansion — only strengthens the determination of the faithful to uphold the faith" (Krugman 2013). Here is Samuelson's (Samuelson 1964) evaluation of Austrian economics: "I tremble for the reputation of my subject" (Samuelson 1964, p.736). Rosen (Rosen 1997) asks if there are any gains to be made from trade (interaction between) Austrian and mainstream economics, and answers very strongly in the negative. These critics cannot seem to understand that there are *a priori* claims "out there" that cannot possibly be refuted by any empirical evidence, and yet are not tautologies.[4]

Linsbichler is entire correct, in his own understated manner, when he writes: "[…] since the standard view maintains that praxeology depends on extreme apriorism, philosophers and economists have condemned praxeological methodology as well as economic claims based on praxeological research" (Linsbichler 2024, p.48). He and I, however, diverge from one another when he opines: "Only the fundamental axiom is a *priori and very little is implied by the fundamental axiom without additional premises"* (Linsbichler 2024, p.48). To the contrary, there are numerous synthetic *a priori* statements in Austrian economics[5]; these are by no means limited to the foundational axiom that man acts.

Linsbichler's next bite at the apple is this: "Praxeology in Mises's tradition faces two problems: (i) if it is extremely aprioristic as most interpretations hold, then it is considered untenable in light of contemporary philosophy of science" (Linsbichler 2024, p.49). There is a quick and easy response to this "problem": contemporary philosophy of science is out to lunch; it is mistaken from the get go. I don't consider it a problem that "contemporary philosophy of science" diverges from praxeology and is thus mistaken.

Here is his next critique: "(ii) Mises's writings seem to allow for radically different interpretations as to how he attempts to justify praxeology and consequently how extreme his apriorism is" (Linsbichler 2024, p.49). Yes, I acknowledge, there are several and diverging accounts for what might be called "meta-praxeology": justifying the entire enterprise, categorizing it[6], accounting for it, looking at its psychological predispositions. However, about praxeology itself, surely a more important issue, there can be little doubt as to its veracity, at least not any emanating from the present quarter.

I think Linsbichler goes off the rails when he "[…] proposes a conventionalist defense of analytic praxeology, first embedded in a broader reconstruction of Mises's methodological views […] and later more focused and detailed on conventionalist praxeology […] The vital step is to construe the fundamental axiom as analytic instead of synthetic *a priori*" (Linsbichler 2024, p.49). No, no, no, "analytic" is tautologous. To be sure, the synthetic *a priori* statement resembles the analytic: it is necessarily true, cannot be tested, only illustrated. But, it also has nothing to do with empirical reality. This construction, then, surrenders half the intellectual power of the synthetic *a priori*.

I take exception to this phraseology of our author: "An explication of praxeology with an analytic fundamental axiom and with limited a *priori scope is only mildly aprioristic"* (Linsbichler 2024, p.54). In my view, there is no either/or here. Like pregnancy, you either exhibit this characteristic or you do not. A "mild" synthetic *a priori* statement only indicates how far apart I am on praxeology from Linsbichler.

Our author avers: "[…] the extension of praxeology to all human activity remains largely programmatic" (Linsbichler 2024, p.58). In one sense, he is correct. The application of praxeology to non economic issues such as war, or games, is largely in the future[7]. But in another sense, instead of "programmatic" he should have said "problematic." For, surely, there are vast areas of economics to which praxeology cannot apply. We know it is a praxeological insight that ceteris paribus, rent controls on residential housing will lead to fewer apartment units. But how big will be the reduction? 5%? 10%? Praxeology can vouchsafe us no answer at all on such matters. And, yet, surely, this is a question of proper interest to economists.

Linsbichler asserts: "Of course, in Rothbard's work, too, the economist is a fallible human being, and critical debate is the key to scientific progress" (Linsbichler 2024, p.61). It cannot be doubted that the praxeologist can err. All human beings are subject to mistakes. But we have to distinguish between statements like "It is raining now" or "team A will beat team B in some sport" on the one hand from the Pythagorean theorem, or that "trade is necessarily mutually beneficial ex ante" on the other. With regard to the first set of empirical statements, people are indeed often wrong. With regard to the second state, while intellectual modesty prevents us from saying these synthetic *a priori* statements are ineluctable, necessarily true, no one who states them can ever be mistaken, in a very different sense: that being the one where we may, possibly, be mistaken when we proclaim that two plus two equals four. Praxeologists are mightily and justifiably tempted to proclaim synthetic a *priori* statements of that sort knowing full-well that to be human is to be subject to error. But it a very different type of mistake, if mistake, indeed, it is.

This conception of science as a search for certainty may be partly responsible, at least psychologically, for the vehemence with which Rothbard and many other Austrians advocate their economic and oftentimes also their political positions. Hayek said he cannot follow Mises there, into the realm of the synthetic *a priori* (Knott 2012; Wozinski 2010; Hayek 1943). Milton Friedman claimed that praxeologists who disagree with one another can only settle their differences via a physical fight.[8] Yes, Rothbard is "vehement" about praxeology, as "vehement" as would be Friedman or Hayek about the truism that two plus two equals four.

Let us delve more deeply into the status of "man acts". Posit that human being never existed. There were apes, but they all died out before man could emanate from them. It is entirely true that under these conditions, a few pieces of wood could have been found in the formation that perfectly resembles this statement, "man acts". Perhaps the wind might have blown some leaves to take up this formation. If an infinite number of monkeys pecking away at a typewriter, nowadays, a lap top, could have come up with one of Shakespeare's plays, this concatenation of events with wood or leaves could have occurred, if enough time had passed for this to come into being. How, then, would a non-human but intelligent being properly interpret this statement. It would be an empirical claim, but a false one. However, if a man wrote it, then it would an empirical claim, and, also, necessarily true, thus, a synthetic *a priori* statement. Any attempt to refute it would strengthen it. For example, to say that man does not act is itself an act, thus self-refutational.

According to Linsbichler, "While Rothbard classifies the auxiliary axioms of praxeology as obviously true in our world, a counterfactual scenario in which they are false can be thought of without contradiction, so they are not necessarily true" (Linsbichler 2024, p.63). But no example is given by this philosopher economist.. No counterfactual scenario can be thought of, apart from the wood or the leaves forming that statement, but, for reasons given, that can hardly count. For, here, such a claim would be empirical, but false.

Saith Linsbichler: "What distinguishes truth from absolute truth in Rothbard's nomenclature is not entirely clear. The formulation can be read as an expression of the lack of differentiation between truth and certainty. In other passages, Rothbard seems to have in mind truth without exception in our world or the much stronger truth in all possible worlds, i.e. necessary truth" (Linsbichler 2024, p.66). What distinguishes truth from absolute truth in Rothbard's nomenclature seems a bit more clear to me than it does to Linsbichler. Take these two claims. One, it is now raining outside. Two, voluntary trade benefits both parties ex ante. The first one is indeed true, since it is now raining. But it is a mere empirical claim which need not always be true. Sometimes, the sun is shining. But it is possible to deny the truth of this correct statement, without becoming enmeshed in self-contradiction. The second statement is absolutely and necessarily true, a higher order of truth if it can be put that way, since the denial is not only false, but self-contradictory. It is true under all possible scenarios, while the first is only true when it rains.

Here is another contribution by this author:

> [...] for the sake of argument let us concede to Rothbard that he has intuited, with necessary truth, that he himself has goals and uses means to achieve them. The main difficulties for establishing the fundamental axiom in a Rothbardian manner arise when one tries to infer statements about the minds of other people from inner experience. How is it possible to draw necessary conclusions about other people from the exploration of one's own consciousness? Since Rothbard requires and considers the fundamental axiom to be empirically meaningful, the term 'human' is at least partly interpreted, i.e. at least for many paradigmatic cases it is determined which physical objects are in the extension of 'human' and which are not. Suppose m is one such human individual and suppose it turns out that m does not act (Linsbichler 2024, p.69).

To me this is akin to supposing the proverbial square circle. We can posit such a supposition all we want but we can succeed in doing no such thing. That is, we can use "square circle" in a sentence, as I am now doing, but that cannot render this concept coherent. In like manner we can use the phrase, "M is a human being who does not act" in a sentence, as, again, I am now doing but that does not render this thought intelligible. The only human being who does not act is a dead human being, and it is not at all clear that this ex-person is still a human being. More to the point, he is a cadaver, and can no longer act. But while he is alive he must necessarily engage in human action, or he will no longer be able to boast of that status.

Linsbichler asserts: "Then, for Rothbard, the potential immunization strategy of simply not calling everything that does not act as a human being is blocked" (Linsbichler 2024, p.69). This is not a mere mistaken definition on Rothbard's part. I think it is incumbent upon Linsbichler to give an instance of a human being who does not act. He fails to do so.

Next, consider this assertion:

> By Rothbard's own standards, not even the proposition 'M acts' is intersubjectively verifiable. How much more problematic is the demand that the fundamental axiom 'All people act' can established as true.
>
> Furthermore, it is dubious how Rothbard's account can show that a falsification of the fundamental axiom is inconceivable... But theories in which other people merely behave instead of acting purposefully can be conceived and formulated without special problems (Linsbichler 2024, pp.70–71).

Let us consider an example. Joe buys a hat from Sally for $20. How much does he value that headgear? If at $15, he would scarcely make the purchase, since he would then lose $5 in value. Nor could he rate this article of clothing at exactly the purchase price of $20, for then there would be no profit in the transaction for him. Why would he bestir himself, get off his comfortable couch, if he anticipated no return at all for his efforts? Clearly he would not. So far, we are discussing purposeful behavior, which undergirds yet another synthetic *a priori* statement: all voluntary trade is mutually beneficial, at least ex ante. Ditto for Sally. She has so many of these hats in her shop, and is desirous of getting rid of them lest styles change. She values the one sold to Joe at minus $1, and thus earns a profit of $21 from the deal. Now, it is entirely possible that the purchaser is indifferent to the hat, or even dislikes it. But there is something about it that he values more than the purchase price. Perhaps he thinks that if he engages in this commercial interaction, he will get a date with Sally. None of this sounds like mere "behavior." It practically reeks of purposefulness. If it was mere behavior, we could not deduce that voluntary trade is mutually beneficial, a synthetic a *priori* statement if ever there was one. If it was mere behavior, like sneezing or the knee reflex, the truth would no longer hold that trade is beneficial ex ante.

States our author: "Rothbard's attempt to establish certainty, intersubjectivity, and truth and for the fundamental axiom is on shaky ground". He predicates this charge on the basis of "the conceptual distinction between certainty and truth" (Linsbichler 2024, p.71). But it is not clear how this distinction can undermine the Rothbardian project. Rothbard's view is, au contraire, compatible with the notion that empirical statements can be true but never certain in that their denial does not imply a logical contradiction. Whereas synthetic a *priori statement can not only be true, but, also, certain, in that their denial does indeed imply a logical contradiction. Trade is not mutually* beneficial ex ante? That leads to the highly problematic claim that people do not seek to better their condition.[9]

Linsbichler now launches into his positive analysis: "A falsification of the fundamental axiom is inconceivable [...] Maybe humans do not have goals, maybe door handles do, but how could we experience this?" (Linsbichler 2024, p.72). Obviously, this is meant more in jest but it stems from this author's skepticism.

He continues: "Even when granting the validity of introspection for one's own mental states, it remains possible and conceivable – in principle – that all other human individuals do not act but merely behave" (Linsbichler 2024, p.73). I cannot see how this move us in a productive direction. As mentioned above, if this were the case, how do we account for the synthetic *a priori* fact that people "behave" in such a manner as to improve their lot?[10]

Here is yet another criticism by Linsbichler:

> Rothbard's criterion requires 'absolute' truth (whatever that exactly amounts to). If we interpret this as being true in all linguistic frameworks, no matter how the terms are defined in them, any justification must obviously fail. The sentence 'Murray is a libertarian' is true if the terms have their usual meaning, but we can easily render the sentence false by changing the meaning of 'Murray' or of 'libertarian'. And to ask for the 'truth' of a sentence, independently of a framework which assigns meanings to the sentence, is unintelligibly with standard notions of truth. No sentence, considered as a purely syntactic string of signs, is true independently of the meaning attached to it (Linsbichler 2024, pp.74–75).

Of course, if we change the meaning of the word "Murray", or "libertarian" we can falsify this obviously true statement. What is its status, without any such "changes?" It is an obviously true assertion, but it is empirical, not synthetic *a priori* and certainly not tautological. That is, it implies to no logical contradiction to say that Murray is not a libertarian. It is even true that at one time in his life, Murray was not a libertarian, I presume; for example, when he was a two year old toddler, with, I strongly suspect, no views at all on political economy.

To conclude: Linsbichler (2024) is an attack on Rothbardian-Misesian-Hoppean "extreme apriorism". This author mistakenly maintains that the fundamental axiom of praxeology is analytic; that is, a tautology; having nothing to do with how the real world operates. I maintain, in contrast, that praxeology is the study of the synthetic *a priori* statements, economic laws, that, yes, are apodictically and necessarily true, just like purely analytic statements, or tautologies, but, also, unlike analytic statements, or tautologies, can explain and help us understand real world events.

Notes

1 I am grateful to two active and incisive referees of this journal. Thanks to their efforts, this paper is much improved. I also thank Emily Threeton for splendid copy editing. The usual caveats of course apply: all remaining errors and infelicities remain with me.

2 This is the view that Linsbichler favors.

3 Semi, demi, quasi, human action.

4 For rejoinders see: (Block 2013; Block, Westley, and Padilla 2008; Rossini 2013).

5 Several of them mentioned above, courtesy of Hoppe. Another problem I have with this essay of Linsbichler's is that all throughout this very long paper of it, he uses but one example of the synthetic *a priori*: man acts. But there are many more, as we have seen.

6 Is it predicated upon empirical, ideological, logical premises.

7 We are now living some 75+ years after the publication of *Human Action*. Why are such developments still non-existent? My only guess is that there is such a thing as specialization and the division of labor. Despite the growth of the number of Austrian economists, evidently it has not been sufficient to call forth any contributions of which I am aware.

8 Friedman (1991, p.18): "[In Mises's view, we] have absolutely certain knowledge of [...] motivations of human actions [...] and [...] we can derive substantive conclusions from that basic knowledge. [...] Facts, statistical or other evidence cannot [...] be used to test those conclusions [...] Suppose two people who share von[202F?]Mises's praxeological view come to contradictory conclusions about anything. [...] The only way they can reconcile their difference [...] is by a purely logical argument. [...] Suppose neither believes he has made a mistake in reasoning. There's only one thing left to do: fight".

9 Suicide is not a counter example to this claim. People who do so seek to improve their lot, not to worsen it.

10 Linsbichler's analysis is marred by his wokism: In this limited sense, criterion (I) is almost fulfilled, as long as the economist stays within *her* linguistic framework. Of course, *she* can step out of *her* linguistic framework, abandon *her* research program, and conceive of behaviorism in a meta-language. Although, to be sure, there are dozens, maybe hundreds of women who have made important contributions to economics and philosophy, they are greatly outnumbered by male contributors. Why make it seem as if the very opposite were the case with the use of this intemperate language? emphasis added by present author.

References

Block, Walter E. 2013. *Is Austrian Economics a Cult? (Critique of James Buchanan and Gary Becker).* http://archive.lewrockwell.com/block/block221.html accessed April 12, 2025.

Block, Walter E., Christopher Westley, and Alex Padilla. 2008. Internal vs. external explanations: a new perspective on the history of economic thought. *Procesos De Mercado: Revista Europea De Economia Politica* 2:35–132. http://www.jsu.edu/depart/ccba/cwestley/Internal.2008.pdf.

Friedman, Milton. 1991. Say 'No' to Intolerance. *Liberty Magazine* 4 (6): 17–20.

Hayek, Friedrich A. 1943. The facts of the social sciences. *Ethics* 54 (1): 1–13. https://www.journals.uchicago.edu/doi/abs/10.1086/290368?journalCode=et.

Hoppe, Hans-Herman. 1995. *Economic Science and the Austrian Method.* Auburn: The Ludwig von Mises Institute.

Knott, Adam. 2012. *Hayek and Praxeology.* https://mises.org/mises-daily/hayek-and-praxeology accessed October 16, 2025.

Krugman, Paul. 2013. *Fine Austrian Whines.* http://krugman.blogs.nytimes.com/2013/02/20/fine-austrian-whines/?_r=%7B0%7D accessed February 20, 2025.

Linsbichler, Alexander. 2024. What Rothbard could have done but did not do: The merits of Austrian economics without extreme apriorism. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 76:43–84. https://doi.org/10.59203/zfn.76.67.

Mises, Ludwig von. 1998. *Human Action: A Treatise on Economics.* Auburn: Ludwig von Mises Institute.

Rosen, Sherwin. 1997. Austrian and Neoclassical Economics: Any Gains from Trade? *Journal of Economic Perspectives* 11 (4): 139–152.

Rossini, Chris. 2013. *Paul Krugman: Hide Thy Enemy.* http://www.economicpolicyjournal.com/2013/12/paul-krugman-hide-thy-enemy.html accessed December 31, 2024.

Rothbard, Murray N. 1962. *Man, Economy, and State.* Auburn: Ludwig von Mises Institute.

Samuelson, Paul A. 1964. Theory and Realism: A Reply. *The American Economic Review* 54 (5): 736–739.

Wozinski, Jakub. 2010. *Hayek and Departure from Praxeology.* https://mises.org/libertarian-papers/hayek-and-departure-praxeology accessed October 16, 2025.

Philosophical Problems in Science (2025), **78**, 13–25
(Zagadnienia Filozoficzne w Nauce)
doi:10.59203/zfn.78.721

ARTICLE

Is the "unreasonable effectiveness of mathematics" a miracle that points to God? Wigner and Craig on the applicability of mathematics

James Franklin
University of New South Wales
Email: j.franklin@unsw.edu.au

Abstract

Eugene Wigner's 1960 article on the "unreasonable effectiveness of mathematics" used the word "miracle" of the fit between abstract mathematics and physical reality. William Lane Craig has developed a theistic argument from Wigner's hints, claiming that the best explanation of the "miraculous" fit is divine creation. It is argued that this argument does not succeed. An Aristotelian realist philosophy of mathematics renders the applicability of mathematics to physical reality unmysterious by showing that mathematics, like any other science, is a study of certain aspects of reality, hence there is no miracle of fit. However, that does not preclude other arguments for the existence of God involving mathematics, for example design arguments from the elegance of the universe's structure, fine-tuning arguments or ones from the nature of mathematical understanding.

Keywords: mathematicity, unreasonable effectiveness, philosophy of mathematics, Aristotelian realism, argument to God from mathematics, William Lane Craig, Eugene Wigner

Introduction

Eugene Wigner's 1960 article, *The unreasonable effectiveness of mathematics in the natural sciences*, now cited 4000 times, is probably the most celebrated philosophy of mathematics article of the last century. Its appeal rests on its case that the fit between mathematics and physical reality is somehow mysterious.

Its much-quoted conclusion is: "The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve" (Wigner 1960, p.14). The word "miracle" appears another eleven times in the article. Wigner does not intend that to be read literally, and simply leaves the mystery unsolved—no doubt contributing to the article's popularity.

But it is reasonable to ask whether "miracle" might be taken literally, and to ask whether a miracle needs a miracle-worker. Is there an argument for the existence of God from the otherwise inexplicable "mathematicity" of the physical world? The theist author William Lane Craig has developed carefully-argued philosophical versions of the argument.

We will examine the phenomena adduced by Wigner, then the argument of Craig. We will argue that this argument is not convincing. It depends on philosophies of mathematics such as formalism and Platonism which imply a gap between the world of mathematics and the world of physical reality, rendering their match a mystery. An Aristotelian realist philosophy of mathematics which sees mathematics as a science of certain aspects of physical (or any other) reality will dissolve the mystery of the "applicability" of mathematics. Mathematics is "applied" from the start, so there are not two realms which need a God to join them in marriage.

The Wigner-Craig argument is distinguished from several others that can be made which take some mathematical aspect of reality to point towards God. They include design arguments from the specially elegant or beautiful mathematical structure of the world, fine-tuning arguments, and arguments from the nature of mathematical understanding. Those arguments may have some credibility, but they are distinct from that of Wigner and Craig.

Finally, there is a reason why the theist should prefer a necessitarian account of the mathematicity of the universe—that is, one in which the mathematical necessities in physical reality are not put there by (and hence point to) God but instead constrain God. As is widely agreed, the most pressing philosophical problem for classical theism is the problem of evil. If, as Leibniz thought, God does not create a better world because (as a matter of mathematical necessity, due to the interconnectedness of the world) there is no better world, then the problem of evil becomes solvable.

Three questions on the relation of mathematics to reality

Before examining the arguments of Wigner and Craig, three questions on the "effectiveness" of mathematics as applied to physics or any other science should be distinguished. As we will see, different authors emphasise one or other of them, so it is helpful to have the distinctions clear before describing particular views.

First, there is the question of why mathematics, including elementary mathematics, is effective in or applicable to physical reality at all. That question has a low profile in Wigner's article but is the focus of, for example, Kant's theories that geometry and arithmetic arise from the mind's imposition of the forms of space and time on an unstructured and unknowable "noumenal" reality. One might attempt to argue to God from the "mathematicity" of the universe, its being describable by mathematics at all; that approach is taken by Heller (2019, p.241), explained further in (Trombik 2025), section 3.2) but it is rather different from that of Wigner and Craig and is not pursued here.

Secondly, there is the question of why many of the abstract and esoteric theories of modern higher pure mathematics, such as abstract groups, Riemannian geometry and Hilbert spaces, turn out long after their discovery to be relevant to high-level physics such as quantum mechanics. It appears surprising that tensor mechanics just happened to be studied by Einstein and proved to be the right mathematical language for general relativity, and that infinite-dimensional Hilbert spaces over the complex numbers were found to be perfect for quantum mechanics. That phenomenon is much more evident in contemporary fundamental physics than in other sciences, which require some forms of mathematics but do not generally benefit from the most abstract mathematical technology (with some possible exceptions such as stochastic differential equations in finance). Wigner and Craig's arguments mostly relate to this second question.

Thirdly, there is the question of accuracy of mathematical predictions. Wigner calls attention to the fact that the predictions of fundamental physics made on the basis of mathematical laws often prove later to be experimentally verified to very many decimal places, sometimes to as many decimal places as powerful experimental measurements can reach. That is not true of mathematical predictions in economics (Velupillai 2005), or in weather prediction.

All of those "fits" of mathematics to physical reality might be regarded as surprising and thought to need explanation, but they are distinct phenomena and *prima facie* their explanations might be different. Indeed Wigner and Craig find elementary mathematics unsurprising but believe the other two questions are either unanswerable or need a divine intervention. Wigner emphasises the difference between the first and second questions: "whereas it is unquestionably true that the concepts of elementary mathematics and particularly elementary geometry were formulated to describe entities which are directly suggested by the actual world, the same does not seem to be true of the more advanced concepts, in particular the concepts which play such an important role in physics" (Wigner 1960, p.2).

It will be argued here however that an Aristotelian realist philosophy of mathematics gives essentially the same answer to all three questions, while undermining the appearance of surprise. Mathematics, whether elementary or advanced, and whether badged as pure or applied, describes absolute mathematical necessities which any physical reality must obey, and obey exactly (where precision is relevant). Since they are necessities, no God is needed to determine them, and hence there is no argument from them to God.

Distinction from design arguments

These questions and the arguments of Wigner and Craig need to be distinguished carefully from design arguments for the existence of God, even though design can have mathematical aspects. In Paley's classic version, the intricate internal arrangements of living organisms are compared to the design of clockwork and an inference is made to the existence of a divine designer. Even though clocks work with "mathematical precision" and their structure could be described mathematically, the argument does not arise from mathematics but from the contingent fact that organisms and clocks have a certain complex structure (which allows them to perform certain impressive tasks).

The same is true with arguments from fine-tuning, which again involve some mathematics but where the argument arises from a contingent property of the universe. It is said that the values of the fundamental constants of nature lie in certain narrow ranges which permit a physical universe with life to appear; it appears to be contingent that those values are as they are, so a divine explanation of their being there is reasonable (Lewis and Barnes 2016). That argument resembles a design argument in beginning with a contingent fact about the universe, but it is not an argument like that of Wigner and Craig from the mathematicity of the universe.

Similar arguments could arise from the beauty of the universe. If it is thought that the universe could be either beautiful or not, then its being beautiful (perhaps in a mathematically describable way) could give rise to an argument that God created it that way (e.g., Kessler

2022). That is again a possible argument from a contingent way the universe is to a divine act of creation choosing that option, but it is not an argument like that of Wigner and Craig.

A different argument begins with the "miracle" that the universe is "rational", "intelligible" or "comprehensible". This means that the observed mathematical structure of the physical universe must be adapted to the cognitive capacities of humans: neither so simple that the world is virtually structureless, nor so transcendent or complex that our minds could not grasp it (Heller 2019, p.242; more extensively in Coyne and Heller 2008). Or it can be argued that human mathematical understanding of necessities is itself miraculous and does not fit into a materialist view of the universe, and hence points to God. Perhaps what is needed to grasp mathematical necessities is something like Aristotle's *nous* (medieval Latin *intellectus agens*), a divinely-granted faculty of the soul capable of grasping the necessity of first principles (Bronstein 2016; Kuksewicz 1988). Those arguments are worth pursuing but like design arguments they proceed from contingent facts about the universe and are quite different from the argument of Wigner and Craig.

Two senses of "surprising"

As well as distinguishing questions on the applicability of mathematics, it is desirable to distinguish two senses of the notion of "surprising" (or "unreasonable") which plays such a prominent role in the arguments. The notion is used without much explanation, especially in Wigner, but his arguments cannot be evaluated without some querying of it.

Is that notion intended in a subjective or objective sense? In the subjective sense, a fact is surprising when we (due often to our cognitive limitations) are in fact surprised by it. That kind of surprise needs no explanation beyond our limitations, and any sense of surprise is dissipated when an explanation appears and we understand it. For example, an amazing observed mathematical regularity is that a number is divisible by 9 if and only if the sum of its digits is divisible by 9. But when one sees a proof of that regularity that explains why it must be so, one's amazement abates. The necessity of the regularity is absolute, and there is no role for a God to impose it. It could not have been otherwise. Hence, there is nothing inherently "unreasonable" in the apparent coincidence and no argument from that regularity to God.

On the other hand, something can be called objectively surprising when, even after the facts are known, something of objectively low probability has occurred—for example, a tossed fair coin comes up heads ten times in a row. What is expected, on the basis of how things actually work, is that a tossed fair coin comes up sometimes heads and sometimes tails, in an unpredictable pattern. If something other than that occurs, it is objectively surprising; the surprise is not merely relative to our stupidity.

In fact the opening paragraph of Wigner's article makes this distinction with a memorable example, though the point does not recur in the article. Wigner imagines a statistician explaining the normal distribution of a population to a mathematically naïve former classmate. The classmate says:

> "And what is this symbol here?" "Oh," said the statistician, "this is pi." "What is that?" "The ratio of the circumference of the circle to its diameter." "Well, now you are pushing your joke too far," said the classmate, "surely the population has nothing to do with the circumference of the circle" (Wigner 1960, p.1).

Wigner comments, "Naturally, we are inclined to smile about the simplicity of the classmate's approach" (Wigner 1960, p.1). That is because "we", the mathematically literate, can understand why the formula for the normal distribution must contain pi—even though the naïve classmate is right to be astonished initially at such an unlikely appearance of pi. Our surprise dissipates when we understand the necessity of the reasons.

We should bear this example in mind when examining applications of higher mathematical concepts of whose grasp we are less certain. Whenever an example of something surprising is given, we should ask if better understanding of mathematics on our part would dissipate that surprise. If there is talk of esoteric concepts like infinite-dimensional Hilbert spaces, we should guard against being in the position of the naïve classmate who understands insufficient mathematics. Any case where surprise could be relieved by more mathematical knowledge is not miraculous and there can be no argument to God from it.

The argument of Wigner's "unreasonable effectiveness"

Wigner advances several cases of the allegedly "unreasonable effectiveness" of mathematics in physics. The first one is Newton's inverse square law of gravitation: that the force of gravity exerted by two bodies is exactly inversely proportional to the square of the distance between them. Wigner writes:

> The law of gravity which Newton reluctantly established and which he could verify with an accuracy of about 4% has proved to be accurate to less than a ten thousandth of a per cent [...] [it is] a monumental example of a law, formulated in terms which appear simple to the mathematician, which has proved accurate beyond all reasonable expectations (Wigner 1960, p.8).

Wigner thus emphasises in this example the accuracy of predictions, rather than the application of mathematical concepts as such. Are the mathematical concepts used in this example in themselves surprising? Wigner says that the most sophisticated one, a second derivative, "is simple only to the mathematician, not to common sense or to non-mathematically-minded freshmen" (Wigner 1960, p.8), that is, it is objectively simple but subjectively not.

So in this first example, the only objectively surprising or miraculous aspect claimed by Wigner is the accuracy of predictions. His conclusion is somewhat different in his second example, the high-powered mathematical concepts used in "ordinary, elementary" quantum mechanics. As Wigner describes it, Heisenberg had devised some rules of computation to summarise experimental results in some simple cases such as the hydrogen atom. Born noticed formal similarities between those rules and the pure mathematics of computation with matrices, and it was then proposed to use the mathematics of matrices instead of the rules of computation. "There was, at that time", Wigner says, "no rational evidence that their matrix mechanics would prove correct under more realistic conditions" (Wigner 1960, p.9). However, it did prove correct, not only in realistic conditions but in much more general conditions. "The miracle occurred only when matrix mechanics, or a mathematically equivalent theory, was applied to problems for which Heisenberg's calculating rules were meaningless" (Wigner 1960, p.9), such as heavier atoms. A calculation with the next heaviest atom, helium, "agrees with the experimental data within the accuracy of the observations, which is one part in ten million. Surely in this case we «got something out» of the equations that we did not put in" (Wigner 1960, p.9).

Wigner thus emphasises again the accuracy of predictions. He does not offer an opinion at this point on whether matrix mechanics is really complex or objectively surprising, though he does note its disconnection from experience. That is an example of a phenomenon Wigner mentioned earlier in the paper, one which has been taken up by many commentators:

> Whereas it is unquestionably true that the concepts of elementary mathematics and particularly elementary geometry were formulated to describe entities which are directly suggested by the actual world, the same does not seem to be true of the more advanced concepts, in particular the concepts which play such an important role in physics [...]. The complex numbers provide a particularly striking example for the foregoing. Certainly, nothing in our experience suggests the introduction of these quantities (Wigner 1960, p.2).

Wigner here calls attention to higher rather than elementary mathematics, the second of the concepts of "mathematicity" noted above. Mentioning also the Hilbert spaces over the complex numbers that appear in quantum mechanics, he claims a "miracle", which is a different kind of miracle from the high accuracy that was the basis of the earlier examples. It is the "miracle" that highly abstract concepts invented by pure mathematicians for largely aesthetic and formal reasons later prove to be the right ones for physics. He writes "It is difficult to avoid the impression that a miracle confronts us here [...] [this] has no reference to the intrinsic accuracy of the theory" (Wigner 1960, p.7).

Wigner thus advances two different kinds of mathematical "miracle" for which we have no explanation. One is the recurrent discovery that concepts of advanced mathematics are the right language for physics; the other is the extraordinary accuracy of mathematically-described physical laws, well beyond the experimental data that first suggested them.

Wigner's formulation of the "gap" between pure mathematical concepts and the physics that they later apply to is driven by his formalist philosophy of mathematics (Ferreirós 2017), which he explains—rather lightly and imprecisely—early in the paper in a section *What is mathematics?*. "Mathematics", he says, "is the science of skillful operations with concepts and rules invented just for this purpose [...] Most more advanced mathematical concepts, such as complex numbers [...] were so devised that they are apt subjects on which the mathematician can demonstrate his ingenuity and sense of formal beauty" (Wigner 1960, p.2). That is not the only possible philosophy of mathematics, and it is one that tends to place a large gap between mathematics and any science of reality such as physics.

Craig's argument from mathematics to God

William Lane Craig argues that Wigner's talk of "miracle" should be taken literally. There is something genuinely miraculous or objectively surprising in the applicability of mathematics to physics, and the best explanation of that is the existence of a miracle-worker, God. His argument relies mainly on the second of the questions above, on the relevance of advanced rather than elementary mathematics, and thus on Wigner's argument that what is miraculous is the applicability of higher mathematical concepts invented for other purposes.

Craig summarises his extension of Wigner's argument as follows:

(1*) Mathematical concepts arise from the aesthetic impulse in humans and have no causal connection to the physical world.

(2**) It would be surprising to find that what arises from the aesthetic impulse in humans and has no causal connection to the physical world should be significantly effective in physics. Therefore, it would be surprising to find that mathematical concepts should be significantly effective in physics.
(3*) The laws of nature can be formulated as mathematical descriptions (concepts) which are often significantly effective in physics.
(4**) Therefore, it is surprising that the laws of nature can be formulated as mathematical descriptions that are often significantly effective in physics.
(5) The fact that the laws of nature can be formulated as mathematical descriptions that are often significantly effective in physics merits explanation.
(6) Theism provides a better explanation of the fact that the laws of nature can be formulated as mathematical descriptions that are often significantly effective in physics than does atheism (Craig 2021, p.209).

Craig plainly intends here an objective concept of "surprising". Like Wigner, he says little about his notion of the aesthetic, but both authors intend it to cover the development of interesting pure mathematical concepts unrelated to experience.

Craig's philosophy of mathematics is not formalist like Wigner's, but (almost) Platonist. Like formalism, however, Platonism sees a large gap between the objects of mathematics, "abstract" Platonic entities like numbers and sets that exist in a non-physical, non-causal realm, and physical reality. As Craig says, Platonism in itself tends to make any applicability of mathematics seem surprising, as there is no way that Platonic objects can constrain what happens in the physical world:

> For the *non-theistic* [Platonist] realist, the fact that physical reality behaves in accord with the dictates of acausal mathematical entities existing beyond space and time is, in the words of philosopher of mathematics Mary Leng, "a happy coincidence". For consider: If, *per impossibile*, all the abstract objects in the mathematical realm were to disappear overnight, there would be no effect on the physical world. This is simply to underscore the fact that abstract objects are causally inert. The idea that realism somehow accounts for the applicability of mathematics "is actually very counterintuitive," muses Mark Balaguer. "The idea here is that in order to believe that the physical world has the nature that empirical science assigns to it, I have to believe that there are causally inert mathematical objects, existing outside of spacetime," an idea which is inherently implausible (Craig 2021, p.204).

Craig's alternative is to place the abstract objects in the mind of God, who is of course not causally inert: "the *theistic* realist can argue that God has fashioned the world on the structure of the mathematical objects He has chosen" (Craig 2021). Craig thus takes the Platonist mathematics-reality gap as an opportunity for God to bridge the gap.

It is true that Craig adds some remarks that point to a different, design-type, argument. He writes:

> What remains wanting on naturalistic anti-realism is an explanation *why* the physical world should exhibit so elegant and stunning a mathematical structure

> in the first place. After all, there is no necessity that a physical world exist at all, in which case mathematical truths would not have been descriptive of the physical world. Perhaps the universe, in order to exist, had to have *some* mathematical structure—though couldn't the world have been a structureless chaos?—but that structure might have been describable by elementary arithmetic (Craig 2021, p.206), elaborated in (Craig 2023).

It is indeed contingent whether the universe should have an "elegant and stunning" mathematical structure rather than a simple one, and a design argument could arise from that. But that is a different argument from the one of Wigner as reconstructed by Craig.

Aristotelian realist philosophy of mathematics

Aristotelian realism is an alternative philosophy of mathematics to nominalism (including formalism) and Platonism. It sees mathematics as directly about the world, a science of certain aspects of physical (or any other) reality, such as quantitative and structural aspects. It thus sees the Wigner "unreasonable effectiveness" problem as largely an artefact of those philosophies of mathematics which place a wide gap between mathematics and physical reality. On an Aristotelian view, mathematics (whether elementary or advanced) describes truly the structure of the world, so of course it is effective. It is no different from the effectiveness of crop science in growing crops, because crop science states the facts about crops.

Islami (Islami 2016) well describes how "gap" philosophies of mathematics see the problem:

> What seems to be puzzling is the underlying difference between the *relata* of this relationship: physics is the study of inanimate nature, concerned with the discovery of laws of nature, while mathematics is the study of concepts (structures) and operations, which seems to be far removed from the empirical study of the natural world (Islami 2016, p.4839).

The most popular philosophies of mathematics do see it as "far removed" from the natural world in that way. Platonism (one version of which is favoured by Craig) regards mathematics as a study of another world, that of acausal "abstract objects". Nominalist views see mathematics as divorced from the natural world because it is not about anything at all, but either just a manipulation of formal symbols (as in the formalism favoured by Wigner), or pure logic or merely the "language of science", not itself contentful.

However, those perspectives do not exhaust the field of philosophy of mathematics, and, as is widely recognised, they all face serious difficulties explaining the applicability of mathematics, the very problem raised by Wigner (Körner 1960, ch.8). The main alternative, which places applied mathematics first, is Aristotelian realism (Franklin 2014a, 2022a) (It is inspired more by Aristotle's general realism about universals than by his specific remarks about mathematics). According to it, mathematics studies certain aspects of the natural (non-abstract) world, quantitative ones (such as ratio) and structural ones (such as symmetry). The ratio of one animal's height to another is a relation that is as real as the heights themselves, but ratio is studied by the pure mathematics of continuous quantity. Symmetry (whether exact or approximate) is a real and easily perceivable feature of physical objects such as faces and crystals, but is studied by a branch of abstract mathematics, group theory. Surprisingly, Wigner himself won his Nobel Prize in physics for his work on the role of symmetry in quantum mechanics, and expressed his understanding of it in realist terms (before taking up formalist philosophy of mathematics). He wrote, in terms very acceptable to Aristotelians:

> Against the group-theoretic treatment of the Schrödinger equation, one has often raised the objection that it is "not physical." But it seems to me that a conscious exploitation of elementary symmetry properties ought to correspond better to physical sense than a treatment by calculation (Ferreirós 2017, pp.66–67).

The case of group theory and symmetry is one example of a wider phenomenon that some historically-informed commentators on Wigner have noticed—that he exaggerated the lack of physical motivation for the concepts of modern higher mathematics. Many of them developed organically from real-world motivations, though of course gradually going beyond them. The ultimate point of them was to deepen our understanding of the mathematical structures that are perceivable in the physical world (Grattan-Guinness 2008; Lützen 2011).

So Aristotelian realism sees mathematics as being directly about the properties of the real world (or sometimes properties that could be realised in the real world but may not be in fact, such as infinities). But it is not an empiricism. As in other fields, Aristotelians emphasise necessities, and the knowability of necessities. For example, ratios are in the world and are easily measurable, but the alternation of ratios is necessary and knowable (that for any non-zero lengths, weights, time-intervals etc, A, B, C, D, if the ratio of A to B equals the ratio of C to D, then the ratio of A to C equals the ratio of B to D).

A classic example provides a template of how Aristotelian realists see mathematical necessities in physical reality. In the eighteenth century, the seven bridges of Königsberg, in East Prussia, connected two islands and two riverbanks as shown in the diagram (Fig. 1).

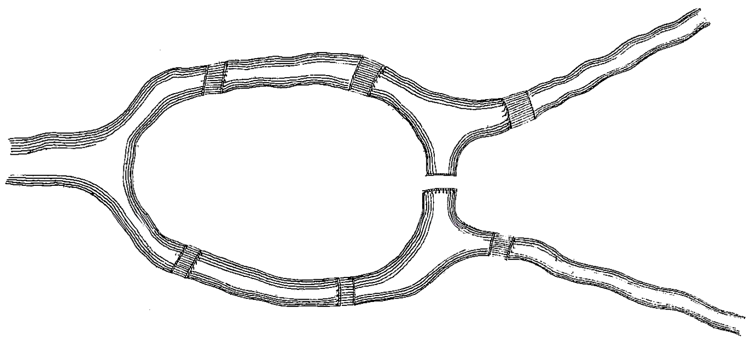

Figure 1. The bridges of Königsberg.

The citizens of Königsberg noticed that it seemed to be impossible to walk over all the bridges once, without walking over at least one of them twice. The mathematician Euler proved they were correct (Euler 1735; Räz 2018). Although new bridges could be built, nevertheless, with the existing bridges and land areas being as they are, it is absolutely impossible to walk over all the bridges once and once only. That is a necessity about the actual physical bridges and actual physical walking, not about some abstract model. The mathematical necessity is not in some Platonic world of abstract objects, or in some formalism, language or logical structure, but in the topological structure of the system of physical bridges. That necessity is inherent in the world and cannot be either imposed by or excepted from by God.

That kind of necessity is absolute and not the same as the "nomic necessity" of laws of nature. It is usually thought that God could have made the laws different so their necessity is less than absolute (and so, in principle there could be arguments to God from the choice of laws, but those arguments would be forms of design argument).

Nor does Aristotelian realism hold that the universe has its individual mathematical properties of necessity. The number of electrons in the universe is a contingent matter, although that number must satisfy all arithmetical properties.

On Aristotelian realism, there is in principle no problem with mathematics being "unreasonably" effective. Mathematics, whether elementary (mainly concerned with quantity) or advanced (mainly concerned with pure structure such as symmetry) describes possible structures that physical reality may have. It is for empirical science to measure whether reality has those structures (for example, whether space is Euclidean), and it is then for mathematics to draw the necessary consequences.

It is not surprising if initial observations only suggest a structure approximately, as in the case of Newton's measurements of gravity or Heisenberg's rules for calculating with hydrogen. If intelligent scientists guess from that the true structure, such as the inverse square law of gravity or the matrix mechanics (or mathematically equivalent) structure of quantum mechanics, that structure will give rise to indefinitely accurate predictions, just because it is the actual structure. Hence there is nothing surprising, in principle, either in the applicability of elementary mathematics, in the applicability of advanced mathematics (that is, mathematics that happens to be harder for humans), or in the high accuracy of predictions.

It is significant that the phenomenon of hyper-accurate prediction that gave rise to the third question above about the mathematics-reality fit occurs in elementary mathematics, and does so in a way that is provably comprehensible and so not a surprise to a mind with enough cognitive power. The digits of π, the ratio of circumference to diameter of any circle, have been calculated to over 100 trillion places. Each digit generates a prediction: that if a physical object is made more and more perfectly circular, the measured ratio of circumference to diameter will eventually be as that digit predicts. To a mind of limited cognitive power, that is amazing, but it is an absolute necessity that is fully comprehensible and does not require or admit of any outside power, divine or otherwise, to make it so.

That case is a model for understanding the necessity of application of the more esoteric concepts of higher pure mathematics, such as Hilbert spaces or Borel sets. A smaller proportion of humanity may understand them than understands the digits of π, but that does not make their necessity in applications any less comprehensible in itself, or any more in need of explanation from outside.

Mathematical necessity and theodicy

The essential motivation for Craig's argument is that theists should prefer a surprising, hence contingent, fit between mathematics and physical reality, in order to postulate God as a cause of the observed fit. There is however a reason why theists should prefer the fit between mathematics and physical reality to be absolutely necessary, namely, that it makes the problem of evil solvable along the lines of Leibniz's theodicy (Franklin 2022b).

Leibniz agrees that God ought to create the best possible world. He poses that as an upfront design problem: "God has ordered all things beforehand once for all, having foreseen prayers, good and bad actions, and all the rest" (Leibniz 1710, p.par.9, 128). As engineers are well aware, design is difficult, and purely logical and mathematical obstacles are among those that must be faced—perhaps, for an omnipotent God, the only ones that must be faced. The best-known of logical problems is the difficulty of reconciling free will for humans with an acceptably good outcome, as granting humans free will leaves God's world design hostage to humans' decisions to do evil.

The relevant mathematical constraints on God's action are those arising from the global-local distinction, one of the great themes of mathematics (Franklin 2014b). It is frequently found that what is possible locally—in some small region—is not possible globally, as the local solutions cannot be fitted together. The Königsberg bridges are an example. In any small part of the system, it is easy to find a path over all the bridges exactly once. But those paths cannot be fitted together to make a path over all seven bridges exactly once. What is possible locally is impossible globally, and the impossibility is mathematical and absolute, hence not subject to exception by God. Although God could make bridges or walking differently, he could not, while keeping the bridges and walking as they are, make it possible to walk over all the bridges exactly once.

Examples are found across many mathematical fields. They include the impossibility of building a circular staircase that goes up all the way round and ends at its starting point: any part can be made ascending, but the parts cannot be fitted together the solve the problem. Tuning and temperament in music involves mathematically unavoidable tradeoffs in dividing the octave into a small number of notes, tradeoffs which are resolved adequately but imperfectly in the standard system of equal temperament. The well-known Arrow impossibility theorems show that certain natural conditions on social choice cannot be satisfied simultaneously. The phenomenon is well-known to mathematical practitioners in many fields, though not to philosophers.

Such examples show what is wrong with the natural thought that it is easy to imagine God's improving the world by tinkering here and there, for example by removing a single toe-stubbing while leaving everything else as it is (Brown and Nagasawa 2005). There is no world in which a toe-stubbing is removed but everything else stays as it is. For a start, causes must act differently in that world, since the actual causes preceding the pain do not produce the same pain, so laws of nature are different. Changes in the world are not isolated but ramify globally.

If we consider from that perspective God's design of creation before starting, taking into account whatever knowledge of the future in different scenarios is possible for omniscience, it becomes clear that purely mathematical constraints make the design problem very difficult. According to the results of physics used in the "fine tuning" argument, God must choose the physical constants of the universe extremely close to their actual values to enable the conditions for life to arise. The tuning needed to do anything more detailed must be even more exact.

Leibniz's best-of-all-possible-worlds theodicy has not been considered credible in recent times. Modern developments in the mathematics of local-global interactions show that it deserves revival. In order to revive it, theists must come to appreciate how mathematical necessities constrain God. That is a way of thinking at odds with the Craig-Heller arguments, which look for contingencies requiring an explanation by divine creation.

Conclusion

Design arguments of all kinds arise from contingent aspects of the universe, which may require divine action to explain them. That is true of classical arguments from teleology or from the adapted complexity of living things, which are unexpected in a purely materialist universe and so may call for an external causal explanation. It is equally true of contemporary fine-tuning arguments, which begin with the apparent contingency of the settings of basic physical constants. As they apparently could easily have been different, their having just the unlikely combined settings that make our existence possible calls for an external causal explanation, such as by divine action. There could in principle be an argument from the universe having one or other mathematical structure which it need not have had, for example from its being finite or infinite, but that is not the kind of argument being considered here.

However, there is no such argument purely from the mathematicity of the universe or the effectiveness of mathematics in describing it. Mathematical truths are absolutely necessary, whether elementary or advanced, pure or applied. The digits of π are what they are necessarily, so they are not subject to God's will and there is no possibility of his writing a message in them. Similarly, the mathematical truths found to be instantiated in the physical universe cannot point to any contingent facts about creation, because they could not have been otherwise. Therefore, their being as they are cannot point to God.

References

Bronstein, David. 2016. *Aristotle on Knowledge and Learning: The Posterior Analytics.* Oxford: Oxford University Press.

Brown, Campbell, and Yujin Nagasawa. 2005. The best of all possible worlds. *Synthese* 143 (3): 309–320. https://doi.org/10.1007/s11229-005-7043-y.

Coyne, George, and Michael Heller. 2008. *A Comprehensible Universe: The Interplay of Science and Theology.* New York: Springer.

Craig, William Lane. 2021. The argument from the applicability of mathematics. In *Contemporary Arguments in Natural Theology: God and Rational Belief,* edited by Colin Ruloff and Peter Horban, 195–215. London: Bloomsbury.

Craig, William Lane. 2023. *Reasonable Faith Podcast: James Franklin Critiques Applicability of Math Argument.* https://www.youtube.com/watch?v=sRz3d5sDgc8 accessed June 30, 2025.

Euler, Leonhard. 1735. Solutio problematis ad geometriam situs pertinentis. In *Graph Theory 1736–1936,* edited by Norman Biggs, E. Keith Lloyd, and Robin Wilson, 3–8. Oxford: Oxford University Press.

Ferreirós, Jose. 2017. Wigner's "unreasonable effectiveness" in context. *Mathematical Intelligencer* 39 (2): 64–71. https://doi.org/10.1007/s00283-017-9719-9.

Franklin, James. 2014a. *An Aristotelian Realist Philosophy of Mathematics: Mathematics as the Science of Quantity and Structure.* Basingstoke: Palgrave Macmillan. 10.1057/9781137400734.

Franklin, James. 2014b. Global and local. *Mathematical Intelligencer* 36 (4): 4–9.

Franklin, James. 2022a. Mathematics as a science of non-abstract reality: Aristotelian realist philosophies of mathematics. *Foundations of Science* 27:327–344. https://doi.org/10.1007/s10699-021-09786-1.

Franklin, James. 2022b. The global-local distinction vindicates Leibniz's theodicy. *Theology and Science* 20 (4): 445–462. https://doi.org/10.1080/14746700.2022.2124481.

Grattan-Guinness, Ivor. 2008. Solving Wigner's mystery: the reasonable (though perhaps limited) effectiveness of mathematics in the natural sciences. *Mathematical Intelligencer* 30 (3): 7–17.

Heller, Michael. 2019. How is philosophy in science possible? *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 66:231–249.

Islami, Arezoo. 2016. A match not made in heaven: on the applicability of mathematics in physics. *Synthese* 194 (12): 4839–4861. https://doi.org/10.1007/s11229-016-1171-4.

Kessler, Volker. 2022. Mathematics declaring the glory of God. *Verbum et Ecclesia* 43 (1): 1–7. https://doi.org/10.4102/ve.v43i1.2432.

Körner, Stephan. 1960. *The Philosophy of Mathematics: An Introductory Essay.* London: Hutchinson.

Kuksewicz, Zygmunt. 1988. The potential and the agent intellect. In *The Cambridge History of Later Medieval Philosophy,* edited by Norman Kretzmann and Anthony John Kenny. Cambridge: Cambridge University Press.

Leibniz, Gottfried Wilhelm. 1710. *Theodicy: Essays on the Goodness of God, the Freedom of Man and the Origin of Evil.* Translated by E.M. Huggard. London: Routledge / Kegan Paul.

Lewis, Geraint F., and Luke A. Barnes. 2016. *A Fortunate Universe: Life in a Finely Tuned Cosmos.* Cambridge: Cambridge University Press.

Lützen, Jesper. 2011. The physical origin of physically useful mathematics. *Interdisciplinary Science Reviews* 36 (3): 229–243. https://doi.org/10.1179/030801811X13082311482609.

Räz, Tim. 2018. Euler's Königsberg: the explanatory power of mathematics. *European Journal for Philosophy of Science* 8 (3): 331–346. https://doi.org/10.1007/s13194-017-0189-x.

Trombik, Kamil. 2025. The Idea of Mathematicity of the Universe as a Locus Theologicus: Michael Heller and Representatives of "The Kraków School of Philosophy in Science". *Religions* 16 (1): 54. https://doi.org/10.3390/rel16010054.

Velupillai, Kumaraswamy Vela. 2005. The unreasonable ineffectiveness of mathematics in economics. *Cambridge Journal of Economics* 29 (6): 849–872. https://doi.org/10.1093/cje/bei084.

Wigner, Eugene. 1960. The unreasonable effectiveness of mathematics in the natural sciences. *Communications on Pure and Applied Mathematics* 13:1–14.

Philosophical Problems in Science (2025), **78**, 27–40
(Zagadnienia Filozoficzne w Nauce)
doi:10.59203/zfn.78.709

ARTICLE

Tractability assumptions and derivational robustness, a match made in heaven?

Gareth Fuller

University of Maine at Presque Isle
Email: gareth.fuller@maine.edu

Abstract

The epistemology of modeling in science is a topic that has received no shortage of attention from both philosophers and scientists. Models, with their ubiquitous unrealistic and false assumptions, raise a myriad of interesting epistemic questions. Many have picked up a concern expressed by biologist Richard Levins, which is that, for many models, it is unclear if any particular result is a product of the realistic assumptions, or if the result, is in some critical way, dependent on an unrealistic assumption. His suggestion of a solution to this problem, the use of robustness analysis, has similarly been picked up by many as a plausible solution to this concern.

Since Levins' time, taxonomies of modeling assumptions and types of robustness analyses have been developed. One common connection made is that of tractability assumption, or assumptions introduced for the purpose of mathematical tractability, and derivational robustness analysis, or a robustness analysis that tests the influence of assumptions on the derivation of some result. Derivational robustness analysis is often singled out as being particularly well-suited for resolving the epistemic concerns introduced by tractability assumptions.

In this paper, I argue that tractability assumptions do not present a consistent set of epistemic concerns. Given this, derivational robustness analysis cannot be used to resolve the epistemic concerns raised by all tractability assumptions. I use this to motivate some concerns about how well-suited the taxonomy that includes tractability assumptions is to discussions about the epistemic concerns of unrealistic modeling assumptions.

Keywords: scientific modeling, robustness analysis, tractability assumptions

1. Introduction

The epistemology of modeling in science is a topic that has received no shortage of attention from both philosophers and scientists. Models, with their ubiquitous unrealistic and false assumptions, raise a myriad of interesting epistemic questions. Many have picked up a concern expressed by biologist Richard Levins (1966), which is that, for many models, it is unclear if any particular result is a product of the realistic assumptions, or if the result is in some critical way dependent on an unrealistic assumption. His suggestion of a solution to this problem, the use of robustness analysis, has similarly been picked up by many as a plausible solution to this concern (Weisberg 2006; Weisberg and Reisman 2008; Kuorikoski, Lehtinen, and Marchionni 2010; Raerinne 2013).[1]

Since Levins' paper, taxonomies of types of modeling assumptions have been developed, often distinguishing types of unrealistic assumptions (Musgrave 1981; Mäki 2000; Hindriks 2006; Cartwright 2007; Kuorikoski, Lehtinen, and Marchionni 2010). The taxonomy of modeling assumptions I will work with includes the categories of substantial assumptions, which are realistic assumptions, and then galilean and tractability assumptions, which are unrealistic. Just like modeling assumptions, taxonomies of robustness analyses have also been developed since Levins' paper (Woodward 2006; Weisberg and Reisman 2008; Raerinne 2013). One particular type of robustness analysis, Derivational Robustness Analysis (DR), has been singled out as being particularly useful for resolving concerns with tractability assumptions (Kuorikoski, Lehtinen, and Marchionni 2010; Raerinne 2013). Given this connection between DR and tractability assumptions, I focus on these two specifically.

The main focus of this paper is to argue that understanding when to apply DR should be separated from the taxonomy of modeling assumptions that includes substantial, galilean, and tractability assumptions.[2] Rather, I argue that understanding the kinds of assumptions that DR can play an important epistemic role, particularly that of resolving concerns about unrealistic assumptions, should focus on particular epistemic features of those assumptions. These epistemic features do not align well with the taxonomy of modeling assumptions commonly used and so this taxonomy is not appropriate for understanding the role of DR. I attempt to do this by showing that tractability assumptions do not consistently capture these epistemic features.

My paper will proceed as follows. In section 2 I introduce DR, focusing on the conditions where it can resolve epistemic concerns of particular modeling assumptions. In section 3 I discuss how the literature singles out tractability assumptions as a unique target for DR. In section 4 I look to several examples of tractability assumptions, arguing that they do not fit well with the required epistemic conditions for the use of DR. I also present an example where DR might be applied to other types of assumptions. In section 5 I draw two general conclusions for the ongoing debates about robustness analysis and the epistemology of models. In section 6 I provide some concluding remarks.

2. Strategies for Resolving Some Epistemic Concerns for Models

In his paper *The strategy of model building in population biology*, biologist Richard Levins outlined the ubiquity, and even necessity, of unrealistic "simplifying" assumptions in models (Levins 1966). The inclusion of unrealistic assumptions, even if necessary, raises some epistemic concerns about these models. Levins raised one particular concern, which is that, for many models, "There is always room for doubt as to whether a result depends on the essentials of a model or on the details of the simplifying assumptions" (Levins 1966, p.423). A result that depends on particular details of simplifying assumptions, rather than the realistic ones, gives us reason to question the value of that result or the model. For instance, if we have a model of a ball rolling down an inclined plane, and this model tells us the ball will have certain velocity at the bottom of the ramp, but the model does not include friction, we might question whether or not this result can be applied to real inclined planes. The result may depend on the lack of friction, which would mean the result does not apply to planes with friction. The general concern is that, If some result is dependent on the unrealistic simplifying assumptions, this can give us reason to discount the result or the model itself.

This concern has been raised by many, and one straightforward solution to this problem is to simply replace the unrealistic assumption with a more realistic assumption. In the case of the frictionless plane, it would include adding friction into the model to make the model more realistic. If the result of the model changes, then we know that the result was dependent, in some way, on the lack of friction. This process of making a model more realistic is the process of de-idealization or concretization (Peruzzi and Cevolani 2022).

While de-idealization is straightforward, it has several drawbacks. First, many models include a great many simplifying assumptions, and it could very well be the case that there are simply too many simplifying assumptions to apply this to. Levins nots that, at least with population biology models, to carry out the process of idealization such that the model lacks any simplifying assumptions, would require a model "using perhaps 100 simultaneous partial differential equations" and this would lead to models that would be too insoluble and the results of which would have no real meaning to us (Levins 1966). Further, it is not the case that all assumptions in a model can be de-idealized, since the true vale of some parameters are not known or are vaguely defined (Levins 1966).

Given these concerns about de-idealization, Levins' solution to this problem is a method known as robustness analysis. In robustness analysis,

> We attempt to treat the same problem with several alternative models each with different simplifications but with a common biological assumption. Then, if these models, despite their different assumptions, lead to similar results we have what we can call a robust theorem which is relatively free of the details of the model (Levins 1966, p.423).

Essentially, the strategy of robustness analysis is to develop several models with a shared "core" of realistic assumptions about the causal mechanism of interest, but different unrealistic or simplifying assumptions. If all of these models produce the same result, or a "robust theorem", then this is meant to show that the robust theorem is derived from the shared core assumptions rather than any unrealistic assumption. This avoids the concerns with de-idealization since none of the models used in robustness analysis are completely de-idealized but rather incorporate *different* simplifying assumptions.

De-idealization and robustness analysis differ in several ways. The most obvious one is that de-idealization focuses on creating more and more realistic models, while robustness analysis does not focus on producing more realistic models, but just different, false models. Epistemically, de-idealization works to boost credence in the idealized model because the less realistic model agrees with a more realistic one. Ultimately, however, once the more realistic model has been developed, the less realistic one is no longer needed. When it comes to robustness analysis, the boost in credence stems from the fact that the result of a model does not depend on any particular set of simplifying assumptions, but, rather, seems to be dependent on the shared, realistic, set of assumptions. Each model in the set for robustness analysis is needed.

Since Levins' paper, a variety of types of robustness analysis have been developed (Woodward 2006; Weisberg and Reisman 2008; Raerinne 2013).Woodward has developed an influential taxonomy that includes Inferential Robustness (IR) and Derivational Robustness

(DR) (Woodward 2006). IR is a method applied when there is a data set and its relation to a particular hypothesis. It is often the case that data alone is not enough, but some further assumptions are needed to derive any conclusions. In some cases, it could be that there are several competing sets of further assumptions. IR would be the process of using the data set with each of these different sets of competing assumptions to see if the results stay the same or change. If the result is robust, or the same, across the different competing sets of extra assumptions, then this can give us reason to accept the result. If the result changes across these competing sets, then we have reason to withhold judgment on the particular hypothesis.

Woodward discusses IR in relation to several different disciplines, but considers how it might be applied in climate modeling (Woodward 2006). Someone who is trying to estimate the annual increase in average ocean temperatures due to greenhouse gasses, will require data about ocean temperature increase, but will also need a climate model to use with this data. If there are several competing climate models, IR would be the process of using the data with each if the climate models and comparing results. If the result stays the same (or within some range of similarity) then this should give us reason to believe the result. However, if the data produce different results with the different climate models, *and* there is no reason to accept one climate model as better than the others, then we cannot draw conclusions about the impact of greenhouse gasses on ocean temperatures from the data.

DR, on the other hand, is a method of determining the influence of particular assumptions in a model by altering the assumption to see how this influences the derivation of a result. As Woodward puts it,

> Suppose that we have a model or theory that allows for the derivation of observed facts P. Suppose the model contains some assumption A, which might concern, e.g., the value taken by some parameter x, or the relationship between two parameters x and y, where A figures in the derivation of P. It is natural to ask how sensitive this derivation is to the assumption of A (Woodward 2006).

The strategy of DR, then, is to alter assumption A and to see how this impacts the derivation of P. If altering A still produces P (or a similar enough result), then it can be said that P is a "derivationally robust" result of the model, or at least derivationally robust in relation to A and its alternatives. Woodward's DR varies slightly from Levins' robustness analysis in that it is not only applied to false or simplifying assumptions. It can be applied to any assumption of the model to check its relative influence on some result.

Now, DR is closely related to both de-idealization and IR in a variety of ways. IR and DR can be distinguished in that IR focuses on testing the "inductive warrant" a data set has provides for some conclusion (Woodward 2006). DR, on the other hand, is not about determining inductive warrant of a data set, but rather about checking the relative influence of a certain assumption on the results of a model. While IR focuses on using the same data with different theoretical models, DR focuses on using the same theoretical model but different values of a parameter to check its relative influence.

DR and de-idealization differ in several ways as well. As mentioned above when comparing de-idealization and robustness analysis as Levins presented it, de-idealization requires that some assumption in a model can be made more realistic while robustness analysis does not. This holds for DR, since the focus is simply on altering some parameter, not necessarily making it more realistic. Further, they function somewhat differently epistemically and find different applications. In de-idealization, it is the more realistic model that does the epistemic heavy lifting. However, when it comes to DR, there is no single value of the altered assumption that does the heavy lifting. It is the set that provides an epistemic boost.

This difference can be made clearer by considering the conditions where a result being robust or not provides an epistemic boost. In particular, Woodward held that failure of a result to be derivationally robust only sometimes raises concerns about a model, depending on what else is known about the phenomenon being modeled, while a result failing to carry over through the process of de-idealization does raise concerns. For instance, if it is known that there is only a single realistic value for some parameter in a model, and this value is stable, *and* this value is known, then showing some result is derivationally robust in relation to multiple values of this parameter would not provide much support for the model (Woodward 2006). What we would be most interested in is what the model does with the known, stable value. This sort of situation, where the value of the parameter is stable and known would be more suitable for de-idealization (assuming the parameter of interest was unrealistic to begin with). For DR, it is when there are multiple equally realistic values of some parameter that some results being derivationally robust provides support for a model. Similarly, it is when there are multiple realistic values for some parameter that failure to be derivationally robust in relation to this parameter raises some concerns about the model. When there are multiple, equally realistic values of some parameter, moving from a less realistic to more realistic assumption cannot be done, and so de-idealization cannot be applied.

Woodward provides as an example of when failure of derivational robustness raises some red flags is with social preference models in economics. The success of these models is highly sensitive to values chosen for a few crucial variables (Woodward 2006). However, there are two key aspects about these kinds of models that make failure of robustness concerning. First,

> (a) there is no real evidence that supports these particular choices of parameter values rather than other choices. Arguably, no choices for these values are consistent with all the data and insofar as the particular values chosen are consistent with some range of experimental results, this is equally true of other possible choices (Woodward 2006).

There are multiple equally acceptable values of the parameters which cannot all be captured at once. If the result of a model is tied to a proper subset of these, then this can indicate a concern for the model. Second,

> (b) we have independent reason to think that the parameter values in question (or their distribution) are unlikely to be stable across different people, across contexts, or perhaps even across time for different people and we have no theory which allows us to predict or understand such variation (Woodward 2006).

Not only are there multiple equally acceptable values of the parameters, but it is also believed that, for any individual, the "real-world" value of the parameter varies across time and context. Once again, if the model is incapable of capturing this variation, this can raise some concerns.

As an example of when a failure of robustness raises *no* concern about a model is that of the inverse square law. The correct predictions about the behavior of the planets produced by the inverse square law fails to be robust across a range of values for the gravitational constant and the exponent of the distance term in the law (Woodward 2006). This, however, is of little concern for the model. For instance, the gravitational constant only takes a single value and this value does not vary across individual planets or a range of contexts. Ultimately, there is only a single realistic or appropriate value for the gravitational constant to take, and so the fact that different values of the gravitational constant fail to accurately predict the behavior of planets is of little concern.[3] What we care about is that the correct results are produced with the correct value.

Very generally, the conditions under which failure of derivational robustness raises no concerns are when "we have a great deal of independent evidence both that the parameter is highly stable across a range of different circumstances, that it does indeed take the claimed value and that we can use it, in conjunction with existing theory, to derive a wide range of observed phenomena" (Woodward 2006, p.232). In this case, if the known, stable parameter is incorrect in the model, we would apply de-idealization. Under conditions where there are multiple equally acceptable values of some parameter, and this value does not vary or fluctuate across contexts and individuals, then failure of robustness *can* indicate a concern.

3. Derivational Robustness and Tractability Assumptions

Woodward's account of DR has been influential and has found development and application to models working with certain taxonomies of modeling assumptions (Kuorikoski, Lehtinen, and Marchionni 2010; Raerinne 2013).[4] In this section, I examine two accounts of DR as applied to a taxonomy of modeling that includes three types of modeling assumptions, substantial, galilean, and tractability assumptions. I outline these two accounts to show that they see the same epistemic value

of DR as Woodward and arguments for why DR is applied to just tractability assumptions. In later sections, I will argue that this pairing between DR and tractability assumptions needs some reconsideration.

Kuorikoski et al. (2010) look to extend the discussion of the role of DR, highlighting two different, albeit related, roles. First, they hold that DR can be used to bolster a belief in a model because "it guards against error by showing that the conclusions do not depend on particular falsehoods" (Kuorikoski, Lehtinen, and Marchionni 2010, p.534). This is a concern in line with concerns raised by Levins and captured in Woodward's discussion. However, they also hold that "it confirms claims about the relative importance of various components of the model by identifying which ones are really crucial to the conclusion" (Kuorikoski, Lehtinen, and Marchionni 2010, p.543). The first of these is in-line with the role that Woodward outlines for DR, where it can highlight potential concerns raised by particular assumptions. The second is a more about exploring the causal mechanism of interest.[5]

The first role also closely follows Levins' concern about unrealistic assumptions. It is recognized that models incorporate assumptions that are known to be unrealistic or false. For at least some of these assumptions, there is a concern that falsehoods may play a significant role in the derivation of some result. By altering these false assumptions, and showing that some result is derivationally robust across these alterations, this can assuage some fears about the influence of any *particular* false assumption. The second role is important since it allows modelers to learn more about the causal relations of the real-world target by examining how they interact in the model.

These roles for DR, however, are embedded in a framework that distinguishes different types of assumptions, a taxonomy that is not part of Woodward's focus. In particular, Kuorikoski et al. and their interlocutors share a taxonomy of modeling assumptions that includes three general types; substantial, galilean, and tractability assumptions (Kuorikoski, Lehtinen, and Marchionni 2010; Odenbaugh and Alexandrova 2011; Raerinne 2013; Lisciandra 2017; Harris 2021). In brief, substantial assumptions are the assumptions of a model that are intended to represent the causal features of interest, picking out the features of the real-world target that the modelers are interested in studying. Galilean assumptions are assumptions made for a model that omit certain causal features of the real-world target from the model. These omitted causal features are those that are expected to be "causal confounders" for the substantial assumptions. Galilean assumptions, then, are assumptions about omission of some real causal features so that the substantial assumptions can be better investigated. Finally, tractability assumptions are assumptions made for a model that are simply done for mathematical tractability.

As an example, we can look at a Lotka–Volterra model.

$$\frac{dV}{dt} = rV - (aV)P$$

$$\frac{dP}{dt} = b(aV)P - mP$$

It is a pair of differential equations used to represent the relationship between a predator and prey population. In this model, $V(t)$ stands for the size of the prey population at time t, $P(t)$ the size of the predator population, the constant r stands for the growth rate of the prey population, the constant m stands for the death rate of the predator population, the constant a stands for the predator attack rate, and the constant b stands for the predator conversion efficiency.

An example of a substantial assumption in this model is that the predator and prey populations are negatively coupled. When the prey population size increases, the predator population size increases while when the predator population increases, the prey population size decreases. An example of a galilean assumption is that there are no influences on the size of the prey population, outside of the predator population. This means that factors like environmental carrying capacity are not found in the model, despite there being carrying capacities in the real-world. These sorts of influences are omitted for the purpose of better understanding how predator and prey populations influence each other. Finally, an example of a tractability assumption would be the *specific* functional form chosen to represent rate of prey capture per predator relative to the size of the prey population. For this model, it is assumed that there is a linear increase in rate of prey capture as the prey population increases (Harris 2021). This is a tractability assumption since any chosen functional form is technically false of any real-world target, but necessary to use the differential equations that simplify modeling.

Within this taxonomy, the galilean and tractability assumptions are unrealistic or false assumptions while the substantial assumptions are (hopefully) realistic. Given that substantial assumptions are intended to capture the causal relations that the modelers want to study, it is hoped that these assumptions accurately capture features of the real-world target. Galilean assumptions, on the other hand, are assumptions that omit some real-world features that might get in the way of studying the substantial assumptions and tractability assumptions can misrepresent the target in a variety of ways.[6]

An important aspect of this taxonomy is that it distinguishes different types of unrealistic assumptions, and provides an opportunity to evaluate how these differences might influence the interpretation of models.[7] Galilean assumptions, being made to omit some causal features, are intended to have a clear causal interpretation

(Kuorikoski, Lehtinen, and Marchionni 2010). This clear causal interpretation means that it is clear how to replace galilean assumptions with a more realistic assumption (i.e. de-idealizing the omitted causal feature). However, tractability assumptions are introduced for reasons of mathematical tractability and "in many cases have no empirical merit on their own" (Kuorikoski, Lehtinen, and Marchionni 2010, p.548). Given this, it is not always clear what it would mean to replace a tractability assumption with a more realistic assumption. Further, the falsehoods introduced by galilean assumptions are often suspected of making some difference to results of the model, given that they omit a causal confounder. However, for tractability assumptions, it is hoped that the falsehoods they introduced have no impact on the results of the model (Kuorikoski, Lehtinen, and Marchionni 2010).

It is the fact that models incorporate unrealistic assumptions like galilean and tractability assumptions that leads Kuorikoski et al. to point out that one of the roles of DR is to guard against error by showing that some result does not depend on a falsehood. However, given their differences, this epistemic role is limited to tractability assumptions. As Kuorikoski et al. point out that "it is only the failure of robustness with respect to tractability assumptions that is epistemically problematic" (Kuorikoski, Lehtinen, and Marchionni 2010, p.548). While they hold that DR can be applied to all types of assumptions, a result being derivationally robust, or failing to be derivationally robust, in regards to a substantial or galilean assumption often "suggests a new empirical hypothesis about a causally relevant feature in the modelled system", which does not, on its own, raise any concerns about the model (Kuorikoski, Lehtinen, and Marchionni 2010).

However, failure of robustness in regards to some tractability assumptions is problematic "because it suggests that the result is an artefact of the specific set of tractability assumptions" (Kuorikoski, Lehtinen, and Marchionni 2010, p.548). The falsehood introduced by tractability assumptions is considered and is treated differently from galilean assumptions because tractability assumptions are supposed to lack a clear empirical interpretation. A failure of derivational robustness cannot be used to generate a new empirical hypothesis about the real-world target because of a failure of robustness with tractability assumptions, unlike galilean assumptions. Further, tractability assumptions are hoped to be negligible to the results of the model, while galilean assumptions are suspected of being causally significant and non-negligible. This is why *failure* of robustness is treated differently between tractability and galilean assumptions, despite both being false.

Similarly, that, "for many tractability assumptions it is often unclear what it would mean to replace them with more realistic ones [...] This is why tractability assumptions are often replaced with assumptions that are also unrealistic, but in a different way" (Kuorikoski, Lehtinen, and Marchionni 2010, p.548), means that they fit into the conditions outlined by Woodward. Since it is unclear what it would mean to replace one tractability assumption with a more realistic one, all potential options are on par, and other strategies like de-idealization are not possible. There are multiple equally acceptable, albeit all false, options. Kuorikoski et al. lay this out, almost exactly as Woodward does, when they state that "For most economic phenomena of interest there might not be a single true functional form, fixed over time, against which the exact form of the assumptions could be compared" (Kuorikoski, Lehtinen, and Marchionni 2010, p.546). It is here that we see agreement between Woodward and Kuorikoski et al on applying DR and its value being in cases where there is no single, best option for the value of a parameter. This, on Kuorikoski et al.'s account captures conditions that are unique to tractability assumptions.

That tractability assumptions are the appropriate target of DR is not only held by Kuorikoski et al. Raerinne, for instance, argues that DR should be considered a method that *solely* focuses on the influence of tractability assumptions, while robustness applied to substantial (and galilean) assumptions is known as "Sufficient Parameter Robustness" (SPR) (Raerinne 2013). This distinction, once again, comes down to epistemic differences. Tractability assumptions, being false, can introduce error into a model. However, in some disciplines, such as biology which is Raerinne's target, there is often redundancy of causes or causal mechanisms, where multiple mechanisms might produce the same effect. The point of SPR is to find a sufficiently abstracted set of parameters that are able to capture the shared effect across these redundant causal mechanisms (Raerinne 2013, p.299). Once again, tractability assumptions are well suited for DR because of the unique epistemic threat they are supposed to pose.

In the literature on DR, tractability assumptions are often singled out as unique. While there are several differences between tractability and galilean assumptions, the significant difference is that tractability assumptions present the conditions that Woodward outlines as those where DR can be used to identify epistemic concerns. For tractability assumptions, it is often understood that there are multiple, equally acceptable, options and no clearly most realistic opportunity. In the next section, I challenge the extent to which this is true.

4. Applying Derivational Robustness to Tractability Assumptions

As discussed in section 3, DR can play two related roles, one to guard against error and the other to determine which assumptions are crucial to the derivation of some result. While these are related, the concern about "guarding against error" is focused specifically on tractability assumptions. However, in order for DR to play this role, the assumption being tested must meet certain conditions. In particular, there must be no, single, correct variation of the assumption.[8] If there is a single correct variation of the assumption, then the model can be de-idealized to ensure that the false assumption is not causing problems. In this section, I intend to show that tractability assumptions, as a category, do not neatly fit into this condition. I will do so by examining examples of tractability assumptions from the literature, showing that several examples of tractability assumptions do not fit this condition. The upshot, in section 5, will be that we should separate discussions of the epistemic import of DR from the common taxonomy of substantial, galilean, and tractability assumptions.

I start with Kuorikoski et al.'s own example of a tractability assumption, that of *specific* functional forms. I will then examine another common example, that of representing populations continuously rather than discreetly (Colyvan and Ginzburg 2003; Colyvan 2013;

Lisciandra 2017). Finally, I consider Hindrik's example of a tractability assumption, the "single planet assumption" in Newton's models of the solar system.

We can start with Kuorikoski et al.'s own example of a tractability assumption, that of a *specific* functional form chosen for a model. However, Kuorikoski et al. present this under conditions that very closely match those required by DR. As quoted in section 3, they draw on examples from economics, and state that, "For most economic phenomena of interest there might not be a single true functional form, fixed over time, against which the exact form of the assumptions could be compared" (Kuorikoski, Lehtinen, and Marchionni 2010, p.546). This lines up well with what Woodward has to say about social preference models.[9] In these conditions, the exchanging or altering of a functional form does not make a model more or less realistic. Rather, each functional form is limited in how much of the real-world target it captures. Given this, applying DR to the functional form of a model seems to be a prime example of such models.

Yet it is not clear that this carries over to all such examples of functional form as a tractability assumption. Harris, discussing the Lotka-Volterra model while raising a different concern about robustness, claims that:

> Although it may be true that any assumed functional form will strictly speaking be false for any real-world predator-prey system, there is certainly a sense in which one particular functional form might be more adequate to describe the rate of prey capture than another, despite both of them being strictly false (Harris 2021, p.14592).

That all tractability assumptions might, technically, be false of a real-world target does not mean that they are all "equally unrealistic". There could, very well, be a functional form that *best* captures the relationship being modeled among the variety of options for functional forms available. For instance, in the example of the Lotka-Volterra model from section 2, it was pointed out that the functional form for rate of prey capture per predator was linear. For any real-world population, it is unlikely that the rate of prey capture will be exactly linear. However, a linear rate of prey capture might best capture the trend or most closely approximate real-world populations.

Now, if this is the case, that there may be some instances where there is a best or most adequate functional form, this does not immediately undercut the value of DR in this case. However, if it is also *known* what the most adequate functional form is, even if it is false, then we would not need or want to apply DR. Rather, modelers would want to apply de-idealization, if the most adequate functional form was not used originally. In Woodward's example of the inverse square law, the fact that some result is lost when the value of the gravitational constant is changed raises no concerns because the model has been made less accurate (Woodward 2006, p.232). Rather than being able to chalk up the loss of the result to its dependence on a particular, false, assumption, we can chalk up the loss of the result to the fact that the new value deviates from the correct value. If there is a most adequate functional form, then the fact that some correct result is lost does not necessarily indicate a fault in the model. Rather, one could chalk up the loss of the correct result to the fact that an inferior functional form was chosen, just like altering the value of the gravitational constant.

This concern about some tractability assumptions does not end with a subset of functional forms. There are a variety of other types of tractability assumptions which are not amenable to the application of DR. Another common example is that of representing populations continuously rather than discreetly (Colyvan and Ginzburg 2003; Colyvan 2013; Lisciandra 2017).[10] Colyvan and Ginzburg discuss the Lotka-Volterra models, and, as they say in footnote 9:

> The use of differential equations in population dynamics is clearly an idealisation; populations are discrete and so the appropriate mathematical machinery is really *difference* equations. Differential equations, however, continue to be used, mainly for their convenience (Colyvan and Ginzburg 2003).[11]

The use of differential equations introduces a tractability assumption. It is an unrealistic assumption in the model, that populations are continuous, that is made for the purposes of mathematical tractability or simplicity.

However, notice that the conditions surrounding this assumption do not meet the conditions outlined for DR to play the guarding role it is intended to. When it comes to this sort of misrepresentation, there is very clearly a more appropriate or correct assumption to make, namely one that allows populations to be modeled discreetly. The conditions where there are multiple, equally unrealistic, options and no single best option is not met by this situation. This kind of tractability assumption simply does not fit into the required conditions.

Another example of a tractability assumption comes from Hindriks (2006). Hindriks' example is that of Newton's model of the solar system. In particular, "Newton assumed that there were no interplanetary gravitational forces. Given his theory, a planet moves around the sun in an ellipse under the assumption that it is the only planet orbiting the sun" (Hindriks 2006). Now, at first this may appear to be a galilean assumption because it is omitting some causal feature. However, the purpose for this omission is not to remove a causal confounder. Rather, the reasons for this assumptions was that:

> The mathematics needed to take account of the gravitational effects of other planets was only developed some time after Newton's death. So Newton was not even in a position to do without the single-planet assumption. This means, I suggest, that the problem of determining the movements of a planet in relation to the sun was intractable for Newton without the single-planet assumption (Hindriks 2006).

The single-planet assumption was one made out of mathematical necessity, since the model could not be constructed otherwise.

Now, just like the example of continuous populations, there is a clear way to correct this assumption. There is a most appropriate assumption to go in place of the single-planet assumption, namely multiple planets and their respective gravitational forces. As far as meeting the conditions laid out for when DR can play a guarding role, this sort of tractability assumption fails. Further, just like the example of functional form where there is a most adequate, a variation of results between the continuous population and a discrete population, or the single planet assumption and one that captures inter-planetary gravitational forces, can be chalked up to the fact that one assumption is more accurate than the other, rather than it simply being an "artifact" of a particular tractability assumption. Essentially, the conditions that Woodward outlines are necessary for DR to play the role of guarding against error, and these conditions are not captured by all tractability assumptions.

5. Why Does This Matter?

Given what I have stated above, there is a very fair question to ask at this point. Without even debating over my proposed examples of tractability assumptions that do not fit the mold, it is fair to ask what the point of this is. I have not established that DR *cannot* play the role of guarding against error and artifacts, because I have accepted Kuorikoski et al.'s point that there are cases where different functional forms are all equally adequate. I do not even want to limit this possibility to just functional forms. Those who argue that DR can be used to assuage fears about tractability assumptions can, with little concern, accept that not all tractability assumptions can be subjected to DR, but DR can still play the role they have laid out for many tractability assumptions.

This pushback is fair, but I think there are two important points to highlight from my discussion. First, and what I see as a less significant point, is that it is always helpful to get a bit more clarity on the methods being used. Knowing that DR cannot be applied to *all* tractability assumptions is useful to know, so that correct applications of DR is carried out. I want to be clear, my intention has *not* been to argue that DR is incapable of playing the epistemic role outlined by Woodward, Kuorikoski et al., or Raerinne. I fully accept that it can play this sort of role. This, however, leads to my more significant point about the value of the taxonomy of modeling assumptions commonly used, and the close connection between DR and tractability assumptions.

This more significant point regarding the use of the class of tractability assumptions. The development of tractability assumptions came out of an exchange across several decades and authors, where the focus was on categorizing the different types of unrealistic assumptions that might show up in a model (Musgrave 1981; Mäki 2000; Hindriks 2006). However, tractability assumptions, as they are understood in Kuorikoski et al. (Kuorikoski, Lehtinen, and Marchionni 2010) and their interlocutors are primarily categorized by the reason that they were introduced into a model, namely for reasons of mathematical tractability.[12] In this way, it makes perfect sense to group the various assumptions together, since this can provide insight into why certain unrealistic assumptions are included into a model and even why models often include so many unrealistic assumptions.

However, categorizing assumptions by the reason they were introduced into a model does not necessarily line up with the epistemic concerns raised by those assumptions. An important difference in the epistemic concerns raised by assumptions seems to be tied to whether or not they can be corrected or de-idealized, if the correct value is known, or if there is a most adequate or correct value for the assumption. Tractability assumptions, as I have tried to argue, are split between some that have a most adequate or correct value and some that do not. In this way, they do not raise consistent epistemic concerns. Given that tractability assumptions do not form a consistent epistemic set in this regard, they may be ill-suited as a category for debates about the epistemic practices of modeling.

This concern about using tractability assumptions can extend to the overall taxonomy that includes substantial and galilean assumptions. Once again, I do not want to deny that this taxonomy has any use. Classifying assumptions by the role they play in a model, such as representing core causal features or omitting causal confounders, can be helpful for better understanding the method of constructing models, particular methods used within specific sciences, as well as why models often end up with certain kinds of misrepresentations. However, once again, this does not mean that the categorizations line up with the epistemic concerns raised by the type of assumption. In particular, the epistemic conditions under which DR can play an important epistemic role are not ones that are clearly a part of any single type of modeling assumption. Rather, talk of the value of DR and when it should be applied should be divorced from the taxonomy of modeling assumptions that is commonly used.

When discussing the epistemic concerns raised by unrealistic assumptions, a taxonomy that better aligns with the features that seem to raise distinct epistemic concerns is superior. In particular, a taxonomy that better captures the difference between the assumptions where there are multiple, equally acceptable options and those where there is a single option.

6. Conclusion

In this paper I have argued that the close connection between DR and tractability assumptions is somewhat tenuous. I have not denied that DR can be used to assuage fears about tractability assumptions. Rather, I have generally looked to establish where DR can be usefully applied to tractability assumptions, as well as evaluated the common taxonomy of assumption types that includes tractability assumptions. I have argued that tractability assumptions do not make an epistemically consistent set, and that the taxonomy that includes tractability assumptions may be of limited use in assessing the epistemic concerns raised by models.

Notes

1 It has also faced its fair share of criticism. See: (Orzack and Sober 1993; Odenbaugh and Alexandrova 2011, 2011; Lisciandra 2017; Harris 2021).

2 I am not arguing that this taxonomy has no value, but when it comes to the epistemology of modeling, I believe it fails to capture what is important.

3 Woodward goes on to point out that this is, in fact, an advantage of the model.

4 Kuorikoski et al. (2010) highlights the role of DR in economic modeling while Raerinne (2013) focuses on biological modeling.

5 Raerinne, despite providing a role for DR, wants to distinguish a different type of robustness analysis for the second role that is called "Sufficient Parameter Robustness". I will discuss this a bit later in this section.

6 I will discuss a few examples in section 4.

7 Musgrave (1981), while not using this taxonomy, provided distinctions of different types of unrealistic assumptions to pursue some debates about economic models, most prominently Milton Friedman's defense of the use of unrealistic assumptions in economic models (Friedman 1966).

8 A further condition to *apply* DR is that we must be able to actually alter the assumption in question. If we cannot alter the assumption, false or not, then we cannot apply DR.

9 An example from economics, to be clear.

10 Raerinne discusses the opposite where "In order to make a model more manageable, we may choose to model a process that appears to change continuously over time, such as population growth, by means of a discrete equation instead of a differential equation" (Raerinne 2013). This example from Raerinne will still run into the same issues that I discuss below.

11 Italics in the original.

12 Lisciandra (2017), in footnote 8 on p. 86, notes Hindriks has a view of tractability assumptions that requires that they remove aspects of the real-world target that are suspected of being non-negligible.

References

Cartwright, Nancy. 2007. *Hunting Causes and Using Them: Approaches in Philosophy and Economics.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511618758.

Colyvan, Mark. 2013. Idealisations in normative models. *Synthese* 190 (8): 1337–1350. https://doi.org/10.1007/s11229-012-0166-z.

Colyvan, Mark, and Lev Ginzburg. 2003. The Galilean turn in population ecology. *Biology and Philosophy* 18 (3): 401–414. https://doi.org/10.1023/A:1024121002194.

Friedman, Milton. 1966. The Methodology of Positive Economics. In *Essays In Positive Economics,* 3–43. Chicago: University of Chicago Press.

Harris, Margherita. 2021. The epistemic value of independent lies: False analogies and equivocations. *Synthese* 199 (5): 14577–14597. https://doi.org/10.1007/s11229-021-03434-8.

Hindriks, Frank A. 2006. Tractability assumptions and the Musgrave–Mäki typology. *Journal of Economic Methodology* 13 (4): 401–423. https://doi.org/10.1080/13501780601048733.

Kuorikoski, Jaakko, Aki Lehtinen, and Caterina Marchionni. 2010. Economic Modelling as Robustness Analysis. *The British Journal for the Philosophy of Science* 61 (3): 541–567. https://doi.org/10.1093/bjps/axp049.

Levins, Richard. 1966. The strategy of model building in population biology. *American Scientist* 54 (4): 421–431.

Lisciandra. 2017. Robustness analysis and tractability in modeling. *European Journal for Philosophy of Science* 7 (1): 79–95. https://doi.org/10.1007/s13194-016-0146-0.

Mäki, Uskali. 2000. Kinds of Assumptions and Their Truth: Shaking an Untwisted F-Twist. *Kyklos* 53 (3): 317–335. https://doi.org/10.1111/1467-6435.00123.

Musgrave, Alan. 1981. Unreal Assumptions in Economic Theory: The F-twist untwisted. *Kyklos* 34 (3): 377–387.

Odenbaugh, Jay, and Anna Alexandrova. 2011. Buyer beware: Robustness analyses in economics and biology. *Biology & Philosophy* 26 (5): 757–771.

Orzack, Steven Hecht, and Elliott Sober. 1993. A critical assessment of Levins's The strategy of model building in population biology (1966). *The Quarterly Review of Biology* 68 (4): 533–546.

Peruzzi, Edoardo, and Gustavo Cevolani. 2022. Defending De-idealization in Economic Modeling: A Case Study. *Philosophy of the Social Sciences* 52 (1-2): 25–52. https://doi.org/10.1177/00483931211049759.

Raerinne, Jani. 2013. Robustness and sensitivity of biological models. *Philosophical Studies* 166 (2): 285–303. https://doi.org/10.1007/s11098-012-0040-3.

Weisberg, Michael. 2006. Robustness Analysis. *Philosophy of Science* 73 (5): 730–743. https://doi.org/10.1086/518628.

Weisberg, Michael, and Kenneth Reisman. 2008. The Robust Volterra Principle. *Philosophy of Science* 75 (1): 106–131. https://doi.org/10.1086/588395.

Woodward, Jim. 2006. Some varieties of robustness. *Journal of Economic Methodology* 13 (2): 219–240. https://doi.org/10.1080/13501780600733376.

Philosophical Problems in Science (2025), **78**, 41–55
(Zagadnienia Filozoficzne w Nauce)
doi:10.59203/zfn.78.720

Philosophical Problems in Science

ARTICLE

“Philosophy in science” and “science as philosophy”. Some selected philosophical views of Małgorzata Głódź

Kamil Trombik
The Pontifical University of John Paul II in Krakow
Email: kamil.trombik@gmail.com

Abstract
This article attempts to reconstruct and analyze some selected philosophical views of the Polish physicist Małgorzata Głódź, especially within the context of her links with the interdisciplinary milieu in Kraków (known as the “Kraków School of Philosophy in Science”). It emphasizes the significant role that Głódź played in the historical development of this intellectual tradition. The main portion of this article highlights the connections between Głódź’s philosophy and the philosophical ideas propagated by Michał Heller, the School’s main founder, and his students and other collaborators. The article also indicates elements of her philosophical achievements that deserve attention from the perspective of today’s philosophical challenges.

Keywords: philosophy in science, incomplete philosophy, Kraków School of Philosophy in Science, Małgorzata Głódź, Michał Heller

1. Introduction

Małgorzata Głódź (born 1946) is one of a group of physicists who engage in philosophical pursuits alongside their scientific work. She developed her philosophical interests primarily through her interactions with thinkers in Kraków, though she conducted scientific research at institutions in Warsaw. At the beginning of her career, she studied at the University of Warsaw, where she obtained the MSc degree in physics in 1969. She continued to be associated with the Warsaw scientific community in the following years. Głódź was affiliated with the Institute of Physics at the Polish Academy of Sciences in Warsaw (*Instytut Fizyki Polskiej Akademii Nauk*) for a number of decades. She obtained her Ph.D. in physics in 1976 and gained habilitation in 1994. She specialized especially in laser spectroscopy (atom–laser–field interactions). Her achievements in the area of physics are unquestionable, but it seems reasonable to also consider her philosophical achievements. Significantly, Głódź has written papers of a strictly philosophical nature. More precisely, they deal with the border between science and philosophy[1].

This paper is intended as an initial attempt at reconstructing and analyzing selected elements of this philosophical legacy, especially in the light of Głódź's connections with the interdisciplinary milieu in Kraków. By this I refer to the so-called "Kraków School of Philosophy in Science" (KSPS), which has developed over decades through the activities of Michał Heller, his students, and other collaborators[2]. Despite her scientific activity in Warsaw, Głódź regularly appeared in Kraków, where her connections with the local community of philosophers and scientists deepened. She has cooperated with this milieu for many years and become a significant figure in promoting interdisciplinary dialogue between science and philosophy, both through her own publications and other scientific activities undertaken on behalf of the Center for Interdisciplinary Studies (*Ośrodek Badań Interdyscyplinarnych* or OBI)[3]. In this paper I will try to point out the connections between Głódź's philosophy and the philosophical ideas propagated by Heller and his close circle. I will also try to highlight valuable elements of her philosophy—valuable not just from a historical point of view, but also useful within the context of the currently challenges emerging at the nexus between science and philosophy. I will precede the main considerations on the views of Głódź with a historical outline of the OBI, emphasizing in particular the role of Małgorzata Głódź in the development of this research institution[4].

2. Małgorzata Głódź and the Center for Interdisciplinary Studies (OBI)

The OBI was founded as an informal research institution through the initiative of Michał Heller and Józef Życiński[5]. From the late of 1970s, both philosophers had been building an interdisciplinary milieu in Kraków that brought together philosophers and scientists (mainly physicists) with an interest in the relationship between science and philosophy, as well as science and religion (e.g., Sierotowicz 1997). Their endeavors included both scientific and didactic activities, as well as organizational aspects (see Heller 2006, p.49). During the OBI's initial activities, so-called interdisciplinary seminars (monthly meetings of philosophers and scientists) and publications (books and the journal Z*agadnienia Filozoficzne w Nauce*[6], which was founded by Heller and Życiński) were especially important for developing philosophy within the context of science among the Kraków milieu. Głódź was a significant figure in the OBI, with her participating in the initiatives undertaken of Heller's circle.

It is worth noting here that Głódź had been interested in Heller's views since the early 1970s, being one of the first person who notice Heller's scientific potential. In 1972, the journal *Znak* published fragments of Głódź's correspondence with Heller at the time. The correspondence concerned Heller's first book *Wobec Wszechświata*. The discussion, or rather polemic, between Głódź and Heller concerned certain aspects of the philosophy of science, with these being related to, among others, the nature of scientific research in the context of the relationship between experimental and theoretical physics (Głódź and Heller 1972)[7]. In the following years, Głódź and Heller continued their collaboration[8], and when the OBI became active, Głódź worked closely with the Kraków interdisciplinary milieu, co-creating it to some extent. Over time, the KSPS was built on the foundation of this activity of the OBI (Polak and Trombik 2022). The term "philosophy in science" refers to a specific style of practicing philosophy proposed by Heller. One of Heller's students and collaborators, Paweł Polak, characterized it as follows: „This philosophy was developed in an interdisciplinary dialogue between philosophy and science. The concept of «philosophy in science» had its origins in Heller's broad understanding of the philosophy of nature. Heller described it as a new philosophy for the philosophical interpretation of science" (Polak 2019, pp.266–267). For more about philosophy in science, see Section 3.

Głódź actively participated in the works of the OBI. She participated in events organized by this institution and was also known as the author of philosophical papers that were published on behalf of the OBI. Her publications included strictly philosophical works (Głódź 1992c, 1996a, 2006), reports from events organized by the OBI (Głódź 1987, 1992b, 1996c), and memoirs on various forms of Heller's and OBI's activities (Głódź 1999, 2012). Her papers were mainlny published in journal *Zagadnienia Filozoficzne w Nauce* and collective monographs devoted to Heller (in connection with his milestone birthdays, multi-author monographs were published that comprised papers by his students and colleagues: (Głódź 1996b, 2006). Some of Głódź's articles were also published in books signed by Heller and Życiński (Głódź 1988)[9]. It is also worth noting that she co-authored paper with Heller (Głódź and Heller 1983).

An important element of the OBI's activity in the 1980s was the translation and publication of important works (papers or fragments of books) by Western scholars that were often unavailable in the political realities of Poland at that time. These translations were published in *Zagadnienia Filozoficzne w Nauce*, as well as in other OBI publications. Głódź's contribution should also be appreciated here, because she translated, among others, works by Erwin Schrödinger, Pierre Teilhard de Chrdin, Charles Misner, and later—Polish astronomer and philosopher Andrzej Pacholczyk, who collaborated with the OBI (Teilhard de Chardin 1982; Misner 1983; Schrödinger 1983; Pacholczyk 1996)[10].

Głódź was deeply involved in the activities of the OBI and was keen to develop an interdisciplinary milieu that would focus on investigating the overlap berween science and philosophy[11]. She not only saw the potential of the OBI in this regard, but also identified with the philosophical program of Heller and his students, and openly expressed her affiliation with this movement. Głódź was also one of the first to notice the process of development of the School (although she did not use this word herself), pointing to the continuity of a certain tradition initiated by Heller and Życiński:

> This first group of enthusiasts, organizers and participants of the seminar[12], led by Michał Heller and Józek Życiński, called by Włodek [Włodzimierz] Skoczny the „undisputed pillars of the OBI", was supplemented over time by subsequent „generations" of their students from the Faculty of Philosophy of the Pontficial Academy of Theology [PAT], and later also students of students. Apart from the strict hardcore of the OBI, there was always a milieu of its more or less faithful supporters and co-workers from the academic milieu outside PAT (especially from the Jagiellonian University) and church circles (especially during the communist era, the publishing activity of the group was possible mainly thanks to Catholic editorial offices, also „church-related").[13]

Głódź's very valuable comments show how the KSPS was formed and how important it is to take into account a wide range of influence in the description of the phenomenon of this School: students of Heller and Życiński, the subsequent generation (i.e., the students of students), collaborators, supporters[14]. It should be noted that she herself also identified with the OBI group. The approach to practicing philosophy in the context of science, something that had formed in the Kraków milieu, was close to her: „The name «philosophy in science» is very relevant and close to my heart" she admitted in 1999, very positively assessing the activities of the OBI as „service to two masters: philosophy in science and cooperation between disciplines" (Głódź 1999, pp.16–17).

Declarations of this type are of great importance when discussing the history of the KSPS, because they prove that there were philosophers among Heller's collaborators who openly admitted being inspired by his philosophy and openly demonstrated their belonging to his circle. I will discuss this issue in more detail in Section 4, meanwhile, in the next part of the paper I will present Głódź's philosophical views against the backdrop of the views of Heller and the OBI circle.

3. Selected philosophical ideas of Głódź in relation to philosophy practiced within the KSPS

The project known as "philosophy in science", outlined by Heller in the 1980s (Heller 1986a)[15], was an important reference point for Głódź in her own philosophical papers. Before I delve further into this, I will give a few remarks on the concept of "philosophy in science", the roots of which lie in the 1970s (as exposed, e.g., by the name of the journal *Zagadnienia Filozoficzne w Nauce* [„Philosophical Problems in Science"], published on the basis of seminars organized by Heller and Zycinski since 1978).

Heller understood "philosophy in science" as a sort of philosophical reflection that continued in a sense what was formerly called the philosophy of nature. This "new" philosophy of nature—precisely as philosophy in science—was more innovative in nature than the traditional forms of this discipline that were widespread in Poland at the time of the OBI's foundation. As a type of non-systemic philosophy—in contrast to the neo-Thomistic philosophy of nature, developed within the framework of the Aristotelian-Thomistic system—it was focused on closer contact with the sciences and the latest philosophy of science. The subject of philosophy, when understood in this way, was to include both the historical and contemporary interactions between science and philosophy (including metaphysics), with an emphasis on the following areas[16]: the influence of philosophical ideas on the development and evolution of scientific theories; traditional philosophical problems intertwined with empirical theories; philosophical reflection over some assumptions of the empirical science (the assumption of the mathematicity of nature, the assumption of the idealizability of nature, the assumptions of an elementary character and the unity of nature).

The analysis of Głódź's philosophical achievements reveals that the area of her interests overlapped with the theme of "philosophy in science" as outlined by Heller, and more broadly—it was simply part of the activities of the wider OBI movement (Trombik 2021, pp.221–226). Although Głódź addressed various issues in her papers, they were all closely related to the problems that dominated within the OBI. It is worth adding that Głódź published these papers over several decades, showing that her philosophical interests were not incidental, but accompanied her throughout her entire scientific career.

Głódź's oeuvre includes papers on certain metaphilosophical aspects (i.e., reflection on philosophy and its connections with science, especially the discussion on the project of "philosophy in science"), contemporary philosophy of science (e.g., the factors influencing the development of science[17], the relationship between theory and experience), ontology (especially the issue of the mathematicity of the world, which was the major issue taken up by the KSPS[18]), the history of science (i.e., selected issues taken from the modern and contemporary history of science[19]). A separate category comprised papers on the border between science and theology, which although not the subject of this article's analysis, touches upon issues discussed by Heller, Życiński and other members of the OBI (Krauze 2008).

Faced with the question of how to practice philosophy in the context of the sciences, Głódź analyzed Heller's "philosophy in science" project several times. Her assessment of Heller's project and what OBI representatives did in practice to develop "philosophy in science" (as understood as a style of practicing philosophy) was positive. Like Heller, she noticed the numerous relationships between science and philosophy (Głódź 1999; 2006, p.370), and just like Heller and the OBI milieu, she did not approve of the rigid methodological isolationism of science and philosophy that was widespread in the views of some philosophers of nature. Głódź also shared the conviction that contemporary philosophizing in the context of the scientific image of the world necessitated rejecting the maximalist ambitions of the old philosophies of nature, which aimed to construct a system of knowledge that encompasses final answers to questions about reality. This was the result of both the fact of temporariness (openness to change and improvement) of the scientific image of the world, as well as more in-depth methodological and historical research, which reveals the nature of the development of knowledge about the world. In this context, Głódź wrote that „OBI philosophizing is not limited to posing problems, but the chances of finding ready-made solutions are assessed realistically. The methodology and history of science, as well as the history of philosophy—treated in the OBI with due attention—allow us to assume that the search for answers is, perhaps, a convergent process, but the result of individual searches will not be a final answer, but, at best, marking one's participation in the chain of human struggles" (Głódź 1999, p.19). Głódź's remarks in this aspect correspond to the declaration and views of Heller and other representatives of the OBI (Heller 1986a, 1998).

Considering the specifics of philosophy, Głódź clearly highlighted a need to consider its tasks and goals in close connection with our scientific image of the world. She therefore posed the question: „What general framework for philosophy is provided by the evolutionary style of thinking imposed by evolving scientific knowledge about evolving reality?" (Głódź 1996a, p.21). In response, she proposes the idea of an "incompleteness philosophy" ("*filozofia niedomknięta*"). Evolution would be inherent in the nature of philosophy understood in this way—in the sense that it would take over certain features of evolution from the process of developing scientific knowledge. Głódź clarified that „evolution excludes closed (final) systems. Evolution postulates «philosophy on the way». There is no reason for philosophy to fear such a state of affairs. The lack of closure in science is a condition for its development. In science, the metaphor of evolution is not indefinite fluidity. It is rather the spreading circles of a wave, encompassing earlier solutions (or perhaps also a metaphor for «evolutionary waste», removed by the logic of new knowledge) [...] The history of human thought gives the impression that—whether individual philosophers wanted it or not—philosophy is somehow open anyway. Why shouldn't it be consciously and programmatically unclosed? Science benefits from this" (Głódź 1996a, p.22).

Głódź's reflections on philosophy can be viewed as part of the project of philosophy practiced within the OBI circle, and fit very well with the idea of "philosophy in science" as an anti-foundationalism proposal for practicing philosophy, or more precisely: as an attempt to practice philosophy without indisputable philosophical foundations that constitute the basis of the theoretical system which is exhaustive and entirety of knowledge. Foundationalism can be understood as a conviction that „a philosophical system that is supposed to provide certainty knowledge and must be built on indisputable foundations" (Heller 1999b, p.84). "Philosophy in science" was intended as a philosophical project based on the results of science, a proposal constituting „a modern equivalent of the traditional philosophy of nature" (Heller 1998, pp.12–13). This modernity was to manifest, among other things, in a programmatic openness to changes triggered by the developing scientific image of the world. Głódź shares this view of philosophy and accepts, just like Heller, the necessity of an interdisciplinary approach to philosophical problems involved in science (Głódź 1999). "Incomplete philosophy" understood this way can, on the one hand, address traditional, great philosophical problems currently entangled in scientific theories while, on the other hand, follow the example of the sciences and focus on certain aspects of reality, ultimately contributing through detailed research to a deeper understanding of certain phenomena. Głódź considered this second aspect of philosophy and posed questions that could be treated as an encouragement to think through the foundations of philosophy when understood as philosophy in science: „philosophy that ignores a series of questions and analyzes only a fragment of reality is contemptuously considered as contributory. Is the contempt justified? Perhaps, at least in some cases, contributory philosophy is a good path to truth? Why is a philosopher afraid to admit: I am programmatically exploring this fragment, and I will probably never know about the other?" (Głódź 1996a, p.22).

An important element of the analyses conducted by Głódź on the philosophy practiced within Krakow's interdisciplinary milieu was the question of referring to the slogan "science as philosophy", which sort of complemented the idea of "philosophy in science" (Głódź 2006, p.370; Heller 1998; 2007, pp.189–191). At this point, Głódź also referred closely to Heller's ideas, proposing the foundations of her own view, which fits into the understanding of "science as philosophy"[20]. The slogan "science as philosophy" is understood in such a way that the sciences, by taking over an important part of the old functions of philosophy, carry information about what reality is like, and the process of interpreting this information has a philosophical value in the end. In the case of Głódź, it should be clarified that she means precisely physics as "science"; physics is the model of science for her (Głódź 2006, p.372). By "philosophy" she refers to ontology here, along with its important epistemological references.

In the case of epistemological issues, especially the problem of the relationship between scientific descriptions and reality itself, her position is close to realism (i.e., scientific realism). She is convinced by the "miracle argument" by John Worrall (Głódź 2006, p.372)[21]. The mathematical–empirical method of investigating the world is an effective procedure that allows us to discover certain properties of the world. It is worth highlighting here some of Głódź's remarks that are important from the perspective of the philosophy of science. Głódź has repeatedly emphasized the close interdependence of theory and experiment (experience). She considers the division into "theoretical physics" and "experimental physics", something that is apparent even in the organizational structures of research institutes, as being invalid and harmful (Głódź and Heller 1972, p.739). She justifies her own position on this matter with research practice itself: experimental physics requires a theoretical foundation, because an experiment cannot be described without referring to the theoretical layer[22]. According to Głódź, this is an expression of the fact that „reality is significantly richer than our common ideas about it. Understanding a number of its properties consists of «reading into» the description made in the language of mathematized theory" (Głódź 1992c, p.107).

The effectiveness of investigating the world using mathematical and empirical methods prompted Głódź to put forward an ontological view. Inspired by Heller's analyses of the ontological aspects of physics (Heller 1986b), she states that physical theory itself—as mathematical structure confirmed empirically—can be considered as „a specific ontological interpretation of the world" (Głódź 2006, p.371). A physical theory carries information about what the world is like, and therefore—as Głódź concludes—it actually carries information of an ontological nature, i.e. about the nature of being (Głódź 2006, p.376). In this sense, physical theory becomes philosophy and the slogan "science as philosophy" becomes legitimate (Głódź 2006, p.371).

Ontological considerations conducted in the OBI circle often led to discussions about the mathematicity of nature (Heller and Życiński 1990). Many scholars who descended from Heller's milieu have followed their inspirator in treating mathematicity as a property or feature of nature, precisely in the ontological sense (Życiński 1987). The problem of the relationship between mathematics and the world itself is also a key element of Głódź's philosophical considerations. This represents an interesting example of the practical implementation of the "philosophy in science" project. It is significant that her position on the issue of mathematicity of nature is also similar to the positions shared by Heller, Życiński and other members of OBI.

Głódź believes that the mathematizability of nature (i.e., the possibility of describing nature through mathematics) is a fundamental methodological assumption of science (Głódź 1996b, p.56). Nevertheless, she did not stop at this obvious fact but rather tried to address the philosophical question about the property of nature itself, which is described mathematically. She sees the possibility of moving from epistemology (i.e., a mathematical description of nature within the context of its investigation) to ontology (mathematicity as a property of nature that is investigated). Głódź seems to have rejected ideas close to phenomenalism, writing that „the discovered regularities actually concern the structure of the world, not our mental states" (Głódź 1996b, p.57). She directly suggests that mathematical structures fit the world, thus raising the legitimate question of whether they are some kind of reflection of the structure of nature.

The correspondence between the structures of mathematics and the physical world is not "pure", i.e. mathematical theory does not describe fragments of reality in all its complexity, but, due to idealization, it only provides an approximate image of a fragment of reality. It is worth emphasizing that Głódź, following the ideas of Heller's project of "philosophy in science", treated the idealizability of nature as a basic assumption of science that one should reflect upon. Heller directly pointed to the idealizability of nature as a key assumption of science that requires philosophical reflection (Heller 1986a, pp.17–18). Głódź conducted such reflection, showing that the possibility of applying idealization in the description of nature allows for approximating its ontological structure through the structures of mathematics used in sciences. At this point, Głódź returns to the interdependence of theory and experience, emphasizing that „in the preparation of a piece of the world so that it is subjected to the research procedure, i.e. idealization, both experience and theory participate, because everything happens [...] in a network of interdependencies" (Głódź 1996b, p.53). It is also worth noting Głódź's remarks on what she called the "mathematicity of experience". In assuming the mathematicity of nature, „we also assume the mathematicality of that part of it which constitutes experience in physics" (Głódź 1996b, p.57). Experience is a part of reality, and its course is subordinated to "Mathematics-Structure" (Głódź 1996b, p.48).

In this case, Głódź adopts the Heller's distinction between mathematics with a small „m" (i.e., mathematics as applied in science) and Mathematics with a capital „M" (i.e, as a property of nature itself). In some fragments of her papers, Głódź seems very close to Platonism in its structuralist version, which Heller also approved of. According to her, scientific practice itself (or rather the successes of mathematized sciences) testify to the fact that the world has a certain property that allows for the reconstruction of its structure. The laws of nature discovered by scientists concern the very structure of the world; mathematical structures correspond to these structures of the world (i.e., its ontology).

In these ontological considerations, Głódź revealed herself as a scientist keenly interested in certain philosophical consequences of scientific discoveries. According to Głódź, practicing science is something that essentially opens one up to metaphysical issues and even further, theological issues (Głódź 1993, p.36) in the sense that awareness of the possibilities (but also certain limitations) of science creates a framework for reflection of a theological nature (e.g., on the nature of God). In relation to religion, science—as a key element of contemporary culture—should serve as an important source of theological inspiration. Głódź openness to issues on the border between science and religion is another significant feature of intellectual formation related to KSPS[23], characteristic also of the "philosophizing scientists" associated with the School, such as Andrzej Fuliński or Leszek Sokołowski (Sokołowski 2011).

The above examples show how closely Głódź's philosophical position dovetailed with the idea of "philosophy in science" and how closely it fit with the philosophical profile of the OBI. Głódź's views should be discussed in more detail—which may happen in a more extensive work—to further cosider these similarities. I assume that the fragments of her papers presented here are representative of her greater philosophical output—in the sense that they highlight certain threads that are common to her and the KSPS, thus allowing to considered her as an prominent member of this School.

4. Some summary remarks about Głódź's philosophical activity

Głódź's philosophical works are not well known within the philosophical community. If we tried to measure her philosophical impact merely by citations, we could say that it is moderate. References to Głódź's works appear mainly within the Kraków milieu (Heller and Życiński 1990, p.121; Heller 1992, p.51; Wszołek 1994, pp.82–83; 1996, p.224; Rodzeń 2007, p.92; Stolarczyk 2018, p.271). On the other hand, the number of citations (especially at the current stage, when Głódź's achievements are only just being exposed) does not necessarily say anything about the quality of these achievements. The history of philosophy includes many examples where the value of a philosophical work was only appreciated years later. One of the goals of this paper was to show that Głódź's articles contain serious philosophical potential worthy of attention in our current reality. Any assessment of her achievement should be nuanced, however. For example, it is debatable as to what extent Głódź's ontological remarks contribute anything new to old disputes about the properties of the world. Nevertheless, certain issues related to the philosophy of science (such as the issue of the theory–experience relationship or the related issue of the "mathematicity of experience") deserve close attention, because Głódź's suggestions in this respect seem relatively original[24]. I believe that in this context, it is also worth emphasizing the fact that Głódź, as a scientist, emphasizes the value of philosophical reflections conducted under the light of scientific knowledge. In her works, one can find many valuable remarks on the understanding of philosophy and the relationship between philosophy and science. These were not just empty declarations, because—as her own achievements show—she put into practice the idea of "philosophy in science".

Głódź was a promoter of interdisciplinary research, within which philosophy plays an important role. She herself participated in such a dialogue between science and philosophy for many years, working within the OBI, where the postulate of interdisciplinarity was put into practice: „I understand interdisciplinary research as the cooperation of specialists of various specialties. However, from the point of view of one specialist, each specialist from a different field is placed in an external position. How can philosophy in science be practiced interdisciplinarily? Probably only in the sense that a scientist must be practiced «interdisciplinary» in at least two fields: he must be suspended between his scientific specialization and philosophy, i.e. in addition to the «scientific sense», he should have «a sense of philosophical perception of problems», as well as philosophical education. It seems, however, that philosophical issues in science cannot be studied in an interdisciplinary way—understood broadly. Meanwhile, the many years of achievements of the OBI contradict this statement" (Głódź 1999, pp.16–17). Głódź is convinced that if philosophers are not ignorant of science, and if scientists do not treat philosophy as a set of meaningless statements, they can meet in creative discussions. According to her, the OBI was an example of such a milieu, one in which interdisciplinary dialogue took place in a creative and fruitful way.

I believe that these remarks regarding the need for interdisciplinarity in philosophy are so currently pressing that they are worth recalling. It is undoubtedly worth emphasizing, for historical reasons, that they are not entirely innovative, however. Even when only considering the OBI milieu, they are actually repetitions of what Heller had been expressing since the 1970s. On the other hand, Głódź's reflections on the so-called "contributory philosophy" seem very promising. In the OBI milieu, the value of philosophical research on certain detailed issues entangled in scientific theories (or from a slightly different perspective, on fragments of reality that were philosophically "penetrated" in attemps to understand them a little better) was eagerly emphasized. If such contributory research is not an end in itself, but simultaneously constitutes an essential basis for reflection on the so-called great questions of philosophy, then it can significantly enrich the traditinal branches of philosophy, and protect philosophers from naive generalizations and excessive speculation. The history of philosophy confirms that such a threat has lingered at various stages. The Kraków School attempted to remedy these threats, and Głódź's position on this matter could have been one of the important voices pointing to the value of such a "contributory" approach, which, after all, does not exclude metaphysical problems from the area of philosophical reflection.

These examples show that in terms of the approach to practicing philosophy, the ideas of Heller and the entire OBI community were very close to those of Głódź. The project of practicing philosophy in science, as an interdisciplinary project in principle, was intended to constitute a new form of philosophy of nature. According to Heller and his collaborators, philosophy in science should not be limited solely to the analyzing the language of science or scientific methods, nor should it merely generalize the results of science. Heller tried to go beyond the relatively narrow understanding of philosophy that was characteristic of, for example, certain forms of positivism. In this respect, the OBI milieu preferred a more maximalist approach, and an expression of this maximalist attitude was openness to metaphysical problems. In other words, philosophy in science was in practice a project of anti-naturalistic philosophy. Heller's metaphysical interests were also shared by his students and colleagues, including Głódź (e.g., her ontological considerations on the properties of nature, which can be interpreted as a manifestation of "science as philosophy"). These interests sometimes even led them to take up issues on the nexus of science, philosophy, and theology. On the other hand, the project of philosophy in science, despite these maximalist tendencies, also had its minimalist profile—in the sense that it was a non-systemic philosophy, not pretending to construct a holistic image of the world (this was, for example, a frequent ambition of systemic philosophies, especially in the Thomistic style, something that Heller's milieu widely polemicized[25]). It was a philosophy conceived as an open project, or—as Głódź used to say—a philosophy that was programmatically open, "incomplete", a "philosophy on the way", the development of which was conditioned by the development of scientific knowledge. Głódź's philosophical achievements show that this balance between minimalist tendencies (in the sense of non-systemic philosophy) and maximalist tendencies (in the sense of a certain anti-naturalism, or at least anti-scientism[26]) was apparent in her papers, and dovetailed with what was generally characteristic of the OBI milieu.

The "philosophy in science" project was implemented in practice by Głódź, and in addition she openly admitted to being inspired by Heller's ideas and demonstrated her awareness of belonging to the Kraków interdisciplinary community (i.e., the OBI). These various factors, i.e., firstly—involvement in the activities of the OBI; secondly—the convergence of Głódź's interests with the research interests of Heller and his other collaborators; and thirdly—a similar (convergent with the project of "philosophy in science", which Głódź herself declared was close to her) way of philosophizing, allow us to consider Głódź as a representative of the Kraków School of Philosophy in Science.

Głódź's philosophical achievements are one of the examples of Heller's idea of "philosophy in science" finding fertile ground among representatives of science. In the context of post-war, 20th-century Polish philosophy, this was a phenomenon—i.e., the visible influence of philosophical ideas among scientists—unseen anywhere else on such a scale. Indeed, Głódź was not an unique figure from the OBI environment in co-creating a specific philosophical School while being active in science. There were many such philosophizing scientists in Heller's circle (e.g. Leszek Sokołowski, Andrzej Fuliński, Alicja Michalik, Andrzej Staruszkiewicz), which leads to the guess that the approach to practicing philosophy in the context of science, proposed by Heller, turned out to be very effective. Philosophical research is still needed to highlight (or deny) the deeper connections between the ideas of scholars such as Głódź and the ideas of Heller and the entire OBI milieu. Further historical research is also undoubtedly needed to shed light on other, possibly even sociological, factors in the development of the School, with an emphasis on the role played by representatives of the sciences in this process.

Importantly, some of Głódź's work holds importance for historical research, because it illustrates the historical development of the intellectual formation known as the KSPS. Głódź not only participated in this milieu as a philosophizing scientist, but she was also a keen observer of it. She was one of the first to clearly emphasize certain unifying factors that emerged among Heller's students and collaborators, which formed this philosophical milieu. Głódź wrote about certain factors as early as the late 1990s (Głódź 1999). This shows that identification with Heller's ideas and the OBI milieu appeared in some—as in Głódź's case—already over 25 years ago. In the context of further research, especially in relation to the discussion on the historical factors involved in the formation of the School (Polak and Trombik 2022), this seems to be of great importance.

Notes

1 At this point, it should be noted that Głódź has also published papers about the relationship between science and religion (e.g., Głódź 1982b, 1992a, 1993, 1994, 1997). Certain aspects of this issue in Głódź's works will therefore also be considered here. Nevertheless, due to the purpose of this paper (the emphasis on the relationship between science and philosophy in Głódź's works with the desire to compare her views with those of Heller and his collaborators), I will leave the subject of the relationship between science and religion in Głódź's works for another occasion.

2 For more information on KSPS, (see e.g., Polak and Trombik 2022).

3 For more on the OBI from a historical and philosophical perspective, see: (Heller 1999a; Heller et al. 1999; Heller 2006; Trombik 2019).

4 The quoted fragments of Głódź's works are translations by the author of this article.

5 Heller and Życiński were academics at the Pontifical Academy of Theology in Kraków (PAT). From the 1970s, they collaborated both on an institutional level (contributing to the creation of the Faculty of Philosophy at the PAT and organizing a local community of philosophers interested in the sciences) and on a philosophical level (initiating and developing the project of philosophy practiced in the context of the sciences). For more information on certain aspects of this cooperation, (see Trombik 2019).

6 The journal is currently called "Philosophical Problems in Science".

7 Głódź's comment on this discussion is significant: "I believe that the misunderstanding between us stemmed, at least in part, from slightly different «sources of admiration» in each of us. For me, the point of discussion about research was closer to the research itself. I like that humans have an effective method for penetrating reality. I didn't actually explicitly ask why this method is effective. For you, the beauty of physics lies one level lower – in the structure of the world itself, in the conviction that it is constructed in such a way that we can investigate it. For me, the beauty of physics lies in the fact that we can investigate the world, that we can decipher the code of its structure, and that when we apply it to reality, we find that it fits quite well. After all, no one said that enthusiasm must necessarily be aroused by causes, not effects" (Głódź and Heller 1972, p.741).

8 Głódź also followed Heller's publications and even wrote reviews of some of his books, in which she discussed various philosophical issues (Głódź 1982a, 2010).

9 Part of this chapter was previously published in the journal *Przegląd Powszechny* (Głódź 1984a, 1984b).

10 Teilhard de Chardin's text (translated by Głódź) was later also published in Heller and Życiński's book *Drogi myślących.*

11 It is worth adding that Głódź was a keen observer of the changes in Polish scientific culture that occurred in connection with the collapse of the so-called Eastern Bloc. For example, she noted with some concern that after 1989, interest in the relations between science and philosophy in Poland was decreasing (Głódź 1992b).

12 Głódź means here the interdisciplinary seminars that I wrote about above. See also: (Liana and Mączka 1999).

13 (Głódź 2012, p.5).

14 It is worth noting that Głódź treated the OBI not so much as a scientific institution, but rather as a group of enthusiasts deeply involved in conducting a dialogue at the interface of science and philosophy: „calling OBI a research and didactic institution, even when the Centre [OBI] was affiliated with PAT [the Pontifical Academy of Theology in Kraków], was, in my opinion, a slight exaggeration. What was happening there intensively was of a «pure» character, precisely outside the institution, untainted by subordination and bureaucracy, supported on the shoulders of consistent and systematic enthusiasm. On the other hand, what kind of institution is it that has no headquarters, no stable income, no paid employees. But despite this, or maybe, as I thought, thanks to the lack of bureaucratic burdens, it works brilliantly" (Głódź 2012, p.6).

15 English version: (Heller 2019). For more on the historical context of the origin of the idea of "philosophy in science", along with a discussion of this concept of practicing philosophy, e.g., (Polak 2019; Pabjan 2019).

16 "Philosophy in Science" initially primarily covered various philosophical issues arising in physics and cosmology. Over time, it also began to touch upon problematic issues of other disciplines: on the border of philosophy and chemistry, philosophy and biology, and in recent decades it has been used as a proposal for pursuing reflection on traditional epistemological and anthropological problems entangled in neuroscience (Dębiec 2006). The project of the so-called "philosophy in informatics" (or "philosophy in technology"), which focuses on the philosophical aspects of new technologies, has also proven innovative (Rodzeń 2007; Polak 2017, 2023; Krzanowski 2023).

17 For example, she took a position in the context of the classical problem of the philosophy of science, namely the issue of scientific development and the factors conditioning this development. Głódź shared the belief about the rationality of the development of knowledge (at least in physics), but she also emphasized the importance of sociological factors (Głódź 1996b, p.50).

18 See e.g., (Heller et al. 1999; Trombik 2025).

19 See e.g., (Mączka et al. 2012).

20 Although Głódź referred to the idea of "science as philosophy", she understood it somewhat differently than Heller's declarations in some of his works (Heller 2007), emphasizing the informational content of scientific theories themselves rather than their philosophical interpretations. In this sense, one could say that her understanding of "science as philosophy" is somewhat narrower than in some of Heller's works, although consistent with the content of his other works (Heller 2006).

21 It is worth noting that the "miracle argument" in support of scientific realism has been a subject of interest for OBI scholars, especially Jacek Rodzeń (Rodzeń 2005, 2006).

22 Compare with: (Życiński 1993, pp.75–93).

23 For a historical perspective see e.g., (Obolevitch 2012; Polak and Rodzeń 2021; Polak 2023).

24 This originality is also noticeable in the context of the views of Heller himself and his close associates, such as Życiński. Some subtle differences can be seen here as a result of slightly different research emphases, which Głódź signalled already in the 1970s (see footnote 7 on the "sources of admiration" in the context of physics' successes). In the case of Głódź, one can speak of a slightly greater emphasis on experimental physics than in Heller's case, and therefore a slightly different view of science. Her reflections on the "mathematicity of experience" can be considered an important supplement to Heller's analyses about the mathematicity of the world.

25 See e.g., (Heller 1996). An example of Głódź's position on Thomism: (Głódź 1996b, p.54). More information about the criticism of Thomism from representatives of the OBI milieu: (Trombik 2021).

26 See e.g., (Głódź 1988, p.141).

References

Dębiec, Jacek. 2006. Struktury świata w strukturach mózgu. In *Struktura i emergencja,* edited by Michał Heller and Janusz Mączka, 173–180. Kraków-Tarnów: Biblos.

Głódź, Małgorzata. 1982a. Człowiek nauki w drodze do syntezy (czytając książkę Michała Hellera „Wszechświat i Słowo"). *Znak* 329:296–303.

Głódź, Małgorzata. 1982b. W 100-lecie Teilharda. *Znak* 336:1439–1442.

Głódź, Małgorzata. 1984a. Zagadka substancji wszechrzeczy i człowieka cz. I. *Przegląd Powszechny,* no. 6, 397–408.

Głódź, Małgorzata. 1984b. Zagadka substancji wszechrzeczy i człowieka cz. II. *Przegląd Powszechny,* nos. 7-8, 112–122.

Głódź, Małgorzata. 1987. Międzynarodowe sympozjum Newtonowskie [sprawozdanie]. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 9:40–45.

Głódź, Małgorzata. 1988. Część III: mechanicyzm Oświecenia. In *Wszechświat - maszyna czy myśl?,* 123–155. Kraków: Polskie Towarzystwo Teologiczne.

Głódź, Małgorzata. 1992a. „Search for" eternal information in science – two perspectives. In *The Science and Theology of Information: proceedings of the Third European Conference on Science and Theology, Geneca, March 29 to April 1, 1990,* edited by Christoph Wassermann, Richard Kirby, and Richard Rodorf. Geneve: Labor et Fides.

Głódź, Małgorzata. 1992b. Spotkanie pod szczytem (Wawelu) [sprawozdanie]. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 14:106–110.

Głódź, Małgorzata. 1992c. Uwagi o współzależności eksperymentu i zmatematyzowanej teorii. In *Matematyczność przyrody,* edited by Michał Heller, Józef Życiński, and Alicja Michalik, 105–114. Kraków: OBI.

Głódź, Małgorzata. 1993. On the Edge of Promise and Mususe: contemporary Dialogue Between Science and Religion. *Theoria et Historia Scientiarum* 3:31–36.

Głódź, Małgorzata. 1994. Unfolding the ordered complexity of reality. Is science relevant to theology? In *Origins, Time and Complexity,* 219–223. Geneve: Labor et Fides.

Głódź, Małgorzata. 1996a. Filozofia niedomknięta. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 19:17–23.

Głódź, Małgorzata. 1996b. Matematyczność doświadczenia. In *Przesłanie Księdza Cogito. Księdzu Michałowi Hellerowi w sześćdziesiątą rocznicę urodzin,* edited by Stanisław Wszołek, 45–57. Tarnów: Biblos.

Głódź, Małgorzata. 1996c. Sympozjum Popperowskie [sprawozdanie]. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 18:125–126.

Głódź, Małgorzata. 1997. Imagination challenged by science. Consequences for theological reflection. In *The Concept of Nature in Science and Theology. Vol. III,* edited by Niels Henrik Gregersen and Michael Parsons, 63–67. Geneve: Labor et Fides.

Głódź, Małgorzata. 1999. Zfwn a OBI: dwom panom służyć. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 25:16–19.

Głódź, Małgorzata. 2006. Filozofia w nauce i nauka jako filozofia. In *Wyzwania racjonalności. Księdzu Hellerowi współpracownicy i uczniowie,* edited by Stanisław Wszołek and Robert Janusz, 370–377. Kraków: WAM.

Głódź, Małgorzata. 2010. Gdy czas narasta [recenzja]. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 46:168–172.

Głódź, Małgorzata. 2012. Zagadnienia po raz pięćdziesiąty. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 50:3–9.

Głódź, Małgorzata, and Michał Heller. 1972. Z korespondencji autora „Wobec wszechświata". *Znak* 215:735–741.

Głódź, Małgorzata, and Michał Heller. 1983. Poza granicami fizyki. In *Drogi myślących,* edited by Michał Heller and Józef Życiński, 168–177. Kraków: Polskie Towarzystwo Teologiczne.

Heller, Michał. 1986a. Jak możliwa jest „filozofia w nauce"? *Studia Philosophiae Christianae* 22 (1): 7–19.

Heller, Michał. 1986b. Ontologiczne zaangażowania współczesnej fizyki. *Analecta Cracoviensia* 18:3–20.

Heller, Michał. 1992. *Filozofia nauki. Wprowadzenie.* Kraków: Wydawnictwo Naukowe Papieskiej Akademii Teologicznej.

Heller, Michał. 1996. *The New Physics and a New Theology.* Vatican: Vatican Observatory.

Heller, Michał. 1998. Nauki przyrodnicze a filozofia przyrody. *Roczniki Filozoficzne* 46 (3): 7–14.

Heller, Michał. 1999a. Po 20-stu latach. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 25:3–5.

Heller, Michał. 1999b. Przeciw fundacjonizmowi. In *Sensy i nonsensy w nauce i filozofii,* edited by Michał Heller, Janusz Mączka, and Jacek Urbaniec, 81–101. Kraków-Tarnów: OBI-Biblos.

Heller, Michał. 2006. O strukturalnym rozumieniu nauki i świata. *Prace Komisji Filozofii Nauk Przyrodniczych Polskiej Akademii Umiejętności* I:9–15.

Heller, Michał. 2007. *Filozofia przyrody. Zarys historyczny.* Kraków: Znak.

Heller, Michał. 2019. How is philosophy in science possible? *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 66:231–249.

Heller, Michał, Zbigniew Liana, Janusz Mączka, Adam Olszewski, and Włodzimierz Skoczny. 1999. Jak filozofuje się w OBI? *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 25:20–29.

Heller, Michał, and Józef Życiński. 1990. *Dylematy ewolucji.* Kraków: Polskie Towarzystwo Teologiczne.

Krauze, Filip. 2008. *Jedna prawda, dwie księgi. Nauki przyrodnicze a teologia w Ośrodku Badań Interdyscyplinarnych Papieskiej Akademii Teologicznej w Krakowie.* Kraków: WAM.

Krzanowski, Roman. 2023. From philosophy in science to information in nature: Michael Heller's ideas. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce),* 83–105. https://doi.org/10.59203/zfn.75.680.

Liana, Zbigniew, and Janusz Mączka. 1999. Z kroniki OBI. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 25:133–152.

Mączka, Janusz, Włodzimierz Skoczny, Andrzej Koleżyński, Paweł Polak, Maria Karolczak, Dorota Sieńko, and Małgorzata Stawarz. 2012. Badania nad historią filozofii przyrody i historią nauki. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 50:10–40.

Misner, Charles. 1983. Niematerialne składowe obiektów fizycznych. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 5:5–13.

Obolevitch, Teresa. 2012. Problem relacji między nauką a religią w OBI. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 50:75–84.

Pabjan, Tadeusz. 2019. Filozofia w nauce. In *Metodologia nauk. Cz. II: typy nauk,* edited by Stanisław Janeczek, Monika Walczak, and Anna Starościc, 265–280. Lublin: Wydawnictwo Naukowe KUL.

Pacholczyk, Andrzej. 1996. *Wszechświat katastroficzny.* Kraków-Tarnów: OBI-Biblos.

Polak, Paweł. 2017. Current perspectives on the development of the philosophy of informatics (with a special regard to some Polish philosophers). *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 63:77–100.

Polak, Paweł. 2019. Philosophy in science: a name with long intellectual tradition. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 66:251–270.

Polak, Paweł. 2023. Philosophy in technology: a research program. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 75:59–81. https://doi.org/10.59203/zfn.75.682.

Polak, Paweł, and Jacek Rodzeń. 2021. The science–religion relationship in the academic debate in Poland (1945–1998). *European Journal of Science and Theology* 17 (6): 1–17.

Polak, Paweł, and Trombik. 2022. The Kraków School of Philsophy in Science: profiting from Two Traditions. *Edukacja Filozoficzna* 74:205–229. https://doi.org/10.14394/edufil.2022.0023.

Rodzeń, Jacek. 2005. *Czy sukcesy nauki są cudem? Studium filozoficzno-metodologiczne argumentacji z sukcesu nauki na rzecz realizmu naukowego.* Kraków-Tarnów: OBI-Biblos.

Rodzeń, Jacek. 2006. Co to jest „argument z cudu"? Próba obrony realizmu naukowego oparta na fakcie sukcesu nauki. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 38:3–34.

Rodzeń, Jacek. 2007. Filozofia w kontekście (naukowym) techniki. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 40:76–110.

Schrödinger, Erwin. 1983. Nauka i religia. In *Drogi myślących,* edited by Michał Heller and Józef Życiński, 178–187. Kraków: Polskie Towarzystwo Teologiczne.

Sierotowicz, Tadeusz. 1997. *Nauka a wiara - przestrzeń dialogu. Obrazy świata jako przestrzeń dialogu pomiędzy nauką i teologią.* Kraków-Tarnów: OBI-Biblos.

Sokołowski, Leszek. 2011. O pewnych podobieństwach filozofii fizyki i religii. In *Czy nauka zastąpi religię?,* edited by Bartosz Brożek and Janusz Mączka, 45–62. Kraków: Copernicus Center Press.

Stolarczyk, Ireneusz. 2018. Nadprzyrodzony wymiar godności człowieka. *Tarnowskie Studia Teologiczne* 37 (1-2): 263–279. https://doi.org/10.15633/tst.3262.

Teilhard de Chardin, Pierre. 1982. Moja wiara. *Znak* 336:1422–1438.

Trombik, Kamil. 2019. The origin and development of the Center for Interdisciplinary Studies. A historical outline by 1993. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* 66:271–295.

Trombik, Kamil. 2021. *Koncepcje filozofii przyrody w Papieskiej Akademii Teologicznej w Krakowie w latach 1978–1993. Studium historyczno-filozoficzne.* Kraków: Wydawnictwo Scriptum.

Trombik, Kamil. 2025. The Idea of Mathematicity of the Universe as a Locus Theologicus: michael Heller and Representatives of „The Kraków School of Philosophy in Science". *Religions* 16 (1): 54. https://doi.org/10.3390/rel16010054.

Wszołek, Stanisław. 1994. Znaczenie eksperymentu w nauce, czyli powrót do Bacona. *Filozofia Nauki* 2 (1): 77–84.

Wszołek, Stanisław. 1996. Teoria i doświadczenie w nauce i teologii. In *Przestrzenie Księdza Cogito. Księdzu Michałowi Hellerowi w sześćdziesiątą rocznicę urodzin,* edited by Stanisław Wszołek, 219–234. Tarnów: Biblos.

Życiński, Józef. 1987. Filozoficzne aspekty matematyczności przyrody. In *Filozofować w kontekście nauki,* edited by Michał Heller, Alicja Michalik, and Józef Życiński, 170–185. Kraków: Polskie Towarzystwo Teologiczne.

Życiński, Józef. 1993. *Granice racjonalności. Eseje z filozofii nauki.* Warszawa: PWN.

Review essays

Philosophical Problems in Science (2025), **78**, 59–64
(Zagadnienia Filozoficzne w Nauce)
doi:10.59203/zfn.78.710

REVIEW ESSAY

Ray Kurzweil's manifesto of technological utopia

Anna Sarosiek

University of the National Education Commission, Krakow
Email: anna.sarosiek@icloud.com

Abstract

The Singularity Is Nearer by Ray Kurzweil is a bold manifesto of technological optimism, envisioning a future where humanity transcends its limits. While thought-provoking, the book is criticized for being overly idealistic, neglecting the ethical, social, and existential dilemmas of such profound changes. It emphasizes possibilities without fully addressing the challenges, highlighting that technology alone cannot solve humanity's most pressing issues.

Keywords: singularity, immortality, artificial intelligence, transcendence, Ray Kurzweil

Ray Kurzweil, *The Singularity Is Nearer: When We Merge with AI*, Viking, Penguin Random House, 2024, pp.420.

In his latest book, *The Singularity Is Nearer* (2024), Ray Kurzweil expands upon and updates the ideas presented in his influential work *The Singularity Is Near* (2005). He offers a vision in which the boundaries between humanity and technology are entirely erased. While grounded in the realities of technological progress, his reflections also touch on fundamental questions about the nature of humanity, the essence of life, and the future of consciousness. The book provides ample inspiration, yet it also raises concerns about the philosophical depth of Kurzweil's engagement with the consequences of the future he envisions.

Kurzweil is renowned for making bold predictions, many of which have failed to materialize within the timeframes he initially proposed (Marcus and Davis 2019). While this new book revisits and updates his earlier concepts, it falls short of offering satisfactory explanations for these earlier inaccuracies. Instead, it presents a manifesto of technological optimism, asserting that humanity is on an inexorable path toward the Singularity: a moment when artificial intelligence will surpass human abilities in all domains. This vision aligns with the central tenets of transhumanism, which regards technology as a means of transcending biological limitations and achieving human enhancement. The book occasionally abandons the tone of analysis in favor of proclamation. One chapter opens with the headline "Diligent People Will Achieve Longevity Escape Velocity by Around 2030" (Kurzweil 2024, p.188), a phrase that encapsulates his faith in technological redemption. The rhetoric of inevitability, framed in the language of salvation and destiny, transforms his vision of technological progress into a secular eschatology. Kurzweil paints a vivid picture of a future where the integration of technology with humanity will eliminate disease, aging, and death, while simultaneously opening new possibilities for human consciousness.

Kurzweil effectively highlights the power of human innovation and the limitless potential of imagination, firmly believing that humanity can overcome its limitations. His visions inspire bold thinking about the future and stimulate the reader's imagination, inviting them to consider what can be achieved through the synergy of science and technology. However, despite the impressive scope of topics covered, the book falls short as a serious analysis of the future. The author overlooks numerous fundamental technological, biological, social, and ethical constraints. At times, his narrative drifts toward techno-utopianism, prioritizing enthusiasm over philosophical reflection on the limits of technological progress.

1. Predictions on Artificial Intelligence and the Singularity

Kurzweil opens his book by revisiting his concept of the law of accelerating returns, which posits that technological development progresses exponentially rather than linearly. As he explains, "each advance makes it easier to design the next stage of its own evolution" (Kurzweil 2024, p.3). Drawing on examples from computing and biotechnology, he argues that each technological breakthrough accelerates subsequent advancements, leading toward an inevitable convergence of biological and artificial intelligence (Tegmark 2017; Russell and Norvig 2020).

Regarding artificial intelligence, Kurzweil predicts that by 2029 AI will reach the level of human general intelligence (AGI), and by 2045 the Singularity will become a reality (Kurzweil 2024, p.2). It is worth noting, however, that his understanding of AGI differs from the traditional Searle's notion of strong AI. For Kurzweil, AGI denotes a system capable of matching or surpassing human cognitive capacities across all intellectual domains. It is a futurist projection rather than an empirically grounded definition.

Kurzweil emphasizes the transformative potential of AI, presenting it as a catalyst for solving humanity's most pressing challenges. Yet this vision downplays the limitations of current AI systems. Models such as Gemini and GPT, though capable of producing coherent linguistic output, lack commonsense reasoning and self-understanding. Their operation depends on vast datasets that approximate but do not embody the complexity of human experience (Arkoudas 2023; Liu et al. 2024), and their meta-ontology differs from human meta-ontology (Krzanowski and Polak 2022). Moreover, the "black-box" nature of machine-learning models renders their decision-making processes opaque, which becomes particularly problematic in high-stakes contexts such as medicine or law (Marcus and Davis 2019).

Kurzweil asserts that AGI is inevitable (Kurzweil 2024, p.286), but he overlooks the technical and conceptual difficulties of simulating intuition, emotion, or intentionality. His optimism regarding the integration of the human brain with cloud computing presupposes breakthroughs in neuroscience that remain purely speculative. While his appeal to historical analogies, Moore's Law above all, adds rhetorical force, it neglects physical and energetic constraints that could slow or even arrest exponential growth (Tegmark 2017).

Kurzweil's confidence in the timeline of 2029 and 2045 reflects less an empirical projection than a metaphysical conviction that acceleration itself has ontological necessity. It is here that the law of accelerating returns becomes less a descriptive principle and more a doctrine of faith, a secular teleology of progress in which technology inherits the role once reserved for evolution. From a biosemiotic standpoint, for example, evolution is not a matter of increasing speed or capacity but of ongoing interpretation of signs exchanged between organisms and their environments (Kull 2015). Technological progress, by contrast, concerns only the velocity and volume of processing; it amplifies computation without necessarily generating meaning. The creation of meaning, so central to intelligence and to the very idea of the Singularity, does not follow automatically from faster or more complex processing. Biosemiotics reminds us that the presence of symbols does not guarantee the presence of meaning, for technological systems manipulate signs but do not interpret them. Without interpretation, there is no subject, and without subjectivity, no intelligence in the human sense.

2. The Vision of Human Immortality

One of the central themes of the book is the vision of human immortality (Kurzweil 2024, pp.93). Kurzweil foresees that with the help of nanobots and biotechnology, humanity will be able to eliminate diseases at the molecular level, regenerate tissues, and counteract aging. He views death as a technical problem that will soon be solved (Kurzweil 2024, pp.191–192). According to his projections, it will be possible not only to eradicate death but to maintain the human body in an ideal and self-repairing state indefinitely. He cites tools such as CRISPR gene editing and nanoscale medical devices, which, in his view, " [...] one of the most profoundly important lifesaving objectives for AI [...]" (Kurzweil 2024, pp.241–242).

However, Kurzweil's approach appears oversimplified. Aging is not merely the result of molecular damage but also of complex interactions between genes, the environment, and the regulatory mechanisms of the body (López-Otín et al. 2013). Reversing these processes technologically may prove far more challenging than Kurzweil assumes. Although CRISPR is a revolutionary method, it still faces fundamental limitations, including the difficulty of precisely targeting edits without unintended consequences and the uncertainty surrounding long-term genetic stability (Doudna and Sternberg 2017). Similarly, nanotechnology within the human body demands a level of understanding of systemic effects that remains beyond current research .

Beneath the scientific optimism lies a deeper philosophical question: what becomes of human existence when mortality is abolished? Classical philosophical thought suggests that life derives its significance precisely from its finitude: awareness of death grants value to experience (Sandel 2007). Kurzweil does not explore whether the abolition of death might render life meaningless or merely endless. Awareness of limitation, not the absence of it, gives form to experience. Without transience, there may be no depth of presence.

What Kurzweil overlooks is that fragility of the body itself is not an obstacle to progress but a condition of meaning. In living systems, vulnerability and limitation are semiotic functions: they mark the boundaries within which interpretation becomes possible. The body, through its perishability, situates the organism in a world of signs: it feels, responds, and assigns value to what sustains or threatens it. To remove mortality from this equation is to remove the very context in which meaning arises. An immortal being might process information indefinitely, but without the tension between preservation and loss, it could no longer interpret the world as something that matters. In this perspective, finitude is not a flaw of life but the ground of its sense-making

3. The Future of Human Consciousness

Kurzweil addresses a fundamental question about the future of human consciousness in a world increasingly mediated by technology. He envisions a time when the human mind will merge with machines, gaining unlimited memory, computational precision, and expanded perception (Kurzweil 2024, p.8). Such integration, he argues, will transform the very nature of experience, allowing consciousness to exist independently of the body. Yet this proposition invites a crucial philosophical question: can consciousness be meaningfully preserved once detached from its embodied and experiential roots?

Kurzweil suggests that consciousness can be transferred as effortlessly as a digital file: "it doesn't matter whether your JPEG files are stored on a floppy disk, a CD-ROM, or a USB flash drive, they look the same and work the same as long as the information is represented with the same sequence of 1s and 0s" (Kurzweil 2024, p.104). The analogy is telling. In reducing consciousness to a data format, Kurzweil treats subjective experience as if it were a picture file, identical wherever copied, indifferent to the medium that holds it. Adapting Wittgenstein's phrase, "a picture held him captive" (Wittgenstein 1958).

To give his view a metaphysical sheen, Kurzweil appeals to what he calls panprotopsychism, borrowing the term from Chalmers (1996). Yet where Chalmers treats panprotopsychism as a speculative attempt to address the hard problem of consciousness, an irreducible quality of experience that eludes physical explanation, Kurzweil instrumentalizes it. He recasts it as a convenient bridge between biology and computation: if consciousness is an emergent field that "awakens" whenever information is sufficiently complex, then silicon can feel as carbon does (Kurzweil 2024, pp.80–81). The gesture toward philosophy serves not to deepen reflection but to naturalize a technological faith. A mind without a body becomes as viewable and as hollow as its digital image.

Philosophers such as Nagel and Chalmers remind us that consciousness cannot be reduced to computational patterns. For Nagel, subjectivity arises from the organism's perspectival position in the world, from what it is like to be embodied (Nagel 1974). For Chalmers, phenomenal experience resists functional explanation because it is not the organization of data that produces awareness but the act of being within that organization (Chalmers 1996). Both perspectives highlight that consciousness is not a system of relations among symbols but a mode of interpretative presence: a process of meaning-making sustained by the living body.

From this standpoint, transferring consciousness into a digital substrate would not preserve experience but abstract it. An uploaded mind might replicate cognitive patterns, yet it would lack the sensory grounding and semiotic reciprocity that constitute awareness. Sensation without embodiment is not perception; continuity of information is not continuity of being. The body, in its fragility, situates the subject in a world of signs and gives coherence to experience through its limitations. Once detached from that context, consciousness ceases to interpret, it only computes.

Kurzweil's vision of collective, networked consciousness follows naturally from this premise (Kurzweil 2024, p.112), yet it raises profound questions about identity and intersubjectivity. If minds merge within a digital continuum, will individuality persist, or will the self-dissolve into informational transparency? Human relations depend on embodied opacity: on gestures, distance, and temporal unfolding. A consciousness without boundaries might know everything, but it would no longer encounter anything.

4. Ethical and Societal Implications of Technological Progress

One of the more philosophical elements of The Singularity Is Nearer is a dialogue with "Cassandra," a symbolic skeptic who questions Kurzweil's optimistic vision of the future (Kurzweil 2024, pp.286–290). In this exchange, she voices concern about the pace and direction of technological advancement: "So you anticipate a neural net with sufficient processing power to exceed all human capabilities by 2029?" to which Kurzweil replies simply, "Correct." The rhythm of this dialogue, Cassandra's questions met with unflinching affirmations, reveals more than its content. It stages optimism as certainty, turning dialogue into doctrine. The figure of Cassandra, mythically condemned to be disbelieved, functions here as a rhetorical prop: skepticism is voiced only to be neutralized.

Cassandra raises fundamental questions about who will control technological development, how it will be distributed, and what consequences it will have for human life. Yet Kurzweil's replies reduce these concerns to technical solvability. His optimism functions as a form of moral insulation, a belief that progress is self-justifying and that its risks will be absorbed by innovation itself.

The ethical dimension of Kurzweil's vision remains underdeveloped. He acknowledges moral challenges but treats them as secondary to innovation, as problems to be engineered away rather than as questions redefining the purpose of technological advancement. A more explicit examination of the political and social tensions inherent in his predictions is needed. Above all, it remains unclear how these technologies might exacerbate existing inequalities: access to advanced biotechnology or cloud-based cognitive augmentation will likely remain confined to the privileged few, deepening the gap between social classes. The promise of a post-scarcity future risks concealing a new form of digital stratification, where intelligence itself becomes a commodity.

These developments also threaten privacy and autonomy on an unprecedented scale. The merging of mind and machine, which Kurzweil envisions as an expansion of freedom, could equally become a mechanism of control. Access to neural data and thought patterns could allow states or corporations to shape, monitor, or even rewrite subjective experience. Technologies meant to extend cognition could render the individual transparent to systems of power.

Moreover, the political implications of such integration remain largely unexamined. Technologies capable of monitoring or influencing cognition could become tools of governance, especially in authoritarian contexts. What Kurzweil frames as collective intelligence might easily devolve into collective surveillance. His reflections, though compelling, remain silent on how to safeguard human dignity (see e.g., Machidon 2024) and pluralism in a world where technological systems may anticipate, and pre-empt, human decision.

From a philosophical standpoint, Kurzweil's faith in progress reveals a deeper assumption: that technology, by virtue of its complexity, carries its own moral trajectory. Yet morality is not an emergent property of systems; it is a mode of interpretation grounded in lived meaning. Once detached from interpretative responsibility, technological evolution risks becoming a semiotic process without ethics, an autonomous production of signs with no subject to answer for them.

5. Conclusion

The Singularity Is Nearer reads less like a study of the future than a manifesto of technological optimism. Kurzweil's vision is ambitious and provocative, yet often marked by an uncritical faith in progress. His arguments rest on the assumption that technological evolution will naturally resolve the problems it creates, and that intelligence, once expanded through computation, will retain meaning and moral orientation.

The book invites readers to imagine a world in which humanity transcends its biological limits, but it leaves unresolved the question of what such transcendence would entail for consciousness, value, and social order. By focusing on possibility rather than complexity, Kurzweil offers an inspiring but incomplete vision, one that omits the interpretative, relational, and embodied dimensions that make human understanding what it is. The challenge his work leaves us with is not to abandon the dream of progress, but to think it through, to ask whether acceleration alone can constitute direction, and whether intelligence without interpretation and understanding can still be called intelligence.

References

Arkoudas, Konstantine. 2023. *GPT-4 Can't Reason.* ArXiv:2308.03762 [cs]. https://doi.org/10.48550/arXiv.2308.03762.

Chalmers, David J. 1996. *The Conscious Mind: In Search of a Fundamental Theory.* Oxford: Oxford University Press.

Doudna, Jennifer A., and Samuel H. Sternberg. 2017. *A Crack in Creation: Gene Editing and the Unthinkable Power to Control Evolution.* Houghton Mifflin Harcourt.

Krzanowski, Roman, and Paweł Polak. 2022. The meta-ontology of AI systems with human-level intelligence. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce),* no. 73, 197–230. https://doi.org/10.59203/zfn.73.610.

Kull, Kalevi. 2015. Semiosis and meaning-making in living systems. *Biosemiotics* 8 (1): 1–12. https://doi.org/10.1007/s12304-014-9223-4.

Kurzweil, R. 2024. *The Singularity Is Nearer: When We Merge with AI.* Viking / Penguin Random House.

Kurzweil, Ray. 2005. *The Singularity Is Near: When Humans Transcend Biology.* Viking.

Liu, Ziyi, Soumya Sanyal, Isabelle Lee, Yongkang Du, Rahul Gupta, Yang Liu, and Jieyu Zhao. 2024. *Self-Contradictory Reasoning Evaluation and Detection.* ArXiv:2311.09603 [cs]. https://doi.org/10.48550/arXiv.2311.09603.

López-Otín, Carlos, Maria A. Blasco, Linda Partridge, Manuel Serrano, and Guido Kroemer. 2013. The Hallmarks of Aging. *Cell* 153 (6): 1194–1217. https://doi.org/10.1016/j.cell.2013.05.039.

Machidon, Octavian-Mihai. 2024. Upholding human dignity in AI: advocating moral reasoning over consensus ethics for value alignment. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce),* no. 77, 25–39. https://doi.org/10.59203/zfn.77.707.

Marcus, Gary, and Ernest Davis. 2019. *Rebooting AI: Building Artificial Intelligence We Can Trust.* Pantheon.

Nagel, Thomas. 1974. What is it like to be a bat? *The Philosophical Review* 83 (4): 435–450. https://doi.org/10.2307/2183914.

Russell, Stuart J., and Peter Norvig. 2020. *Artificial Intelligence: A Modern Approach.* 4th. Pearson.

Sandel, Michael J. 2007. *The Case Against Perfection: Ethics in the Age of Genetic Engineering.* Cambridge, MA: Harvard University Press.

Tegmark, Max. 2017. *Life 3.0: Being Human in the Age of Artificial Intelligence.* New York: Knopf.

Wittgenstein, Ludwig. 1958. *Philosophical Investigations.* 3rd ed. Translated by G.E.M. Anscombe. Oxford: Basil Blackwell.