dialectica

International Journal of Philosophy

PROOF

Contents

European Society for Analytic Philosophy

Vol 77 : 2
June 2023

dialectica

International Journal of Philosophy
Official Organ of the European Society of Analytic Philosophy

founded in 1947 by Gaston Bachelard, Paul Bernays and Ferdinand Gonseth dialectica

vol 77
issue 2

June 2023

Contents

Artifact Concept Pluralism

Alper Güngör

We have a rough idea of what artifacts are: artifacts are objects made to serve a certain purpose. However, there is no consensus on how to specify this definition. Essentialists argue that objects are grouped into artifact kinds by sharing non-trivial artifact essences, while anti-essentialists argue that there is no such essence to be found. However, the prominent essentialist and anti-essentialist accounts suffer from extensional and definitional problems. I argue that the problems current essentialist and anti-essentialist accounts face mainly stem from the assumption of *artifact concept monism*. According to artifact concept monism, there is only one way to group objects into artifact kinds. To remedy the problems that stem from artifact concept monism, this paper offers an alternative framework by drawing parallels from the debates on species concept pluralism and art concept pluralism.

The rapidly growing literature on artifacts revolved mostly around finding non-trivial artifact essences, while dissenting voices pointed out the plurality of artifact kinds and raised legitimate concerns about the applicability of any essence for artifacts and artifact kinds. I call the first endeavor *artifact essentialism* and the latter *artifact anti-essentialism*. Both essentialists and anti-essentialists, implicitly or explicitly, share the same assumption: that there is only one legitimate artifact concept that we can profitably use. I call this view *artifact concept monism*. I argue that the current state of artifact essentialism cannot provide an extensionally adequate and definitionally coherent over-arching concept. The extensional and definitional problems I point out led some anti-essentialists to give up on classificatory aims and others to doubt the primacy of metaphysics on the topic of artifacts. In this paper, I aim to offer an alternative to artifact concept monism. I call my view *artifact concept pluralism*. I argue that artifact concept pluralism provides a better framework to deal with the problems artifact essentialism faces. Furthermore, it enables us to bring metaphysical and epistemic considerations together without giving up on the classificatory aims and requiring a significant revision in our taxonomical practices.

That said, this paper's main methodological leaning is clear: practices come first. According to David Davies (2004), an ontologist of art should not put forward metaphysical principles before examining the practices closely; art practices impose a "pragmatic constraint" on metaphysical accounts. As Davies (2004, 18) describes this pragmatic constraint, "Artworks must be entities that can bear the sorts of properties rightly ascribed to what are termed 'works' in our reflective and critical and appreciative practice…" Similarly, in this paper, I assume that artifact practices impose a "pragmatic constraint" on the metaphysics of artifacts. This does not mean that artifact practices are the final arbiter of our best metaphysical account; rather, our rational reconstruction of the output of the relevant practices determines our metaphysical accounts. However, as artifact practices are (even) less uniform than art practices and given the problems current monistic accounts face, I argue that a responsible form of pluralism is needed to account for artifact practices.

Following Kathrin Koslicki (2008, 201), I take kinds as "taxonomic classifications under which particular objects may be grouped based on shared characteristics of some sort." Accordingly, an artifact concept is what singles out the relevant characteristics required for artifact-kind membership. Artifact concept monism assumes that there can only be one way of grouping entities under artifact kinds, and thus, it assumes that there is an overarching artifact concept. Artifact concept pluralism rejects this assumption. I construct a model of artifact concept pluralism following Christy Mag Uidhir and P.D. Magnus's proposal on the art concept pluralism. According to Mag Uidhir and Magnus (2011, 91–92), there are at least four art concepts; in other words, there are four ways of grouping art objects, and each way of grouping has its own strengths and weaknesses. Mag Uidhir and Magnus (2011) draw their art concept pluralism on the model of species pluralism. According to species pluralism, there are several ways of grouping organisms into species. Both models guide this project of artifact concept pluralism. Drawing on these models and taking the output of relevant practices seriously, artifact concept pluralism proposes that there are multiple correct ways of grouping entities into artifact kinds.

1 Artifact Essentialism

John Locke famously distinguished the real essence of things from their nominal essences (Locke 1689, bks. 3, chap.3, para.15; cf. Reydon 2014, 127). The former is generally construed as the mind-independent nature of things, whereas the latter depends on how the relevant minds conceive of entities (Reydon 2014, 127). Although Locke was pessimistic about finding real essences of things, in the case of natural kinds, those authors who prefer the semantics put forward by Kripke (1980) and Putnam (1975b) seek out kinds whose nature is constituted by mind-independent essences (Thomasson 2007, 54). For instance, in the case of a natural-kind term like gold, all gold atoms share the same atomic structure, and this structure is discoverable by the relevant scientific practices. This mind-independent essence of gold, in turn, fixes our reference to the term 'gold' and enables us to distinguish genuine gold from fool's gold (Reydon 2014, 127).

Some suggest that a similar strategy applies to artifact-kind terms and claim that functions can serve a reference-fixing role for artifact-kind terms (Putnam 1975b; Kornblith 1980). Others argue further that some artifact kinds have a mind-independent nature akin to natural kinds (Elder 2007; Franssen and Kroes 2014). However, it is not at all clear that the traditional distinction between mind-dependent and mind-independent essences and its bearing on reality is uncontroversial (Reydon 2014, 130). Not all natural kinds neatly follow this distinction. For instance, it is now commonly taken for granted that biology failed to provide genetic essences unique to species simply because species are found to be subjected to constant evolutionary change (Reydon 2014, 131). A new form of essentialism is on the rise in the philosophy of science (Boyd 1999; cf. Reydon 2014).

According to the new essentialism, the essences need not be non-relational properties. The paradigmatic cases are biological kinds. Historical and relational properties are now considered part of biological kinds' essences (Reydon 2014, 130–131). The new form of essentialism is also suitable to accommodate artifact kinds. After all, possible candidates for artifact essences refer to how artifacts are being used, why they are reproduced, etc. Having briefly elucidated both forms of essentialism, I formulate essentialism about artifact kinds broadly as follows:

doi: 10.48106/dial.v77.i2

> Artifact Essentialism. Necessarily, for all x, if x is an artifact, then there's some essence E such that x has E, and x is a member of artifact kind K in virtue of E.

I will consider Artifact Essentialism as a condition about kind essences as opposed to individual essences. There are at least two distinct construals of individual essences. First, it might mean a particular instance i of the kind K essentially belongs to K. On this understanding, i cannot exist without being a K (Bird and Tobin 2022). According to the other construal, individual entities might have essential properties besides the essential properties shared with the other instances of the kinds they belong to. For instance, if we agree with Kripke (1980) on origin essentialism, then being a child of my parents is an essential property of mine, while it is not an essential property of the kind *human*. Having made the distinction between individual and kind essences, we can state that throughout this paper, a kind essence E indicates a non-trivial essential property or a set of properties that are shared by the members of an artifact kind. For instance, if artifact essentialism is best understood in terms of functions, one would expect individual chairs to have the function of *seating a single individual*, and by this functional property, one could assess whether a given chair is a proper chair, a malfunctioning chair, or a non-chair (e.g., a chair beyond repair).

To make my discussion more exhaustive, I take artifact essentialism to be neutral on traditional and new forms of essentialism. The most commonly discussed artifact kind essences (E) are the following (Grandy 2007; Vega-Encabo and Lawler 2014; Koslicki 2018): i) functions, ii) maker's intentions. I will not provide a detailed explanation of any individual account. Having provided the general essentialist outline, I raise two problems against artifact essentialism, namely, the extensional problem and the definitional complexity problem. Both problems are raised by Mag Uidhir and Magnus (2011) in their attack against art concept monism. I follow a similar argument.

1 *Function Essentialism*

A quick survey both of the literature and pre-theoretical intuitions shows that functions are the most favored artifact essences.[1] Even many familiar

1 Juvshik (2021b) formulates "function essentialism" and attempts to refute it. In this section, I largely benefit from his discussion.

artifacts around us are named after their functions (Baker 2008). To list a few: screwdriver, corkscrew, pencil sharpener. Kornblith (1980, 112) writes, "At least, for the most part, it seems that what makes two artifacts members of the same kind is that they perform the same function." Kornblith's statement provides us with the basic intuition behind function essentialism.

According to Tim Juvshik, function essentialism favors function as the best candidate for artifact essences. To elaborate by way of an example, a triangle screwdriver and a magnetic screwdriver have distinct designs and perform their functions differently. The former's design is more safety-oriented, whereas the latter, with the help of magnetic force, performs a better job with smaller screws. Yet they both drive screws. Given the significant multiplicity of form and design, according to function essentialists, functions provide a prima facie suitable artifact essence that can bind various artifacts under a single artifact kind (Preston 2013).

However intuitive the functional characterization of artifacts and artifact kinds is, there is no consensus on how to characterize functions. The first attempt to characterize functions may be taking functions as answers to "what is it there for" questions, which in turn explains "how the thing got there" (Wright 1973, 146–156; Vega-Encabo and Lawler 2014; Juvshik 2021b). For instance, I can use a towel as a cover for my favorite snacks, yet a towel is *for* drying hands, just as the heart is there *for* pumping blood, not *for* producing a unique sound. Wright (1973) calls the former function of my towel function *as* and the latter *the* function.[2] The main difference between these two senses of functions is that the latter has the explanatory force that accounts for the historically successful reproduction of, say, towels, which the former lacks.

Wright's distinction is more or less retained in the subsequent theories of function. Benefiting from the literature on functions, philosophers recently put forward elaborate theories on artifacts. The attempts can be largely divided into two camps: *etiological functionalism* and *intended functionalism*. Emphasizing the etiological aspect of functions while eschewing the intentional properties, Elder, one of the champions of etiological functionalism, suggests that many artifact kinds share a similar nature with natural kinds; these kinds essentially instantiate a cluster of properties that are copied among the members (Elder 2007, 37). The cluster of properties for artifact kinds includes three main elements: particular shape, proper function, and historical placement

2 The same distinction is used by many under different headings. Vermaas and Houkes (2003, 262–266; cf. Juvshik 2021b) use standard/accident functions, and Evnine (2016) calls it kind-associated/idiosyncratic functions.

doi: 10.48106/dial.v77.i2

(Elder 2007, 43). The kinds of objects that satisfy all these elements, in Elder's view, are *copied kinds*. Copied kinds include both natural and artifact kinds without having any ontologically significant difference between them.[3]

However, etiological functionalism leaves us with conclusions that are at odds with our ordinary linguistic practices (Thomasson 2007; Juvshik 2021b). In Elder's view, for instance, a familiar artifact kind such as *corkscrew* turns out not to be a copied kind since its nature is not specific enough because the shape shows high variations among corkscrews. Thus, this view admits only specifiable artifact kinds like *winged corkscrew*, which has a certain shape (e.g., winged), proper function (e.g., *to remove corks*), and historically proper placement (e.g., H.S. Heely's 1888 patent) (Thomasson 2007). This result is controversial for those who try to account for intuitive artifact kinds such as *corkscrew* and *chair* (Thomasson 2007; Juvshik 2021b).

Many philosophers, on the other hand, emphasize the intentional aspect of functions rather than the etiological aspect. Artifacts, after all, for intended functionalists, are, in a significant sense, dependent on the activities of conscious agents. Given the importance of intentions of the relevant agents, intended functionalists claim that artifacts have functions that make necessary reference to our "needs, desires, and plans" (Thomasson 2009, 205). Thus, according to intended functionalists, artifacts have *the* functions because their makers bestow them those very functions.

However, this quickly leads to the following problematic cases: Some corkscrews are only produced or used for aesthetic purposes and are not intended to remove any cork. Similarly, some ships and chairs are produced as exhibition ships and chairs (Bloom 1996, 5). We can add motors, cars, guitars, and many other artifacts to the list. Bloom presents these cases as a threat to intended functionalism. Because in such cases, either one should admit that artifact kinds are not united by a shared intended function or that those particular entities are not members of the relevant artifact kinds.

One can defend intended functionalism by underlying the feature of reproductive success that is associated with *the* functions. Chairs, after all, are reproduced throughout history because they were highly useful in seating people, not because they are good decorative pieces in exhibitions. This is, for

3 Elder (2007) favors the traditional form of realism, according to which an entity is real only if it has a mind-independent nature. That is why he emphasizes the three mind-independent features that are mentioned here. His account, in fact, shares many interesting elements with the anti-essentialist HPC view I discuss in section 2.1. It is important to point out that one can also formulate an essentialist HPC view based on, for instance, Elder's remarks.

instance, the route Evnine takes in his distinction between kind-associated functions and idiosyncratic functions (2016, 119–124). For Evnine, the kind-associated function of *chair* is to be sat upon, while if someone produces a chair for exhibition purposes, then *that* chair has an idiosyncratic function (*being an exhibition piece*) in addition to its kind-associated function (*seating a single individual*). Thus, for Evnine, artifact functions are still present even when they are not performed or not intended to be performed (2016, 121–124).

Although Evnine's distinction seems to secure kind-associated functions for Bloom's cases, it still suffers from a more serious case: artworks. Artworks are considered the epitome of artifacts. However, if artifacts are grouped under an artifact kind by their kind-associated functions, then many highly esteemed artworks (especially the modern works after Marcel Duchamp's *The Fountain*) of the twentieth and twenty-first centuries turn out not to be artifacts simply because they lack functions (Koslicki 2018, 218; Juvshik 2021b). Furthermore, even if specific paintings have functional properties such as invoking religious feelings (e.g., religious paintings), *painting* kind does not seem to have unifying functional properties (Juvshik 2021b). Thus, functional theories can only account for specific art kinds that are produced to fulfill certain functions.

To sum up, etiological function essentialism faces the extension problem because the view is extensionally inadequate—it can only provide an arbitrary fineness of grain at best and, thus, leaves out many familiar artifact kinds. In contrast, intended functionalism is better at dealing with intuitive artifact cases; nonetheless, the view suffers from the extension problem as it cannot easily explain Bloom's cases (e.g., exhibition ships). Even if there is a possibility to parry Bloom's cases, many non-functional artworks still constitute a deep extensional worry.

Given the heterogeneity of the artifactual world, some proponents of intended function restricted their domain of inquiry only to cover "technical artifacts" (Baker 2007, 49). This, however, leads to a further problem, namely, the definitional complexity problem (Mag Uidhir and Magnus 2011, 85). Mag Uidhir and Magnus (2011, 85) write, "In order to capture art's plurality and thereby avoid extensional worries, definitions often become dangerously complex, borderline arbitrary, or circular." Similarly, in the case of artifacts, delineating a distinction between technical artifacts and non-technical artifacts is not principled (Koslicki 2018, 235; Juvshik 2021b, 19). Because appealing to the "technical artifact" restriction cannot be profitably defined to exclude "technical" artworks (Juvshik 2021b). For instance, the cases of computer art discussed in Lopes (2009) show that there are technically complex artifacts that have no obvious function (Juvshik 2021b). Therefore, given the definitional complexities and extensional problems, it seems that both etiological and intentional theories of functions fail to serve as an overarching artifact concept. Acknowledging this problem, Evnine (2016, 129) also admits a kind of pluralism by considering artworks as *sui generis* artifact kinds.

1.2 *Intention Essentialism*

The basic motivation behind intention essentialism is rooted in Hilpinen (1992) and Bloom (1996). Bloom (1996, 10) writes, "Although someone can create a chair without intending anybody to sit on it, it is difficult to see how someone can create a chair without intending it to be a chair." The upshot of Bloom's insights is that function and shape do not provide a stable ground for artifact groupings, but the maker's intention does.

Based on Bloom's insights, Amie Thomasson further defends the essentiality of intentions (Thomasson 2003, 2007, 2009, 2014). According to her, what lies at the core of artifacts is the maker's intentions:

> Necessarily, for all x and all artifactual kinds K, x is a K only if x is the product of a largely successful intention that (Kx), where one intends (Kx) only if one has a substantive concept of the nature of Ks that largely matches that of some group of prior makers of Ks (if there are any) and intends to realize that concept by imposing K-relevant features on the object. (Thomasson 2003, 600)

Unlike functionalist essentialist accounts, Thomasson's intentionalist account does not imply any strict necessary *and* sufficient condition. Even if intention essentialism does not impose strict necessary *and* sufficient conditions, nonetheless, as the above quote shows, Thomasson claims that the maker's intentions are necessary for *all* artifacts. Assuming that Thomasson's intentionalist account constitutes some form of essentialism, it faces several problems. As I focus on the cases that seem to be artifact cases but fail to be one given the definitional restrictions of essentialist accounts, I will leave the discussion of other problems aside.[4] Intention essentialism leaves out what I will call *twilight kinds*.[5] Twilight kinds include kinds such as *path*, *village*, *trail*, *foot-*

4 See Koslicki (2018, 226–237) for an extensive list.

5 Twilight kinds are discussed in Margolis and Laurence (2007) and Koslicki (2018, 219–220). I derive the name "twilight kind" from Koslicki's discussion. Koslicki (2018, 235) claims that if the

print, *doodle*, etc. Members of these kinds are not exhaustively products of intentions. For instance, a path can unintentionally come into existence as a result of many agents' repeated movements from one place to another via the same way (Koslicki 2018, 219). Similarly, people might decide to build shelters in a close range without any intention to create a member of the village kind, yet might end up unintentionally creating a village. Although some members of twilight kinds come into existence unintentionally, still, as a kind *path* or *village*, we seem to agree on their status as artifact kinds. If some members of these artifact kinds are not intentionally created, then this means those artifact kinds do not share the necessary condition of "intending to create a kind K" Thomasson (2003) puts forward.[6]

Acknowledging the twilight kinds, Thomasson (2007, 58, n.5) slightly restricts her account by limiting her account to cover only "the essentially artifact kinds," members of which are exhaustively produced with the right sorts of intentions. To my knowledge, this exclusion is not defended thoroughly, except in Juvshik (2021a), to some extent.[7] According to Juvshik (2021a), there are two lines of argument against the intention-dependent nature of artifact kinds:

> (1) Artifacts aren't necessarily mind-dependent, but most of the artifacts around us happen to be.
> (2) Artifacts are necessarily mind-dependent, but don't need to be intention-dependent. (Juvshik 2021a, 9316)

To defend intention essentialism, Juvshik (2021a) considers five cases: Regarding (1), swamp and modal cases. Regarding (2), accidental creation, mass production, and automated production. Not all of these cases are relevant to my purposes. Leaving out mass production and automated production, I will discuss swamp and modal cases later. For now, I will focus on accidental creation. My ultimate critique of intention essentialism will take the form of (2).

law of excluded middle does hold, then these cases cast confusion since they seem to be neither natural kinds nor artifact kinds.

6 It might be useful to note that the twilight kinds also raise an extensional worry to the functionalist essentialist accounts that take intentions as necessary.

7 Hilpinen (1992, 66), in a short paragraph, suggests twilight cases should be taken as "natural cultural objects," echoing what some archeologists and anthropologists call "naturefact." These are objects crafted by natural forces put into human use, such as rocks used as hammers. Also, like Thomasson, Evnine (2016, 19–20) and Grandy (2007, 24) rule twilight kinds out of their discussion.

doi: 10.48106/dial.v77.i2

The closest case discussed by Juvshik to the twilight cases is the case of accidental creation. Accidental creation is distinct from proper creation because, in the former, the intention to create *that* item is lacking. His discussion of accidental creation mostly revolves around the cases of failed-attempts-turned-into-new-artifacts. For instance, the piece of bread I forgot in the toaster turns out to be pretty good charcoal for my new drawing. So, I accidentally create a new piece of drawing charcoal. However, Juvshik aims to show that there is neither a toast nor a piece of charcoal, unless they are *appropriated* in the right sort of way. The moment of my appropriation of the failed toast as a piece of drawing charcoal marks the moment of the new artifact's coming into existence. Appropriation also requires me to have, at least, a basic awareness of the relevant success conditions of making a piece of drawing charcoal.

However, twilight cases do not result from failed attempts. Instead, their coming into existence does not involve *attempting* to create an artifact. Yet, Juvshik might respond that, even if some members of twilight kinds are not failed-attempts-turned-into-new-artifacts, they are still non-artifacts, unless they are correctly appropriated. If that is the case, then the path formed as a result of my repeated commuting from the barn to the house is not actually a member of the *path* kind. Unlike Thomasson, Juvshik rules out not the kind itself but the unintentional cases. However, this will end up admitting that a large number of twilight cases, even though they share a similar morphological structure with their intentionally created counterparts, are ultimately waiting for an appropriator to confer them a status of artifactuality. I do not think that an archeologist or an anthropologist would accept the result that the unintended path was not created, say, one thousand years ago, but at the moment they approve it as a path. Archeologists and anthropologists discuss the significance of the path for that culture regardless of it being a product of specific intentions. Thus, contrary to Juvshik, I think that the twilight cases amount to genuine artifact cases without requiring a strict intention dependence. Twilight cases can be considered mind-dependent without being intention-dependent since their coming into existence requires the presence of agents with cognitive capabilities.

An intention essentialist might also respond by weakening their account only to require mind-dependence. However, the weakened account would not be helpful in distinguishing many other mind-dependent entities from artifacts. For instance, since the existence of many kinds of plants (e.g., seedless grapes) requires human activity, these plants and animals would be wrongly included in the domain of the overarching artifact concept, for which the only necessary condition is being mind-dependent. This strategy, therefore, would not be desirable for an intention essentialist who works in a monistic framework.

Even if one agrees that the twilight cases pose a legitimate worry against intention essentialism, a proponent of intention essentialism can still point out that those cases are a burden for everyone and thereby suggest that those cases are best left out until our most promising theory can account for them (Juvshik 2021a). However, we should not opt for the inference to the best explanation without examining other alternatives in depth. There *is* a neglected alternative. I will outline artifact concept pluralism as an alternative to the artifact concept monism after I challenge artifact anti-essentialism in the next section.

2 Artifact Anti-essentialism

The preceding discussion indicates that there seems to be a plethora of essentialist accounts. In contrast, unfortunately, there is not any fully developed anti-essentialist account. This is the reason why Koslicki (2018, 237–240) discusses general anti-essentialist frameworks that might apply to the case of artifacts. Here, I will focus on the artifact literature in order to extract some anti-essentialist views.[8]

8 As an anonymous reviewer rightly points out, this discussion of anti-essentialist views is not exhaustive. For instance, David Wiggins (2001) rejects artifact kinds as real for lack of determinate identity and persistence conditions. See Soavi (2009) for a more elaborate discussion of Wiggins' views. Leaving out the discussion of anti-essentialist anti-realist views, here I limit my discussion to realist views. However, here is a foreshadowing of how pluralism might be considered a realist position: Those who hold neo-Aristotelian views argue that artifact kinds are primary. According to these views, without knowing which artifact belongs to which primary kind, it is hard to distinguish the allegedly substantial kinds, such as *coin*, from the phasal kinds, such as *coin-in-a-pocket* (Baker 2004, 100). Baker (2004, 100) argues that there is a crucial ontological difference between objects essentially belonging to primary kinds (e.g., *coin*) and merely conventional groupings (e.g., *coin-in-a-pocket*). The former kinds are real, but our ontology cannot accommodate adding the latter. Because adding the latter would result in the proliferation of all sorts of imaginary entities. Pluralism by adopting context relativity seems to disrupt this hierarchy. Given that pluralism is not compatible with hierarchical classification, does this commit pluralism to some form of anti-realism about artifacts or artifact kinds? It certainly commits pluralism to a form of anti-essentialism, at least in the sense that there is not a unifying essential structure that applies to artifact kinds. I think for those who assume artifact concept monism, the result is worrying. The reason is that artifact concept pluralism leads to the non-existence of overarching artifact concept. However, I believe that pluralism requires one to be anti-realist, neither about artifact kinds nor individual artifacts. Consider that, in the case of species pluralism advanced by Ereshefsky, anti-realism targets only the "category" of species (1998, 114). Here, category means "the class of all species taxa," where species taxa are groupings of organisms (e.g., Homo

doi: 10.48106/dial.v77.i2

One anti-essentialist strategy takes artifact groupings as context-relative. Reydon (2014, 133) defers the task of grouping artifacts to particular relevant epistemic contexts. For Reydon (2014, 137), "These epistemic contexts include academic disciplines such as archeology, art history, cultural anthropology, museum studies as well as engineering and design practices." As explicated in the previous section, etiological functions, intended functions, and maker's intentions fail to provide an overarching account. Given the problems they face, each requires some form of domain restriction, and, thus, for Reydon, to avoid counter-intuitive or arbitrary restrictions, we should settle down the ontological questions only after determining the epistemic context (2014, 141). Thus, the main task of a metaphysician (or, in this case, an anti-metaphysician) is to track how the different artifact concepts are used in the relevant epistemic contexts.

According to Koslicki (2018), pure context-relative solutions of artifact anti-essentialists are not plausible in the case of artifacts. Koslicki (2018, 239) writes, "Empirical questions only arise once we have taken as fixed that screwdrivers are primarily intended to be used by agents who wish to engage in certain kinds of actions, viz., to tighten and loosen screws." This implies that we engage with artifacts not on an explanatory basis but on practical grounds (Koslicki 2018, 239–240). For Koslicki, while we engage with the members of natural kinds to discover their shared properties, what it means for an entity to be an artifact is something we decide before we engage with the candidate entities.

Reydon agrees with Koslicki that the metaphysics of artifacts primarily aims at specifying the general nature of artifacts before we engage with them. However, Reydon (2014, 141) argues that metaphysical approaches, so far, have failed to agree on how to specify the general nature of artifacts; that's why it is better left "open." One implication of leaving the nature of artifacts open is that if metaphysical approaches are far from settling on the general nature of artifacts, we should better track how epistemic contexts fare with artifacts; only then can it be decided whether "an overarching metaphysics of artifact kinds is feasible or a pluralist metaphysics is required" (Reydon 2014,

sapiens) (Ereshefsky 2007, 404). Ereshevsky remarks that biologists and philosophers discuss the definition of the species category when they discuss the definition of "species" (2007, 404). Thus, species pluralism only rejects that there is a single species category without eliminating species taxa. Similarly, I think artifact concept pluralism needs only to reject that there is a single artifact category without eliminating artifact kinds from the picture. The pluralism I outline in this paper modestly suggests that there are at least four ways of grouping entities into artifact kinds.

142). Agreeing with Reydon, I believe context relativity can help us solve the definitional and extensional problems artifact essentialism faces. However, I do not believe that the solution is purely epistemological. In the remainder of this paper, I will argue for an epistemically informed pluralist metaphysics for which Reydon's discussion paves the way. Once I explicate the form of artifact pluralism I have in mind, I will qualify this claim in section 3. For now, note the following points by Mag Uidhir and Magnus that make pluralism suitable for both the species concept and the art concept. I adapt the following points for artifacts.

> Multiple concepts are profitably used by practitioners [1]… Even without a settled [artifact] concept, we are able to agree on the rough boundaries of many [artifact kinds] [2]… [N]o [overarching] concept can profitably apply to all instances [3]… Some of the concepts involve an arbitrary fineness of grain [4]… (Mag Uidhir and Magnus 2011, 90)

Artifact anti-essentialists seem to endorse [1] and [3]. They use [2] to argue that the nature of the artifact concept is better left open. However, they miss the fact that not only do we agree on the rough boundaries of many artifact kinds but also on the ways individual artifacts can be grouped under those artifact kinds. The pluralism I motivate in section 3 is also similar to the anti-essentialist proposals in spirit. I take that there is no single way of dividing the artifactual world. Recent theories concentrate on at least four productive artifact concepts: *morphological artifact concept*, *purely intentional artifact concept*, *intentionalist functional artifact concept*, and *residual artifact concept*. I argue that even though none of these concepts are extensionally or definitionally unproblematic, they still play distinct yet significant roles both in ordinary talk and other disciplines. Instead of completely withdrawing from classificatory aims or leaving the nature of artifacts unspecified, I suggest that by adopting artifact concept pluralism, we can rather focus on the merits of artifact concepts individually. For now, I will turn to another anti-essentialist account that might be based on Richard Boyd's Homeostatic Property Cluster (HPC) view, which aims to account for the extensional and definitional problems artifact essentialism suffers from.

doi: 10.48106/dial.v77.i2

2.1 *The Homeostatic Property Cluster View*

Reydon (2014) outlines the second anti-essentialist strategy by considering the possibility of artifact kinds being homeostatic property clusters. But he does not expand on it. I think it would be informative to explicate the HPC view briefly and contrast a possible anti-essentialist view based on the HPC view with the pluralistic metaphysics I have in mind.

Boyd (1999) develops the HPC view for natural kinds. According to the HPC view, members of a certain kind are not united by virtue of necessarily instantiated essences but by virtue of similarities. The similarities among the members of a kind are stable enough to sustain our taxonomical practices. Furthermore, these similarities are not clustered arbitrarily; as Boyd (1999) argues, they result from some "underlying homeostatic mechanisms." One advantage of the HPC view over essentialist proposals might be that it accounts for the flexibility and change in both natural kinds and artifact kinds. The reason is that the HPC view takes the nature of species as open. This means that the HPC view takes the nature of species, contrary to traditional species concepts, is not fixed by some essential properties (Reydon 2014, 134).

However, a quick concern regarding the kind membership conditions arises against the HPC view: How do we assess whether a given organism or an artifact belongs to a certain kind? The answer is not straightforward. The HPC view suggests that there is a property cluster associated with a kind. The properties are not necessary or essential to a given cluster because it can lose some of the associated properties or gain others over time (Reydon 2014, 134). Furthermore, Boyd (1999, 143) claims that not all members of a kind need to instantiate all the properties of a given. For instance, assuming that the kind *chair* has the functional property of *seating a single individual* necessarily, a functional essentialist would expect all individual chairs to have *that* functional property. However, since the HPC view takes properties as neither essential nor necessary when adapted to artifact kinds, the view admits the possibility of non-functional chairs. Thus, an exhibition chair or a malfunctioning chair (or a chair beyond repair) can be considered as a member of the *chair* kind. The reason is that the HPC view still seems to work if artifacts instantiate only some properties associated with an artifact kind. Adapting the HPC view to artifacts, one can leave which conditions are minimally necessary and sufficient for an artifact to be a member of an artifact kind as unspecified. Although there are no minimally fixed necessary and sufficient conditions that entities need to satisfy, this still does not mean that the nature of artifact kinds is determined arbitrarily. Similar to the case with species, according to the HPC view, the properties associated with a certain artifact kind might result from certain causal-historical relations. These causal-historical relations might, for instance, include the reproductive history of an artifact kind being selected for a certain intended function over a certain period, which, in turn, might not result in associated properties as stable as in the natural kinds. However, this might be the price an anti-essentialist who argues in the line of the HPC view might be willing to pay to account first for the extensional problem artifact essentialism faces and second for the evolutive nature of artifact kinds. One benefit, or for some philosophers an additional cost, of the HPC view is that this form of anti-essentialist account, in turn, might admit accidental creations as well as byproducts that lack intentional properties. Simply because, in this view, artifact kinds do not have their associated properties necessarily or essentially.

Although an anti-essentialist view advanced in these lines seems to account for the extensional problems, the cost is worrying. Eliminating the necessary and essential features from artifact kinds leaves us with vague boundaries, as Reydon (2014, 140) acknowledges: “the HPC view fails to provide membership criteria for kinds.” I believe this cost stems partly from assuming the monistic framework as the backdrop because, according to anti-essentialists, if it is not possible to come up with an extensionally adequate overarching artifact concept, then the nature of the overarching artifact concept should be left open.

Consider the following case with anti-essentialism about art concepts. To account for the revolutionary artworks of the twentieth century that defied the limits attributed to the preceding artworks and art traditions, Weitz (1956) argues that we should regard art as an open concept. This does not mean that the nature of art is lacking; rather, it means that there is not any property such that it is necessary for something to be an artwork (Mag Uidhir and Magnus 2011). Similarly, an anti-essentialist view based on the HPC view proposes to account for the flexibility that artifact kinds show at the cost of denying necessary properties. However, just as being an artwork seems to require something to be an artifact, being an artifact seems to require, at least, one necessary property: *being mind-dependent*.

If artifact kinds are not necessarily mind-dependent, in other words, if artifact kinds do not require the presence of agents with cognitive capabilities, then there seems to be no basis for discarding the swamp and modal cases from our artifact ontology (Juvshik 2021a). Swamp artifact cases are cases in

doi: 10.48106/dial.v77.i2

which an entity structurally similar to paradigm cases of artifacts comes into existence by sheer luck. Modal artifact cases are artifact cases occurring in a possible world that lacks agents with cognitive capabilities (Juvshik 2021a). A proponent of the HPC view might respond to modal and swamp cases by claiming that those cases lack the causal and historical mechanism required for the existence of the members of the HPC clusters. However, this answer is in tension with the principle claim of the HPC view. Consider the following case: Due to a strange accident of nature, a swamp village comes into existence at time *t*. Then, what would preclude one from arguing that the nature of the *village* kind is changed in a way that, after *t*, the *village* kind does not have *being mind-dependent* among its associated properties? I can imagine that the proponent of the HPC view might deny that a single case suffices by itself to change the nature of an artifact kind. However, it is not hard to twist the example so that many modal and swamp villages come into existence over a certain period of time. The point is that I do not see a reason why sufficient frequency of modal and swamp cases would not participate in determining the associated properties of a given artifact kind. As a response, one can insist on the necessity of causal links between human activity and the artifact kinds, but this undermines the HPC view's main thrust as a form of anti-essentialism.

The pluralism I outline in the next section shares the main motivation of an anti-essentialist account based on the HPC view briefly outlined in this paper. That is, to account for the extensional problems without restricting the scope of the term artifact. However, instead of completely eliminating necessary or essential features from the picture, I suggest that we should adopt pluralism without giving up on the mind-dependence condition. Pluralism takes note of the benefits of the artifact concepts individually. Moreover, there are only a limited number of candidate artifact concepts that direct us to fruitful taxonomic practices.

3 Motivating Pluralism

It is not surprising that a single characterization cannot easily capture the nature of all artifacts. This is already implied in many philosophers' discussions. For instance, Thomasson (2014, 46) writes, "The very term 'artifact' is itself used quite loosely, and in many different ways, so there may be no single characterization of what is essential to artifacts that fits best." Bloom, in a similar vein, states that intentions provide the best source for what is essential to artifacts, but not the one that is exactly correct (1996, 20). However, the background assumption of monism remains unchallenged despite the extensional problems monism leads to.

In this section, by outlining how species and art concept monism leave out other widely used senses of these concepts, I aim to draw a parallel to the artifact concept. I argue that in the case of the artifact concept, too, the multiplication of senses is not a vice but an advantage. However, this does not necessarily lead us to an unrestricted proliferation of the senses. Classifications such as "objects that can be used either as doorstops or as cleaning supplies" do not guide us to a useful concept (Koslicki 2008, 202).

3.1 *Pluralism in Other Fields and Artifact Concept Pluralism*

Biology provides many different species concepts, such as the ecological species, the phylogenetic species, and the biological species, just to name a few. Ereshefsky (1998) picks out three prominent species concepts that are used by biologists. However, different versions of each concept have pitfalls that leave certain organisms or significantly shared characteristics of those organisms out of the picture.

The phenotypical (i.e., morphological) species concept uses exhibited characteristics of organisms to sort them into species at a given time while ending up disregarding the evolutionary history of species. The biological species concept sorts organisms according to their sexual reproductive capabilities, simply leaving out asexual organisms that reproduce by other means (e.g., vegetative reproduction). The phylogenetic species concept traces the evolutionary ancestry of organisms to situate species in the evolutionary tree of life; however, due to the evolution, the phylogenetic concept does not provide a stable taxonomy (Ereshefsky 1998, 104–106; Mag Uidhir and Magnus 2011, 89).

Similarly, Mag Uidhir and Magnus (2011) argue that there are at least four distinct art concepts that are gainfully used by philosophers of art. These concepts do not overlap while agreeing in many cases. The aesthetic art concept emphasizes the formal properties of artworks and provides a valuable source of information primarily for perception-related cognitive inquiries. The historical art concept emphasizes the historical properties of artworks, useful for historical inquiries. The conventional art concept traces the norms governing the art world's institutions and practices, providing significant information for sociological and anthropological studies. The communicative art concept focuses on the "representational, semantic, or expressive content"

doi: 10.48106/dial.v77.i2

of artworks, serviceable for learning and emotion-related cognitive inquiries (Mag Uidhir and Magnus 2011, 92).

According to Mag Uidhir and Magnus (2011, 92), in both types of pluralism, insisting on monism ends up in a parochial understanding of the relevant domains. Arguing for a single overarching concept disregards the other fruitful senses of both the species concept and the art concept. As explicated above, for instance, in the case of the species concept, the biological species concept does not range over asexual organisms, whereas the phenotypical species concept does. Similarly, in the case of the art concept, the conventional art concept excludes outsider art, whereas the aesthetic art concept can range over those cases (Mag Uidhir and Magnus 2011, 92). However, admitting pluralism does not mean that all senses of art or species are fruitful. The relevant senses that pluralism should include are epistemically informed; in other words, these concepts must already be in use among the practitioners (e.g., biologists, art critics, historians, and philosophers of art). Mag Uidhir and Magnus (2011, 90) name this form of pluralism "responsible pluralism" to distinguish it from "anything goes" approaches. Granted that an epistemically informed responsible pluralism is possible for both species and art concepts, in the remainder of this section, I try to motivate a similar form of pluralism for the artifact concept and defend it against possible objections in section 4.

My aim in this paper is to outline a rough guide for artifact concept pluralism. It is enough for pluralism if I can show at least two different artifact concepts are well-motivated. I state four. These are morphological, purely intentional, intentionalist functional, and residual artifact concepts. I choose to focus on these four concepts as I believe the combination of these four concepts provides the best result extensionally. Before turning to the relevant domains and purposes, let me first briefly state the candidate concepts I have in mind:

> Morphological Artifact Concept. Considerations regarding shape are undeniably important when it comes to artifacts. According to Malt and Sloman (2007), artifact categorization is not settled on a single feature that artifacts display. Shape, function, and intended category membership all play a role in our various ways of artifact groupings. Shape plays an indispensable role in Franssen and Kroes's (2014) and Elder's (2007) respective artifact ontologies. Franssen and Kroes's and Elder's fine-grained ontologies can accommodate only highly specific artifact kinds, such as

Pasha Seatimer grand modèle automatique Cartier watch (Franssen and Kroes 2014) and *Eames 1957 desk chair* (Elder 2007). Under the essentialist framework, the shape is mixed into functions and makers' intentions. This, I believe, stems from the monistic assumption in the background. This need not be the case if we shift the framework to pluralism. I suggest that a morphological concept needs to be fleshed out in order to accommodate morphological classifications in certain domains and inquiries. For instance, in archeology, classifications based on morphological properties play a crucial role in artifact classification. These classifications do not necessarily involve reference to makers' intentions or to functions. Archeologists Kelly and Thomas (2014, 99–100) remark that morphological classification is highly used by practitioners alongside functional and temporal classifications. Depending on the task and the object at hand, an archeologist can classify an object under a coarse-grained grouping, such as "flat-bodied-with-protruding-legs" (Kelly and Thomas 2014, 99–100). According to Kelly and Thomas (2014, 100), morphological classification requires an item to show similarity in displayed characteristics; also, the item should be laden with information regarding the past culture.

Thus, under the morphological artifact concept, we can say that artifacts are grouped into artifact kinds based on their displayed similarities to other members of artifact kinds. These objects need not have functional properties or be intentionally created, but they are mind-dependent. The notion of similarity is vague, and it is left unspecified purposefully as some variations of the morphological concept may require more strict similarity and thus result in a finer-grained classification, whereas others, depending on the inquiry, may involve a coarse-grained classification (Vermaas and Houkes 2013; Franssen and Kroes 2014; Elder 2007).

Purely Intentional Artifact Concept. Intentions provide a better understanding of the normative aspects of artworks compared to the other two concepts. For instance, David Friedell (2020) argues that since Bruckner's unfinished Ninth Symphony is intended to be produced as a member of the *symphony kind* in the Western classical music tradition, a subsequent composer could finish the work posthumously. This is because the relevant convention (e.g.,

doi: 10.48106/dial.v77.i2

Western classical music tradition) allows for such a change in a given symphony while sustaining the work's identity. Thus, it seems that what is essential to artworks is determined by the intentions of their makers and the conventions in which these intentions are situated. If that's the case, then a purely intentional concept would better capture the nature of these artifacts. Under the purely intentional concept, we can say that artifacts are mind-dependent objects that are made to be a member of a certain artifact kind. These objects may or may not have functional properties (Thomasson 2003, 2007, 2014; Juvshik 2021a).

Intentionalist Functional Artifact Concept. The intentionalist functional concept successfully sorts artifacts that show significant form variations under the same kind (Baker 2004, 2007; Dipert 1993; Hilpinen 1992; Evnine 2016). However, it cannot be profitably used in the case of artworks (e.g., conceptual art). Intended functions are used both in folk classification and engineering practices. Thus, under the intentionalist functional artifact concept, artifacts are mind-dependent objects that are made to perform certain functions.

It must be noted that the list of artifact concepts briefly elaborated above is not exhaustive; it only aims to cover the widely used senses of artifact concept. As expected, these artifact concepts share many of their extensions. In the case of species and art concepts, people can use "species" and "art" distinctly without specifically stating the concept they use (Mag Uidhir and Magnus 2011, 92). Similarly, in the case of artifact concepts, folk classifications, as well as social sciences and engineering practices, use the artifact concept quite liberally.

Residual Artifact Concept. One important result of accepting pluralism is that pluralism accounts for the problematic cases of artifacts such as byproducts and residues. Woodchips, sawdust, and midden heaps are all indiscriminately considered to be artifacts by archeologists and anthropologists. Since these artifacts lack shared morphological structure, function, or intentional features, they do not fit neatly into the previous artifact concepts, and so they are ruled out by monists.

By shifting the focus, we do not have to settle down the problem cases as "spoils to the victor" (Juvshik 2021a). The winner-take-all approach flat-out rejects the problematic senses of the artifact concept. However, in a pluralistic framework, we can fruitfully approach specific kinds of problem cases within the boundaries of a specific artifact concept and see to what extent that concept manages to account for such cases (Mag Uidhir and Magnus 2011, 92–95). Many consider artworks as artifacts (Dickie 1984; Levinson 2007; Mag Uidhir 2013). If some artworks are not functional, then we can better approach the philosophy of art with a purely intentional artifact concept as the backdrop.

> The substantive necessity of intention-dependence should be seen as posing a philosophical constraint not just for any theory of art but also for the philosophy of art itself. That is, we ought to expect any and all philosophical enquiry into art and its associated *relata* (i.e., the nature of art, artworks, art forms, art practices, art ontology, art interpretation and evaluation, etc.) to yield conclusions at least minimally consistent with, if not directly informed by, the basic background assumption that intention-dependence is a *substantive* necessary condition for being art. (Mag Uidhir 2013, 5–6, emphasis in the original)

According to Mag Uidhir, the intention to create an artwork provides significant information regarding the nature of that artwork. Thus, even though a certain snowy hill may have more exciting aesthetic properties than Pieter Bruegel's *Hunters in the Snow*, with the purely intentional artifact concept in mind, we can rule out such cases since they are not artifacts, hence not artworks.

This means that depending on the inquiry, we may need distinct concepts to classify certain artifacts. For instance, in the historical inquiries conducted by archeologists, shape may play a crucial role in evaluating the cultural significance of the found object. Archeologist Steven Mithen (2007, 290) notes that "Polly Wiessner (1983), for instance, studied the arrowheads of the !Kung bushmen of Southern Africa and documented how their specific shapes are not only effective at killing game but define individual and social identity." !Kung bushmen's arrowheads thus belong to different artifact kinds under the morphological artifact concept. In this case, it is not the function, but the shape plays a more important role in determining the membership conditions. One may object that it is not the shape itself, but the intention to create an arrowhead that has a certain shape is what plays this role. However, we can imagine a scenario in which a !Kung bushman can find an arrowhead-shaped stone in the forest; still, that arrowhead would provide a valuable source of information for archeologists. Furthermore, archeologists may not only classify found objects as artifacts but also accidental or unintentional creations, such as woodchips that result from making wooden spears, are considered to be artifacts (Fullagar and Matheson 2014).

Three things should be noted. First, the variations of the morphological concept result in arbitrary fineness of grain. For instance, depending on the inquiry and context, artifacts can be partitioned into fine-grained artifact kinds, such as *Pasha Seatimer grand modèle automatique Cartier watch* (Franssen and Kroes 2014, 78), or a coarse-grained classification, such as *flat-bodied-with-protruding-legs* (Kelly and Thomas 2014, 100). Counter-intuitively, as the !Kung Bushmen case exemplifies, the morphological concept might admit accidentally created or unmodified objects as artifacts, granted that they share a similar morphological structure to members of a certain artifact kind and show cultural significance. The intentionalist functional concept provides a stable taxonomy used both in folk classification and engineering practices; however, it leaves out artifacts that lack function (e.g., artworks). The purely intentional concept performs better in the case of artworks compared to the other two concepts. Given that none of the concepts can single-handedly capture the plurality of artifacts, then this can give us a reason to challenge the monistic framework itself.

Second, even though the pluralism I formulated suggests four concepts, these are not the only viable concepts. Depending on the context or inquiry, a more refined concept might be needed. So, even though I strongly suggest adopting pluralism in the case of artifacts, my wish is not to leave it static. There is no reason to reject that we might require more concepts in the future as taxonomic practices change. Consequently, a pluralistic framework that methodologically privileges actual practices should be flexible enough to capture the dynamicity of the taxonomic practices.

Lastly, all viable artifact concepts share a necessary condition: *being mind-dependent*. Given the methodology, this condition is needed to account for the current taxonomical practices. As our artifact practices dictate, the items that the concepts pick out should be those that have causal links to the human culture. That is why pluralism cannot afford to admit swamp and modal cases to the artifact ontology. To rule out such cases, therefore, pluralism needs to adopt mind-dependence as a necessary condition.

4 Objections

Pluralism seems to avoid the problems monism faces with relative ease. As we see in the previous section, pluralism shifts the focus from providing the best possible overarching artifact concept to retaining the merits of four individual artifact concepts. By shifting focus, pluralism offers a greater scope. Furthermore, pluralism does not need to appeal to the definitional restrictions to which essentialist accounts commit. However, the general worries regarding the nature of pluralistic approaches make pluralism undesirable. Here, I defend pluralism against three objections one can raise to make it more desirable.

First, one may object by arguing that adopting pluralism or any disjunctive supplementation brings its own complexities, and, thus, instead of clarifying the concepts, pluralism might end up adopting the "disadvantages of those concepts" (Vermaas and Houkes 2003, 275). Furthermore, Ockham's Razor dictates we eliminate the murkier senses of a notion, not propagate them—the simpler, the better. However, the artifactual world is not less divergent than the biological world and the art world. Considering the heterogeneity of the artifactual world, I think a unified account is possible only in the case of ad hoc domain restrictions. Even in the case of domain restrictions (e.g., technical artifacts), there is a considerable amount of evidence from psychological research and engineering practices that led Vermaas and Houkes (2013) to argue for pluralism in the categorization of technical artifacts.

Vermaas and Houkes (2013) argue that certain classificatory practices in engineering coincide with psychological findings presented in Malt and Sloman (2007). Malt and Sloman's experiment shows that there are, roughly, three major features that play significant roles in artifact classification: form (i.e., shape), functions, and intended category membership. Correspondingly, from their experience in the philosophy of technology, Vermaas and Houkes (2013) formulate three types of categorization principles for technical artifacts: id-made-product categorization, functional and goal categorization, and use-plan and make-plan categorization. Even though there are certain similarities worth mentioning, I will not get into details of Houkes and Vermaas's account since here I attempted to motivate pluralism not only for technical artifacts but artifacts in general and across different disciplines. Each artifact concept I briefly pointed out provides partial partitioning; in other words, the success of a concept is not constrained by its scope, as each concept can only range over a certain portion of artifacts depending on the inquiry.

doi: 10.48106/dial.v77.i2

Second, one may point out that pluralism only amounts to a verbal dispute and claim that it is only a *linguistic fact* that we use distinct artifact concepts. So, according to this objection, pluralism only tracks people's different usage of the term artifact rather than metaphysically important features, and there might be a metaphysically salient use of "artifact." For instance, to account for the metaphysically salient features of artifacts, Dipert (1993, 23ff.) suggests a tripartite distinction between *tools*, *instruments*, and *artifacts proper*. Leaving out the details, according to this distinction, artifacts are items that are made to be recognized as functional objects; as Dipert (1993, 31) puts it, they are "distinctively social." However, his conceptual distinction results in an even more restrictive artifact concept than the restrictions we have seen so far ("technical artifact" and "essentially artifact kind"). Given that the aim is to account for taxonomical practices, the same extensional worries that apply to the previous accounts *mutatis mutandis* apply in Dipert's case. So, Dipert's distinction is not helpful. Going back to verbal dispute objection. Since pluralism tracks important metaphysical distinctions, I think this objection does not pose a threat to artifact concept pluralism. For instance, residual artifacts are not produced with the intention to create those items; they also do not have a specific morphological structure, so they are metaphysically different from intentionally created functional objects such as computers and airplanes. So, we need at least two different concepts to account for the metaphysical differences of these cases.

Lastly, one may doubt the accuracy of the analogy between species/art concept pluralism and artifact concept pluralism along the following lines: Our aim with artifact classifications is not primarily inferential or explanatory, whereas taxonomy for species and art concept is provided by the relevant specialists (Koslicki 2018, 239). Thus, our artifact classifications need not be based on specialists' vocabulary. I agree that in the case of artifacts, folk classifications are not ultimately determined by the relevant disciplines and practices. For instance, I would not wait for archeologists' validation for calling my favorite sitting device a "chair," nor do I think I would be in error if that device turns out not to be a chair in some engineers' classifications. However, the pluralism explored in this paper aims not only to describe folk classifications but give a more encompassing picture across different domains in which the term artifact plays an important role. Pluralism aims to provide distinct concepts for different inquiries and hence be an alternative to the arbitrary domain restrictions that stem from artifact concept monism. By changing the question from "What concept of artifact can best capture all cases?" to "What specific artifact concept can best capture the specific problem cases?" we need not approach a urinal, Duchamp's *Fountain*, a toast, archeological woodchips, and nuclear reactors under an overarching artifact concept (Mag Uidhir and Magnus 2011, 92). Otherwise, as Preston (2014) points out, the gap between metaphysicians' and other disciplines' classificatory practices will continue to widen. This, in turn, may result in the philosophical term of artifact having no informative use outside of philosophy.

5 Conclusion

Artifact essentialists focused on finding an artifact essence. Artifact anti-essentialists claimed that there is none. In this paper, I challenged the monistic assumption that pervades the debate. I argued against artifact concept monism first by showing that the prominent essentialist proposals currently at play suffer from major extensional and definitional problems. Second, I aimed to show that current anti-essentialist accounts suffer from eliminating all necessary properties, which results in the proliferation of cases, as shown by the modal and swamp cases. Metaphysical literature on artifacts is a productive field. There are both compelling essentialist and anti-essentialist proposals yet to come. Adopting a pluralistic framework motivates a new focus on the neglected aspects of the artifactual world. I pointed out some of those aspects. Obviously, artifact concept pluralism invites many questions that I could not touch upon or give a detailed answer to. It requires a greater elaboration to properly flesh out the details; however, considering the significantly diverse roles artifacts play in our lives, I believe such effort is both needed and fascinating.*

Alper Güngör
0000-0001-7146-6834
McGill University
alper.gungor@mail.mcgill.ca

* I would like to thank two anonymous referees for their detailed comments. For their helpful feedback on earlier drafts, I am especially grateful to Nurbay Irmak, Kathrin Koslicki, Sun Demirli, Ata Türkoğlu, Aslı İdil Kaynar, and Alex Stooshinoff.

doi: 10.48106/dial.v77.i2

References

Baker, Lynne Rudder. 2004. "The Ontology of Artifacts." *Philosophical Explorations* 7(2): 99–111, doi:10.1080/13869790410001694462.

—. 2007. *The Metaphysics of Everyday Life: An Essay in Practical Realism*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press, doi:10.1017/cbo9780511487545.

—. 2008. "The Shrinking Difference Between Artifacts and Natural Objects." *APA Newsletter on Philosophy and Computers* 7(2), https://cdn.ymaws.com/www.apaonline.org/resource/collection/EADE8D52-8D02-4136-9A2A-729368501E43/v07n2Computers.pdf.

Bird, Alexander and Tobin, Emma. 2022. "Natural Kinds." in *The Stanford Encyclopedia of Philosophy*. Stanford, California: The Metaphysics Research Lab, Center for the Study of Language and Information, https://plato.stanford.edu/archives/spr2022/entries/natural-kinds/.

Bloom, Paul. 1996. "Intention, History, and Artifact Concepts." *Cognition* 60(1): 1–29, doi:10.1016/0010-0277(95)00699-0.

Boyd, Richard N. 1999. "Homeostasis, Species, and Higher Taxa." in *Species: New Interdisciplinary Essays*, edited by Robert A. Wilson, pp. 141–185. Cambridge, Massachusetts: The MIT Press, doi:10.7551/mitpress/6396.003.0012.

Davies, David. 2004. *Art as Performance*. Oxford: Blackwell Publishers, doi:10.1002/9780470774922.

Dickie, George. 1984. *The Art Circle: A Theory of Art*. New York: Haven Publications.

Dipert, Randall R. 1993. *Artifacts, Art Works and Agency*. Philadelphia, Pennsylvania: Temple University Press.

Elder, Crawford L. 2007. "On the Place of Artifacts in Ontology." in *Creations of the Mind: Theories of Artifacts and Their Representation*, edited by Eric Margolis and Stephen Laurence, pp. 33–51. Oxford: Oxford University Press, doi:10.1093/oso/9780199250981.003.0003.

Ereshefsky, Marc. 1998. "Species Pluralism and Anti-realism." *Philosophy of Science* 65(1): 103–120, doi:10.1086/392628.

—. 2007. "Foundational Issues Concerning Taxa and Taxon Names." *Systematic Biology* 56(2): 295–301, doi:10.1080/10635150701317401.

Evnine, Simon J. 2016. *Making Objects and Events: A Hylomorphic Theory of Artifacts, Actions, and Organisms*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780198779674.001.0001.

Franssen, Maarten and Kroes, Peter. 2014. "Artefact Kinds, Ontological Criteria and Forms of Mind-Dependence." in *Artefact Kinds: Ontology and the Human-Made World*, edited by Maarten Franssen, Peter Kroes, Thomas A. C. Reydon, and Pieter E. Vermaas, pp. 63–83. Synthese Library n. 365. Dordrecht: Springer, doi:10.1007/978-3-319-00801-1_5.

Friedell, David. 2020. “Why Can’t I Change Bruckner’s *Eighth Symphony*?” *Philosophical Studies* 177(3): 805–824, doi:10.1007/s11098-018-1207-3.

Fullagar, Richard and Matheson, Carney. 2014. “Stone Tool Usewear and Residue Analysis.” in *Encyclopedia of Global Archaeology*, edited by Claire Smith, pp. 7062–7065. New York: Springer, doi:10.1007/978-1-4419-0465-2_842.

Grandy, Richard E. 2007. “Artifacts: Parts and Principles.” in *Creations of the Mind: Theories of Artifacts and Their Representation*, edited by Eric Margolis and Stephen Laurence, pp. 18–32. Oxford: Oxford University Press, doi:10.1093/oso/9780199250981.003.0002.

Hilpinen, Risto. 1992. “On Artifacts and Works of Art.” *Theoria* 58(1): 58–82, doi:10.1111/j.1755-2567.1992.tb01155.x.

Juvshik, Tim. 2021a. “Artifacts and Mind-Dependence.” *Synthese* 199(3-4): 9313–9336, doi:10.1007/s11229-021-03204-6.

—. 2021b. “Function Essentialism About Artifacts.” *Philosophical Studies* 178(9): 2943–2964, doi:10.1007/s11098-020-01594-w.

Kelly, Robert L. and Thomas, David Hurst. 1991. *Archaeology: Down to Earth*. 1st ed. Belmont, California: Wadsworth Publishing Co.

—. 2014. *Archaeology: Down to Earth*. 7th ed. Belmont, California: Wadsworth Publishing Co. First edition: Kelly and Thomas (1991).

Kornblith, Hilary. 1980. “Referring to Artifacts.” *The Philosophical Review* 89(1): 109–114, doi:10.2307/2184866.

Koslicki, Kathrin. 2008. *The Structure of Objects*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199539895.001.0001.

—. 2018. *Form, Matter, Substance*. Oxford: Oxford University Press, doi:10.1093/oso/9780198823803.001.0001.

Kripke, Saul A. 1980. *Naming and Necessity*. Oxford: Basil Blackwell Publishers.

Levinson, Jerrold. 2007. “Artworks as Artifacts.” in *Creations of the Mind: Theories of Artifacts and Their Representation*, edited by Eric Margolis and Stephen Laurence, pp. 74–82. Oxford: Oxford University Press, doi:10.1093/oso/9780199250981.003.0005.

Locke, John. 1689. *An Essay concerning Human Understanding*. London: Eliz. Holt, for Thomas Basset, at the George in Fleet Street.

Lopes, Dominic McIver. 2009. *A Philosophy of Computer Art*. London: Routledge, doi:10.4324/9780203872345.

Mag Uidhir, Christy. 2013. *Art and Art-Attempts*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199665778.001.0001.

Mag Uidhir, Christy and Magnus, P. D. 2011. “Art Concept Pluralism.” *Metaphilosophy* 42(1/2): 83–97, doi:10.1111/j.1467-9973.2010.01678.x.

Malt, Barbara C. and Sloman, Steven A. 2007. “Artifact Categorization: The Good, the Bad, and the Ugly.” in *Creations of the Mind: Theories of Artifacts and Their

doi: 10.48106/dial.v77.i2

Representation, edited by Eric Margolis and Stephen Laurence, pp. 85–123. Oxford: Oxford University Press, doi:10.1093/oso/9780199250981.003.0006.

Margolis, Eric and Laurence, Stephen, eds. 2007. *Creations of the Mind: Theories of Artifacts and Their Representation*. Oxford: Oxford University Press, doi:10.1093/oso/9780199250981.001.0001.

Mithen, Steven. 2007. “Creations of Pre-modern Human Minds: Stone Tool Manufacture and Use by *Homo habilis*, *heidelbergensis*, and *neanderthalensis*.” in *Creations of the Mind: Theories of Artifacts and Their Representation*, edited by Eric Margolis and Stephen Laurence, pp. 289–311. Oxford: Oxford University Press, doi:10.1093/oso/9780199250981.003.0016.

Preston, Beth. 2013. *A Philosophy of Material Culture: Action, Function, and Mind*. London: Routledge, doi:10.4324/9780203069844.

—. 2014. “Ethnotechnology: A Manifesto.” in *Artefact Kinds: Ontology and the Human-Made World*, edited by Maarten Franssen, Peter Kroes, Thomas A. C. Reydon, and Pieter E. Vermaas, pp. 145–163. Synthese Library n. 365. Dordrecht: Springer, doi:10.1007/978-3-319-00801-1_9.

Putnam, Hilary. 1975a. *Mind, Language and Reality: Philosophical Papers, Volume 2*. Cambridge: Cambridge University Press, doi:10.1017/cbo9780511625251.

—. 1975b. “The Meaning of ‘Meaning’.” in *Minnesota Studies in the Philosophy of Science, Volume VII: Language, Mind, and Knowledge*, edited by Keith Gunderson, pp. 131–193. Minneapolis, Minnesota: University of Minnesota Press. Reprinted in Putnam (1975a, 215–271).

Reydon, Thomas A. C. 2014. “Metaphysical and Epistemological Approaches to Developing a Theory of Artifact Kinds.” in *Artefact Kinds: Ontology and the Human-Made World*, edited by Maarten Franssen, Peter Kroes, Thomas A. C. Reydon, and Pieter E. Vermaas, pp. 125–144. Synthese Library n. 365. Dordrecht: Springer, doi:10.1007/978-3-319-00801-1_8.

Soavi, Marzia. 2009. “Antirealism and Artefact Kinds.” *Techné* 13(2): 93–107, doi:10.5840/techne20091329.

Thomasson, Amie L. 2003. “Realism and Human Kinds.” *Philosophy and Phenomenological Research* 67(3): 580–609, doi:10.1111/j.1933-1592.2003.tb00309.x.

—. 2007. “Artifacts and Human Concepts.” in *Creations of the Mind: Theories of Artifacts and Their Representation*, edited by Eric Margolis and Stephen Laurence, pp. 52–73. Oxford: Oxford University Press, doi:10.1093/oso/9780199250981.003.0004.

—. 2009. “Artifacts in Metaphysics.” in *Philosophy of Technology and Engineering Sciences*, edited by Anthonie Meijers, pp. 191–212. Handbook of the Philosophy of Science n. 9. Amsterdam: Elsevier Science Publishers B.V., doi:10.1016/B978-0-444-51667-1.50012-4.

—. 2014. “Public Artifacts, Intentions, and Norms.” in *Artefact Kinds: Ontology and the Human-Made World*, edited by Maarten Franssen, Peter Kroes, Thomas A. C.

Reydon, and Pieter E. Vermaas, pp. 45–62. Synthese Library n. 365. Dordrecht: Springer, doi:10.1007/978-3-319-00801-1_4.

Vega-Encabo, Jesús and Lawler, Diego. 2014. "Creating Artifactual Kinds." in *Artefact Kinds: Ontology and the Human-Made World*, edited by Maarten Franssen, Peter Kroes, Thomas A. C. Reydon, and Pieter E. Vermaas, pp. 105–124. Synthese Library n. 365. Dordrecht: Springer, doi:10.1007/978-3-319-00801-1_7.

Vermaas, Pieter E. and Houkes, Wybo. 2003. "Ascribing Functions to Technical Artefacts: A Challenge to Etiological Accounts of Functions." *The British Journal for the Philosophy of Science* 54(2): 261–289, doi:10.1093/bjps/54.2.261.

—. 2013. "Pluralism on Artefact Categories: A Philosophical Defence." *The Review of Philosophy and Psychology* 4(3): 543–557, doi:10.1007/s13164-013-0149-0.

Weitz, Morris. 1956. "The Role of Theory in Aesthetics." *The Journal of Aesthetics and Art Criticism* 15(1): 27–35, doi:10.2307/427491.

Wiggins, David. 2001. *Sameness and Substance Renewed*. Cambridge: Cambridge University Press, doi:10.1017/cbo9780511612756.

Wright, Larry. 1973. "Functions." *The Philosophical Review* 82(2): 139–168, doi:10.2307/2183766.

doi: 10.48106/dial.v77.i2

Axiomatization of Galilean Spacetime

Jeffrey Ketland

In this article, we give a second-order synthetic axiomatization $\mathsf{Gal}(1,3)$ for Galilean spacetime, the background spacetime of Newtonian classical mechanics. The primitive notions of this theory are the 3-place predicate of betweenness $\mathtt{Bet}$, the 2-place predicate of simultaneity $\sim$, and a 4-place congruence predicate, written $\equiv^{\sim}$, restricted to simultaneity hypersurfaces. We define a standard coordinate structure $\mathbb{G}^{(1,3)}$, whose carrier set is $\mathbb{R}^4$, and which carries relations (on $\mathbb{R}^4$) corresponding to $\mathtt{Bet}$, $\sim$, and $\equiv^{\sim}$. This is the standard model of $\mathsf{Gal}(1,3)$. We prove that the symmetry group of $\mathbb{G}^{(1,3)}$ is the (extended) Galilean group (an extension of the usual 10-parameter Galilean group with two additional parameters for length and time scalings). We prove that each full model of $\mathsf{Gal}(1,3)$ is isomorphic to $\mathbb{G}^{(1,3)}$.

This article provides a synthetic (and second-order) axiom system, which I call $\mathsf{Gal}(1,3)$, which describes Galilean spacetime and does so categorically.[1] Galilean spacetime is a system $\mathbb{P}$ of points on which three physical geometrical primitives are defined, satisfying certain conditions.[2] Galilean spacetime can be thought of as the background geometry of the system of spacetime events for Newtonian classical mechanics:

---

1 The parameters "1" and "3" in $\mathsf{Gal}(1,3)$ mean: "1 time and 3 space dimensions." Recall that an axiom system is called *categorical* when it has exactly one model up to isomorphism. Second-order Peano arithmetic, PA_2, is categorical, its unique model being $(\mathbb{N}, 0, S, +, \times)$. The proof (essentially given in Dedekind 1888) is that if $M \vDash \mathsf{PA}_2$, we may define using Dedekind's Recursion Theorem a function $\Phi : \mathbb{N} \to \mathrm{dom}(M)$ by $\Phi(0) = 0^M$ and, for all $n \in \mathbb{N}$, $\Phi(n+1) = S^M(\Phi(n))$. The axioms of PA_2 then imply that Φ is a bijection, which is an isomorphism from $(\mathbb{N}, 0, S, +, \times)$ to M. In addition to PA_2, the theory ALG of the complete ordered field is also categorical (essentially given in Huntington 1903; using methods developed in Dedekind 1872; Cantor 1897; Hölder 1901). Various second-order geometrical theories are also categorical. These include the systems denoted $\mathsf{BG}(4)$ and $\mathsf{EG}(3)$ below. Theorems 62 and 63 in appendix B establish the categoricity (and standard models) of these two systems. The proofs are due to Hilbert (1899), Veblen (1904), and Tarski (1959).

2 I think, informally, of a Galilean spacetime *modally*: a *physically possible world* with certain distinguished, or built-in, geometrical (spatio-temporal) relations. Such metaphysical issues, however, don't matter here, as our whole discussion below is about models of $\mathsf{Gal}(1,3)$.

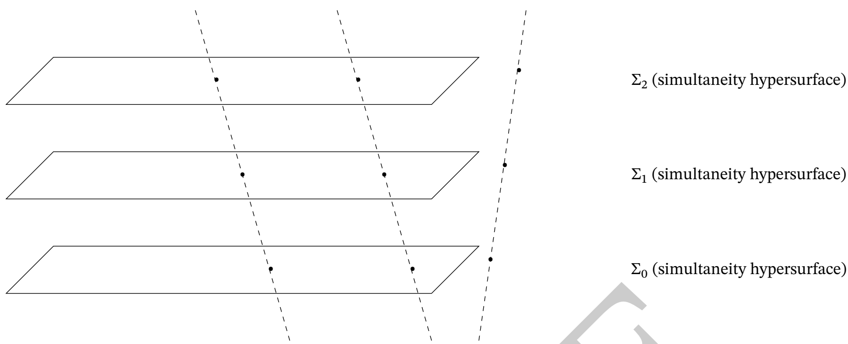

Figure 1: Galilean Spacetime

I shall call the carrier set of Galilean spacetime $\mathbb{P}$: this is the domain of "spacetime points" or "events." Going ahead of ourselves a bit, there are three distinguished physical relations on $\mathbb{P}$. A three-place *betweenness relation B*, which gives the whole system an affine "straight-line" structure;[3] a binary *simultaneity relation* $\sim$, which induces a partition of $\mathbb{P}$ into a system of non-intersecting simultaneity hypersurfaces, $\Sigma_0, \Sigma_1, \ldots$, arranged as a "foliation"; and a special four-place *congruence* relation: this is the four-place *sim-congruence* relation, $\equiv^{\sim}$, which induces three-dimensional Euclidean geometry on each hypersurface.[4]

An especially important subset of straight lines are "time axes": a time axis is a *straight line* in the affine geometry that does not lie within a simultaneity hypersurface. Physically, a time axis is the *trajectory of a material point acted on by no forces*—this is Newton's First Law or the Law of Inertia.[5]

3 It is isomorphic to the standard four-dimensional affine space usually called $\mathbb{A}^4$ (see Gallier 2011), which is gotten from the vector space $\mathbb{R}^4$ by "forgetting its origin." In Gallier's notation, $\mathbb{A}^4$ is $(\mathbb{R}^4, \mathbb{R}^4, +)$, where the first $\mathbb{R}^4$ is the point set, the second $\mathbb{R}^4$ is the vector space, and $+$ is the action of vectors in $\mathbb{R}^4$ on points in $\mathbb{R}^4$. For the reader whose algebra is rusty, the notion of a *group action* is explained nicely in Dummit and Foote (2004, 41), Gallier (2011, 11), or Saunders (2013, 29).

4 A valuable semi-formal mathematical description of Galilean spacetime, incorporating what has just been said, is given in Arnold (1989, chap. 1).

5 Why do material points move (four-dimensionally) along these "grid lines" in Galilean spacetime? The *physical* answer is that such trajectories *minimize the action*. I.e., $\delta \int dt\, (\dot{q})^2 = 0$.

We can bundle the carrier set of Galilean spacetime and the aforementioned three distinguished physical relations on Galilean spacetime together: $(\mathbb{P}, B, \sim, \equiv^{\sim})$. Our aim in this paper is to give a *synthetic* axiomatization of this structure $(\mathbb{P}, B, \sim, \equiv^{\sim})$.[6] This means that, in contrast with analytic geometry, the axioms do not quantify over the reals, introduce a metric function (like a Riemannian metric g_{ab}), or talk about coordinate systems. Instead, the axioms use a number of basic physical predicates on spacetime. And then the existence of special mappings $\Phi : \mathbb{P} \to \mathbb{R}^4$—that is, coordinate systems—becomes a theorem, not an assumption.

Hartry Field (1980) has carefully studied this approach in order to try and vindicate *nominalism*: this is the claim that there are no mathematical objects at all, and insofar as numbers, functions, sets, vector spaces, Lie groups, and so on are used in physics and science more generally, they can be dispensed with. It is the claim that physical theories can, in principle, be replaced with theories that are "nominalistic" and that the normal use of mathematics is "useful *but false*." It is to Field's enormous credit to have pinned down the two essential uses. These are:

> EXPRESSIVENESS. We can express physical laws by, e.g., "$\nabla \cdot \mathbf{B} = 0$" and so on. So, $\mathbf{B}$ is a *mixed function* that maps each point to some *numbers*. As Feynman put it, "From a mathematical view, there is an electric field vector and a magnetic field vector at every point in space; that is, there are six numbers associated with every point" (Feynman, Leighton and Sands 2005, chaps. 20, sec.3).

6 I have tried to write this paper so that it can be read by those unfamiliar with some of the somewhat arcane details of synthetic geometry. A very useful summary of the main ideas behind the construction of coordinate systems may be found in Burgess and Rosen (1997, 102–111). In my view, a very clear and nice introduction to the topic of affine and projective *incidence* geometry is Bennett (1995), where "geometric addition" and "multiplication" of points on a fixed line are explained clearly, and the core result is proved, that the line, with those operations, is a division ring (if Desargues's Theorem is assumed) and a field (if Pappus's Theorem is assumed). Notable reference works more generally are Coxeter (1969) and Hartshorne (2000). A fairly advanced treatment is Borsuk and Szmielew (1960). Tarski's papers (1959; and Tarski and Givant 1999) are very accessible. The first of these sketches the representation theorem for first-order Euclidean geometry and for the second-order Euclidean geometry EG(3) used below. Tarski focuses on the two-dimensional, first-order ("elementary") case. The book Schwabhäuser, Szmielew and Tarski (1983) is very detailed (it is in German, and there is no English translation). Some recent works have implemented Tarski Euclidean geometry in theorem provers, just as one can implement arithmetic, set theory, and type theory in such provers. I have no doubt that this can, in principle, be generalized to our Galilean spacetime geometry and to one or other axiomatization of Minkowski spacetime geometry.

doi: 10.48106/dial.v77.i2

> PROOF-THEORETIC. Mathematically reasoning is generally conservative over non-mathematical premises, but using mathematics, we can get "quicker proofs" of a non-mathematical conclusion C from a non-mathematical premise P.

As regards the second, in mathematical logic, this is called "speed-up," and it was discovered by Kurt Gödel (1936) as a spin-off from his incompleteness results. Perhaps the most remarkable example of this phenomenon was given in Boolos (1987), a first-order valid inference with a short mathematical proof (it uses second-order comprehension), but whose shortest purely logical derivation, using the rules for the connectives and quantifiers, has vastly more symbols than the number of baryons in the observable universe.[7]

The best survey, and overall evaluation, of a large variety of nominalist approaches for both mathematics and science is Burgess and Rosen (1997).[8] I'm not recommending this as an approach to studying the geometrical assumptions of physical theories, as my own view here is the usual mathematical realist view ("useful *because true*"). Indeed, *Riemannian geometry* is here to stay! Riemannian geometry provides incredible flexibility by assuming the existence of a metric tensor g_{ab} on spacetime.[9] However, for the two special cases of Galilean spacetime and Minkowski spacetime, the *synthetic* approach helps provide a nice example of how the physics (i.e., the basic physical relations: betweenness, congruence, and so on) and mathematics (i.e., real numbers, coordinate systems, vector spaces, and so on) get "entangled."

The basic machinery for the introduction of coordinates is the *Representation Theorem*. Given a synthetic structure satisfying a series of conditions, one proves the existence of an isomorphism to a standard coordinate structure:[10]

$$\Phi : \text{synthetic structure} \to \text{coordinate structure}. \tag{1}$$

7 See Ketland (2022) for a formalization of the quicker proof in the Isabelle theorem prover.

8 In that book, Field's approach is called "geometrical nominalism." A technical difficulty that arises for Field's program in Field (1980) concerning the problem of maintaining *both* a conservativeness condition *and* representation theorems is briefly described in remark 14 below.

9 As Einstein showed, the laws of gravitation amount to certain differential equations constraining g_{ab} and the energy-momentum tensor T_{ab}. The "low energy limit" of Einstein's field equation is Newton's Law of Gravitation. Two standard textbooks on general relativity are Weinberg (1972) and Wald (1984).

10 Cf. Terence Tao (2008): "More generally, a coordinate system Φ can be viewed as an isomorphism $\Phi : A \to G$ between a given geometric (or combinatorial) object A in some class (e.g. a circle), and a standard object G in that class (e.g. the standard unit circle)."

That is, the isomorphism Φ takes each point p in the synthetic structure to its coordinates $\Phi^i(p)$ (usually in $\mathbb{R}^n$) in such a way that a distinguished synthetic relation R holds for $p, q, \ldots$ iff a separately defined coordinate relation R' holds for $\Phi(p), \Phi(q), \ldots$ (see, for example, (5) below). Because the synthetic and coordinate structures are *isomorphic*, the latter is a kind of *map* or *representation* of the former: they share the same *abstract* structure.[11]

However, historically, the analysis of Galilean spacetime did not proceed like this. Modern analysis of Galilean spacetime (sometimes called "neo-Newtonian" spacetime or just "Newtonian spacetime") was developed using the differential geometry methods developed to study General Relativity: what are now called "*relativistic spacetimes*." This began in the 60s and 70s, with work by Trautman, Penrose, Stein, Ehlers, Earman, and others (based on earlier work, such as Cartan's).[12] In Malament (2012, chap. 4), David Malament provides details of the differential geometry formulation of this topic. Galilean (or Newtonian) spacetime is defined as a structure of the form

$$\mathcal{A} = (M, \nabla, h^{ab}, t_{ab}), \tag{2}$$

where M is a manifold diffeomorphic to $\mathbb{R}^4$, ∇ is a flat (torsion-free) affine connection on M, and h^{ab}, t_{ab} are tensor fields on M satisfying compatibility conditions, from which one can construct temporal and spatial metrics and simultaneity surfaces.[13]

The approach we develop here is entirely *synthetic*. The underlying geometric relations are *betweenness* (written $\texttt{Bet}(p, q, r)$), *simultaneity* (written

---

11 To be clear, the synthetic and coordinate structures are isomorphic structures *of the same signature*, say σ. This is because it doesn't make mathematical sense to talk of an isomorphism from A to B unless they are both σ-structures. E.g., it doesn't make sense to say a group $(G, \oplus)$ is isomorphic to a ring $(R, +, \times)$ outside the special case where $\times$ is definable from $+$ or vice versa. Isomorphisms have to "match up" corresponding relations (operations and constants) in the signature. In logic, automated theorem proving, and so on, even seemingly small changes of the signature of the structures in question can make a large difference. For example, the structure $(\mathbb{N}, 0, S, +)$ is decidable (Presburger 1930), but $(\mathbb{N}, 0, S, +, \times)$ is undecidable (Gödel 1931; Tarski 1935). I'm grateful to a referee for mentioning this point, as related ones have arisen in the philosophy of physics.

12 See Trautman (1966), Stein (1967), Penrose (1968), Earman (1970, 1989), Ehlers (1973), Friedman (1983). One may also find mathematically precise descriptions in Arnold (1989, chap. 1) and in Kopczyński and Trautman (1992, 31–32).

13 Here, I am referring to such things as manifolds, diffeomorphisms, affine connections, tangent spaces, tensor fields, and whatnot. An excellent textbook on differential geometry, oriented towards advanced physics students, is Schutz, B.F. (1980). Also, Malament (2012) and Wald (1984). For useful surveys of some of the surrounding philosophical issues, see Huggett and Hoefer (2015) (absolute vs. relational theories of spacetime) and DiSalle (2020) (inertial frames).

doi: 10.48106/dial.v77.i2

$p \sim q$), and *sim-congruence* (written $pq \equiv^{\sim} rs$): these are relations on points. Inertial coordinate systems are then proved to exist by a Representation Theorem. An inertial coordinate system Φ is nothing more than an *isomorphism* from the *synthetic* geometrical structure $(\mathbb{P}, B, \sim, \equiv^{\sim})$ of Galilean spacetime (with carrier set $\mathbb{P}$) to a suitable "coordinate structure" built on the carrier set $\mathbb{R}^4$. Below, we shall call this standard coordinate structure $\mathbb{G}^{(1,3)}$ (Definition 4). So, we shall obtain, by analogy with (1),

$$\Phi : \overbrace{(\mathbb{P}, B, \sim, \equiv^{\sim})}^{\text{synthetic structure}} \to \overbrace{\mathbb{G}^{(1,3)}}^{\text{coordinate structure}}. \tag{3}$$

Euclidean geometry, of course, was also first set out synthetically in Euclid's *Elements*. However, Euclid's *Elements* does not quite meet modern adequate standards of formal rigor. In particular, Moritz Pasch (1882) noted that certain *betweenness* properties of space were merely implicit in Euclid's treatment. Influenced by Pasch and others, the synthetic axiomatization for Euclidean geometry was first made rigorous in Hilbert (1899), which was modified, extended, or simplified in a number of ways, one of which is Veblen (1904) (which extracted the purely betweenness part of Hilbert's system: sometimes called the "axioms of order").

Synthetic axiomatization for *Minkowski spacetime geometry* appeared soon after the classic work of Albert Einstein and Hermann Minkowski (i.e., Einstein 1905; Minkowski 1909) in Alfred Robb's (1911) book. This led to a series of later synthetic developments, including Robb (1936), Ax (1978), Mundy (1986), Goldblatt (1987), Schutz, J.W. (1997), and, most recently, Cocco and Babic (2021). As is now known, Minkowski spacetime can be axiomatized using a single *binary relation*, usually called λ, with $p\lambda q$ meaning "points p and q can be connected by a light signal"—the light-signal relation.[14] As the reader probably knows, this induces a "light cone structure" on the carrier set of points. So, Minkowski spacetime can be defined as a structure $(\mathbb{P}, \lambda)$ satisfying certain axioms, and one may prove that there is an isomorphism $\Phi : (\mathbb{P}, \lambda) \to (\mathbb{R}^4, \lambda_{\mathbb{R}^4})$.[15] Such an isomorphism is called a "Lorentz coordi-

14 In Goldblatt (1987), a relation of "spacetime orthogonality," $pq \perp rs$, is used, but $\perp$ and λ are interdefinable, as Goldblatt shows.

15 Where the standard coordinate relation $\lambda_{\mathbb{R}^4}$ on $\mathbb{R}^4$ is defined as follows: for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^4$, $\mathbf{x}\lambda_{\mathbb{R}^4}\mathbf{y}$ holds iff $\sum_{i=1}^{3}(x^i - y^i)^2 - (x^4 - y^4)^2 = 0$ (i.e., the Minkowski interval is equal to 0). I have set $c = 1$.

nate system.” Then the automorphism group $\mathrm{Aut}((\mathbb{R}^4, \lambda_{\mathbb{R}^4}))$ of $(\mathbb{R}^4, \lambda_{\mathbb{R}^4})$ is the Poincaré group.[16]

Galilean spacetime, however, is the basic spacetime of classical Newtonian (pre-relativistic) physics. In retrospect, it is a kind of “low energy limit” of Minkowski spacetime (when we let the speed of light approach infinity and all the light cones get “squashed” into simultaneity surfaces). But, unlike the case with Minkowski spacetime, the synthetic approach did not appear for a long time. As far as I know, the first brief sketch of a synthetic axiom system for Galilean spacetime appeared in Hartry Field’s *Science Without Numbers* (1980, chap. 6), some 80 years after Hilbert’s classic monograph, *The Foundations of Geometry* (1899), and close on three hundred years after Newton’s *Principia* (1687). Shortly after, John Burgess added further work on this in Burgess (1984) and then again in Burgess and Rosen (1997). Our work here is a descendant of and stimulated by theirs.[17]

The axiom system $\mathrm{Gal}(1, 3)$ we shall arrive at can be written as follows (see table 1 in section 3):

Gal1	BG(4).
Gal2	$\mathrm{EG}(3)^{\sim}$.
Gal3	$\sim$ is an equivalence relation.
Gal4	$\equiv^{\sim} \subseteq [\sim]^4$.
Gal5	$\equiv^{\sim}$ is translation-invariant.

Here, BG(4) is a group of nine axioms, the subsystem of order axioms for betweenness (see appendix A below). And $\mathrm{EG}(3)^{\sim}$ is a group of eleven axioms, a relativized subsystem of axioms for “sim-congruence” and betweenness, obtained from Tarski’s formulation of Euclidean geometry for three dimensions (see appendix A). The three further axioms, Gal3, Gal4, and Gal5, “tie together” these subsystems.[18]

16 In fact, to be a bit more accurate, I believe it is the “*extended*” Poincaré group, allowing global scaling, $x^\mu \mapsto \alpha x^\mu$ ($\alpha \neq 0$), of coordinates. This is because $(\mathbb{P}, \lambda)$ does not have a special “unit length.”

17 Field states his four axioms very briefly, in a footnote (1980, chaps. 6, 54, n.33). Field remarks, “Given the Szczerba-Tarski axiom on ‘Bet’, it is quite trivial to impose requirements on the two new primitives ‘Simul’ and ‘S-Cong’ so as to get the desired representation and uniqueness theorems” (1980, 54). Although Field takes a slightly different congruence relation as primitive (which he calls S-cong), I am reasonably sure that Field’s axiom system is definitionally equivalent to the one given here, $\mathrm{Gal}(1, 3)$. I hope to publish the equivalence proof elsewhere. Burgess’s sketch of the geometry of Galilean spacetime (Burgess 1984; Burgess and Rosen 1997) uses our physical primitives and I believe Burgess must have separately established this equivalence.

18 $[\sim]^4$ is defined to be: $\{(p, q, r, s) \mid p \sim q \wedge p \sim r \wedge p \sim s\}$. See definition 12 below.

doi: 10.48106/dial.v77.i2

To summarize, then, how the rest of this paper goes, we shall use the two separate Representation Theorems for BG(4) and EG(3). The first of these (theorem 62 in appendix B below) asserts the existence of a "global" bijective coordinate system:

$$\Phi : \mathbb{P} \to \mathbb{R}^4, \tag{4}$$

on any (full) model $(\mathbb{P}, B)$ of BG(4), matching any given "*4-frame*" O, X, Y, Z, I and satisfying the betweenness representation condition, for any points $p, q, r \in \mathbb{P}$:[19]

$$B(p, q, r) \leftrightarrow B_{\mathbb{R}^4}(\Phi(p), \Phi(q), \Phi(r)), \tag{5}$$

where $B_{\mathbb{R}^4}$ is the standard betweenness relation on $\mathbb{R}^4$. The second Representation Theorem (theorem 63 in appendix B) asserts the existence of a global coordinate system ψ on any (full) model $(\mathbb{P}, B, \equiv)$ of *three-dimensional Euclidean geometry* EG(3), matching a given "*Euclidean 3-frame*" O, X, Y, Z and satisfying the representation condition for congruence:

$$pq \equiv rs \leftrightarrow \psi(p)\psi(q) \equiv_{\mathbb{R}^3} \psi(r)\psi(s), \tag{6}$$

where $\equiv_{\mathbb{R}^3}$ is the standard congruence relation on $\mathbb{R}^3$. In our system, the axioms EG(3) are *relativized* to simultaneity hypersurfaces, yielding EG(3)$^\sim$. The relativization implements the requirement that each simultaneity hypersurface is a three-dimensional Euclidean space.

We can then combine these two Representation Theorems, applied to any full model $M \vDash_2 \mathrm{Gal}(1,3)$, to obtain the Representation Theorem for Gal(1, 3), which is our main theorem (theorem 55 in section 5). That is, assuming $(\mathbb{P}, B, \sim, \equiv^\sim)$ is a (full) model of Gal(1, 3), the existence of an isomorphism as stated in (3) above:

$$\Phi : \overbrace{(\mathbb{P}, B, \sim, \equiv^\sim)}^{\text{synthetic structure}} \to \overbrace{\mathbb{G}^{(1,3)}}^{\text{coordinate structure}}. \tag{7}$$

The crux of the proof of the main theorem are the Chronology Lemma (lemma 52) and the Congruence Lemma (lemma 54).

19 A 4-frame is an ordered quintuple of points that are not in the same 3-dimensional hypersurface. See definition 58 below.

1 Definitions

Definition 1. The standard Euclidean inner product $\langle ., . \rangle_n$ and norm $\|.\|_n$ on $\mathbb{R}^n$ are defined as follows:[20] For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\langle \mathbf{x}, \mathbf{y} \rangle_n := \sum_{i=1}^{n} x^i y^i$, and $\|\mathbf{x}\|_n := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_n}$. The standard Euclidean metrics $\Delta_n : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ are defined as follows:

$$\Delta_n(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|_n. \tag{8}$$

The *standard Euclidean metric space* with carrier set $\mathbb{R}^n$ is:

$$\mathbb{EG}^n_{\text{metric}} := (\mathbb{R}^n, \Delta_n). \tag{9}$$

Definition 2. The following relations are the *standard betweenness relation* $B_{\mathbb{R}^n}$, *standard simultaneity relation* $\sim_{\mathbb{R}^n}$, *standard congruence relation* $\equiv_{\mathbb{R}^n}$, and *standard sim-congruence relation* $\equiv^{\sim}_{\mathbb{R}^n}$ on $\mathbb{R}^n$. For $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \in \mathbb{R}^n$:

$$\begin{aligned}
B_{\mathbb{R}^n}(\mathbf{x}, \mathbf{y}, \mathbf{z}) &:= (\exists \lambda \in [0,1])(\mathbf{y} - \mathbf{x} = \lambda(\mathbf{z} - \mathbf{x})); && (a)\\
\mathbf{x} \sim_{\mathbb{R}^n} \mathbf{y} &:= x^n = y^n; && (b)\\
\mathbf{xy} \equiv_{\mathbb{R}^n} \mathbf{zu} &:= \Delta_n(\mathbf{x}, \mathbf{y}) = \Delta_n(\mathbf{z}, \mathbf{u}); && (c)\\
\mathbf{xy} \equiv^{\sim}_{\mathbb{R}^n} \mathbf{zu} &:= \Delta_n(\mathbf{x}, \mathbf{y}) = \Delta_n(\mathbf{z}, \mathbf{u}) \,\&\, \mathbf{x} \sim_{\mathbb{R}^n} \mathbf{y} \,\&\, \mathbf{x} \sim_{\mathbb{R}^n} \mathbf{z} \,\&\, \mathbf{x} \sim_{\mathbb{R}^n} \mathbf{u}. && (d)
\end{aligned} \tag{10}$$

For the one-dimensional case, we have two alternative but equivalent definitions. First, $B_{\mathbb{R}}(x, y, z) := (x \leq y \leq z)$; second, $B_{\mathbb{R}}(x, y, z) := |x - y| + |y - z| = |x - z|$.[21]

Definition 3. It will be useful below to define the following special five points in $\mathbb{R}^4$:

$$\mathbf{O} := \begin{pmatrix} 0\\0\\0\\0 \end{pmatrix}, \quad \mathbf{X} := \begin{pmatrix} 1\\0\\0\\0 \end{pmatrix}, \quad \mathbf{Y} := \begin{pmatrix} 0\\1\\0\\0 \end{pmatrix}, \quad \mathbf{Z} := \begin{pmatrix} 0\\0\\1\\0 \end{pmatrix}, \quad \mathbf{I} := \begin{pmatrix} 0\\0\\0\\1 \end{pmatrix}. \tag{11}$$

In other words, the *origin* and the "unit points" on the four axes. I call the ordered tuple $\mathbf{O}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{I}$ the *standard (4-)frame* in $\mathbb{R}^4$.

20 We use the abbreviation $\mathbf{x} = (x^1, \ldots, x^n)$ for n-tuples in $\mathbb{R}^n$. Similarly, for $\mathbf{y}, \mathbf{z}, \ldots$. Hopefully, it will be clear that these don't mean *powers* of x.

21 The second of these, in fact, generalizes to $n > 1$ *if* we have a metric function: $B_{\mathbb{R}^n}(\mathbf{x}, \mathbf{y}, \mathbf{z}) := \Delta_n(\mathbf{x}, \mathbf{y}) + \Delta_n(\mathbf{y}, \mathbf{z}) = \Delta_n(\mathbf{x}, \mathbf{z})$.

doi: 10.48106/dial.v77.i2

Definition 4. The *standard coordinate structures* are:[22]

$\mathbb{BG}^n$	Betweenness geometry in n dimensions over $\mathbb{R}$	$:= (\mathbb{R}^n, B_{\mathbb{R}^n})$.
$\mathbb{EG}^n$	Euclidean space in n dimensions over $\mathbb{R}$	$:= (\mathbb{R}^n, B_{\mathbb{R}^n}, \equiv_{\mathbb{R}^n})$.
$\mathbb{G}^{(1,n)}$	Galilean spacetime in $n+1$ dimensions over $\mathbb{R}$	$:= (\mathbb{R}^{n+1}, B_{\mathbb{R}^{n+1}}, \sim_{\mathbb{R}^{n+1}}, \equiv^{\sim}_{\mathbb{R}^{n+1}})$.

Our central interest is $\mathbb{G}^{(1,3)}$, the *standard coordinate structure for four-dimensional Galilean spacetime*. The carrier set of $\mathbb{G}^{(1,3)}$ is $\mathbb{R}^4$. Its distinguished relations are betweenness (10, a), simultaneity (10, b), and sim-congruence (10, d) on $\mathbb{R}^4$. Note that $\mathbb{G}^{(1,3)}$ does *not* carry a metric or distance function.

2 Derivation of (Extended) Galilean Transformations

What is the *symmetry group* of the standard coordinate structure $\mathbb{G}^{(1,3)}$ for Galilean spacetime? We will see that its symmetry group is a certain Lie group $\mathcal{G}^e(1,3)$, a 12-dimensional Lie group that extends the usual Galilean group $\mathcal{G}(1,3)$ by two additional parameters, which determine coordinate scalings.

Definition 5. A is an element of the *extended Galilean matrix group* $\mathsf{Mat}^e_{\mathsf{Gal}}(4)$ if and only if A is a 4×4 matrix with real entries and has the (block matrix) form

$$A = \begin{pmatrix} \alpha_1 R & \vec{v} \\ 0 & \alpha_2 \end{pmatrix}, \tag{12}$$

where

$$R = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} \tag{13}$$

is in $O(3)$, $\vec{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$, and $\alpha_1, \alpha_2 \in \mathbb{R} - \{0\}$. The $O(3)$ matrix R is called the *rotation* of A, the 3-vector $\vec{v}$ is called the *(relative) velocity* of A, the constant α_1 is called the *spatial scaling factor* of A, and the constant α_2 is the *temporal scaling factor* of A.

Lemma 6. $\mathsf{Mat}^e_{\mathsf{Gal}}(4)$ *is a subgroup of* $GL(4)$*.*

22 Regarding the definitions of $\mathbb{BG}^n$, $\mathbb{EG}^n$ and $\mathbb{G}^{(1,n)}$. These still make sense if we replace $\mathbb{R}$ in the definition by a Euclidean ordered field F (an ordered field where all non-negative elements are squares). Cf. Szczerba and Tarski (1979, 160, Definition 1.5), who call a space $\mathbb{BG}^n(F)$ a "Cartesian affine space" over F.

Proof. This is a routine verification. The main part is to check that $\mathsf{Mat}^{e}_{\mathrm{Gal}}(4)$ is closed under matrix multiplication and each element in $\mathsf{Mat}^{e}_{\mathrm{Gal}}(4)$ has an inverse in $\mathsf{Mat}^{e}_{\mathrm{Gal}}(4)$.

$\square$

Definition 7. Let $h : \mathbb{R}^4 \to \mathbb{R}^4$. We say that h is an *extended Galilean transformation* just if there exists an extended Galilean matrix A and a displacement $\mathbf{d} \in \mathbb{R}^4$ such that, for all $\mathbf{x} \in \mathbb{R}^4$,

$$h(\mathbf{x}) = A\mathbf{x} + \mathbf{d}. \tag{14}$$

Lemma 8. *The set of extended Galilean transformations forms a group.*

Proof. This is a detailed verification of the group properties, analogous to the above.

$\square$

Definition 9. $\mathcal{G}^e(1,3) :=$ the group of extended Galilean transformations.

Theorem 10 (Automorphisms of $\mathbb{G}^{(1,3)}$)**.** *Aut*$(\mathbb{G}^{(1,3)}) = \mathcal{G}^e(1,3)$.

Proof. I give a sketch of the proof. To show $\mathcal{G}^e(1,3) \subseteq \mathrm{Aut}(\mathbb{G}^{(1,3)})$, we verify that each extended Galilean transformation is a symmetry of $\mathbb{G}^{(1,3)}$. Since $\mathbb{BG}^4$ is a reduct of $\mathbb{G}^{(1,3)}$, and each extended Galilean transformation is affine, it follows that betweenness is invariant. The special form of extended Galilean matrices then ensures that simultaneity and sim-congruence are invariant.

To show that $\mathrm{Aut}(\mathbb{G}^{(1,3)}) \subseteq \mathcal{G}^e(1,3)$ is more involved. Since $\mathbb{BG}^4$ is a reduct of $\mathbb{G}^{(1,3)}$, it follows that any symmetry h of $\mathbb{G}^{(1,3)}$ must be affine, and so there exists a $GL(4)$ matrix A and displacement $\mathbf{d} \in \mathbb{R}^4$ such that, for any $\mathbf{x} \in \mathbb{R}^4$,

$$h(\mathbf{x}) = A\mathbf{x} + \mathbf{d}. \tag{15}$$

To determine the sixteen components A_{ij} of A, one must then examine the conditions that simultaneity and sim-congruence be invariant. By examining certain choices of points, the invariance of simultaneity enforces that A must have the form

$$A = \begin{pmatrix} C & \vec{v} \\ 0 & \alpha_2 \end{pmatrix}, \tag{16}$$

where C is a 3×3 matrix, and α_2 is a non-zero constant. The invariance of sim-congruence enforces that the upper 3×3 block C must be a multiple $\alpha_1 R$

doi: 10.48106/dial.v77.i2

of an $O(3)$ matrix R by a non-zero real factor α_1:

$$A = \begin{pmatrix} \alpha_1 R & \vec{v} \\ 0 & \alpha_2 \end{pmatrix}. \tag{17}$$

But this is an extended Galilean matrix. Consequently, $\mathrm{Aut}(\mathbb{G}^{(1,3)}) \subseteq \mathcal{G}^e(1,3)$.

Together, these results imply that $\mathrm{Aut}(\mathbb{G}^{(1,3)}) = \mathcal{G}^e(1,3)$.

□

The constants α_1, α_2 in any *extended* Galilean matrix A determine *scalings* of the spatial and temporal coordinates, respectively. So, given some A in the extended Galilean matrix group and any $(\vec{x}, t) \in \mathbb{R}^4$,

$$A(\vec{x}, t) = (\alpha_1 R\vec{x} + \vec{v}t, \alpha_2 t). \tag{18}$$

Let's set the relative rotation R to be $\mathbb{I}$ and set the relative velocity $\vec{v}$ to be zero:

$$A(\vec{x}, t) = (\alpha_1 \vec{x}, \alpha_2 t). \tag{19}$$

Thus, the spatial coordinates are scaled by α_1, and the temporal coordinate is scaled by α_2. Instead, let us set these scalings α_1, α_2 at 1 and consider the image $(\vec{x}', t')$ of the point with coordinates $(\vec{x}, t)$ under an extended Galilean transformation:

$$\vec{x}' = R\vec{x} + \vec{v}t + \vec{d}, \tag{20}$$

$$t' = t + d_t. \tag{21}$$

These are the usual Galilean transformations as given in physics textbooks, in usually simplified form (e.g., Sears, Zemansky and Young 1979, 252; Longair 1984, 87; or Rindler 1977, 3). The conventional Galilean group $\mathcal{G}(1,3)$ is normally understood to be this 10-parameter Lie group: the ten parameters are these: four parameters for the spatial and temporal translations, **d**; three parameters (i.e., determined by the three Euler angles) for the rotation matrix R; three parameters for the velocity $\vec{v}$.

As we have defined it, the *extended* Galilean group $\mathcal{G}^e(1,3)$ is a 12-parameter Lie group: the two additional parameters, α_1, α_2, permit coordinate scalings. These two extra degrees of freedom are a consequence of our synthetic treatment, and this is completely analogous to Euclidean betweenness and congruence being invariant under coordinate scaling. Indeed, α_1 and α_2 are gauge parameters in the oldest sense of the word.

3 Axiomatization of Galilean Spacetime: $\mathsf{Gal}(1,3)$

To begin, we state the informal physical meanings of our three primitive symbols:[23]

> Betweenness Predicate: $\texttt{Bet}$. $\texttt{Bet}(p,q,r)$ means that q lies on a straight line inclusively between p and r (allowing the cases $q = p$ and $q = r$).

> Simultaneity Predicate: $\sim$. $p \sim q$ means that the points p, q are simultaneous.

> Sim-Congruence Predicate: $\equiv^{\sim}$. $pq \equiv^{\sim} rs$ means the points p, q, r, s are simultaneous, and the length of the segment pq is equal to the length of the segment rs.

We are now ready to state the (synthetic) axioms for Galilean spacetime.

Definition 11. The theory $\mathsf{Gal}(1,3)$ is a two-sorted theory with sorts $\{\texttt{point}, \texttt{pointset}\}$ and variables $\mathsf{Var}_{\texttt{point}} = \{p_1, p_2, \dots\}$ and $\mathsf{Var}_{\texttt{pointset}} = \{\mathsf{X}_1, \mathsf{X}_2, \dots\}$. The signatures σ_{Gal} and $\sigma_{\mathsf{Gal},\in}$ are given by $\sigma_{\mathsf{Gal}} = \{\texttt{Bet}, \sim, \equiv^{\sim}\}$ and $\sigma_{\mathsf{Gal},\in} = \{\texttt{Bet}, \sim, \equiv^{\sim}, \in\}$. By $L(\sigma_{\mathsf{Gal}})$, I shall mean the first-order language with restricted signature σ_{Gal} over the single sort $\texttt{point}$. Its atomic formulas are of the four forms: $p_1 = p_2$, $\texttt{Bet}(p_1, p_2, p_3)$, $p_1 \sim p_2$, and $p_1 p_2 \equiv^{\sim} p_3 p_4$, where "$p_i$" are point variables, and the remaining formulas are built up using the connectives $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$, and quantifiers $\forall$ and $\exists$, as per the usual recursive definition of "formula of $L(\sigma)$."[24] By $L(\sigma_{\mathsf{Gal},\in})$, I mean the "monadic second-order" language, with signature $\sigma_{\mathsf{Gal},\in}$. Its atomic sentences include those above along with formulas: $p_i \in \mathsf{X}_j$ and $\mathsf{X}_i = \mathsf{X}_j$. (A parser for this language counts the strings $p_i = \mathsf{X}_j$, $\mathsf{X}_j = p_i$, and $\mathsf{X}_i \in p_j$ and $p_i \in p_j$ as ill-formed.) The remaining formulas are built up using the connectives $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$, and quantifiers $\forall$ and $\exists$, including the new quantifications $\forall \mathsf{X}_i\, \varphi$ and $\exists \mathsf{X}_i\, \varphi$.

In discussing a full model M of, say, $\mathsf{BG}(4)$, I shall generally write "$M \vDash_2 \mathsf{BG}(4)$" to make it clear that M is a *full* model of $\mathsf{BG}(4)$. In other words, if

23 Cf. the "interpretive principles" given in Malament (2012, 120–121).

24 Informally, we liberalize notation for point variables, occasionally using "p," "q," "r," "s," "u," "x," "y," "z," and the like, with natural number subscripts.

doi: 10.48106/dial.v77.i2

$M = (\mathbb{P}, \dots)$, then $M \vDash_2 \forall \mathsf{X}_i\, \varphi(\mathsf{X}_i)$ if and only if, for *every* subset $U \subseteq \mathbb{P}$, $\varphi[U]$ is true in M.

Definition 12. $(p, q, r, s) \in [\sim]^4$ iff $p \sim q, p \sim r, p \sim s$.

Definition 13. The (non-logical) axioms of Gal$(1, 3)$ are as follows:

Table 1: The axiom system Gal$(1, 3)$.

Gal1	BG(4).
Gal2	$\mathrm{EG}(3)^{\sim}$.
Gal3	$\sim$ is an equivalence relation.
Gal4	$\equiv^{\sim} \subseteq [\sim]^4$.
Gal5	$\equiv^{\sim}$ is translation-invariant.

BG(4) is really an axiom group of nine axioms for `Bet`.[25] These are given in definition 56 in appendix A. But, to simplify the description here, one may take their conjunction.[26] EG(3) is also an axiom group, this time of eleven axioms. These are given in definition 57 in appendix A. The axiom $\mathrm{EG}(3)^{\sim}$ listed above requires further explanation.[27]

25 I use the moniker "BG" to mean "*betweenness geometry*" (n dimensions) for several reasons. First, because there doesn't seem to be a standard name for these geometries. Second, they are sometimes called "affine geometries," but the word "affine" has too many meanings, including two different meanings, each having nothing to do with the betweenness relation. These are "*affine plane*" (see, e.g., Bennett 1995) and "*affine space*" (see, e.g., Gallier 2011). Sometimes, the terminology "ordered geometry" is used (Pambuccian 2011). But "OG" seems to me ugly. Since the terminology is not entirely uniform, I use "betweenness geometry" and, hence, BG(4), etc. I should note that these axiom systems contain Euclid's Parallel Postulate in some form.

26 The system BG(4) corresponds precisely to what Burgess called GEOM_4 in Burgess (1984). The system BG(4) also corresponds to what Szczerba and Tarski called $\mathrm{GA}_4^* + \mathrm{Euclid}$ in Szczerba and Tarski (1979, 1965). The term "GA" is used to mean a system of *absolute* or *neutral geometry* (i.e., without the Parallel Postulate), which is why (Euclid) is added. Note that (Euclid) is formulated entirely using `Bet`, and congruence does not appear. The subscript denotes the dimension, and the asterisk denotes that the axiom system is second-order; this means the *Continuity Axiom* is second-order rather than a scheme. A system essentially equivalent to GA_3^* is studied carefully in the monograph Borsuk and Szmielew (1960). The axioms of BG(4) are the result of simplifying the categorical system of "order axioms" given in Veblen (1904), where the relevant categoricity or representation theorem (i.e., our theorem 62 in appendix B) was first given.

27 EG(3) itself corresponds to the *second-order* version of the *three-dimensional* version of Tarski's system for synthetic Euclidean geometry in Tarski (1959), somewhat simplified in Tarski and Givant (1999). As with "BG," I use the moniker "EG" to mean "*Euclidean geometry*." In my notation, Tarski's 1959 paper is mostly about the first-order theory $\mathrm{EG}_0(2)$, which is EG(2) "little's brother."

This construction is sketched, very briefly, in Field (1980, 54, n.33). First, one replaces $\equiv$ by $\equiv^{\sim}$ in each $\mathsf{EG}(3)$ axiom. Next, one *relativizes* each axiom to the formula $p \sim z$ (treating z as a parameter) so that the resulting axiom says that it holds for all points simultaneous with z.[28] Next, one prefixes the result with $\forall z$ and then takes the conjunction of the axioms. For example, under relativization, the $\equiv$-Transitivity axiom (E3) and the Pasch axiom (E6) become:

$\equiv^{\sim}$-Transitivity	$\forall z[(\forall p, q, r, s, t, u \sim z)(pq \equiv^{\sim} rs \wedge pq \equiv^{\sim} tu \rightarrow rs \equiv^{\sim} tu)]$.
Pasch	$\forall z[(\forall p, q, r, s, u \sim z)(\mathtt{Bet}(p, q, r) \wedge \mathtt{Bet}(s, u, q) \rightarrow (\exists x \sim z)(\mathtt{Bet}(r, x, s) \wedge \mathtt{Bet}(p, u, x)))]$.

In addition to the given non-logical axioms, we also have the customary axioms for second-order logic (table 3):

Table 3: Axioms for second-order logic.

Comprehension	$\exists \mathsf{X}_1 \, \forall p \, (p \in \mathsf{X}_1 \leftrightarrow \varphi)$	(variable X_1 not free in φ).
Extensionality	$\forall \mathsf{X}_1 \, \forall \mathsf{X}_2 \, (\forall p \, (p \in \mathsf{X}_1 \leftrightarrow p \in \mathsf{X}_2) \rightarrow \mathsf{X}_1 = \mathsf{X}_2)$.	

I shall, in effect, however, assume an *ambient set theory*.[29] The reason is that I am not concerned with narrow proof-theoretic matters concerning the whole theory (for example, completeness) but rather with establishing some facts about the *full* models of the theory $\mathsf{Gal}(1,3)$. Since we consider just full models, Comprehension and Extensionality are satisfied more or less by fiat.[30] This is completely analogous to our approach in giving the usual proof, essentially that of Dedekind (1888), of the categoricity of second-order

28 The relativization is more precisely defined as a translation $^{\circ}$, which acts as the identity on atomic formulas, which commutes with the Boolean logical connectives, and, for quantifiers, maps $\forall p \, \varphi$ to $(\forall p \sim z) \, \varphi^{\circ}$, maps $\exists p \, \varphi$ to $(\exists p \sim z) \, \varphi^{\circ}$, maps $\forall \mathsf{X} \, \varphi$ to $(\forall \mathsf{X} \subseteq \Sigma_z) \, \varphi^{\circ}$, and maps $\exists \mathsf{X} \, \varphi$ to $(\exists \mathsf{X} \subseteq \Sigma_z) \, \varphi^{\circ}$.

29 See also Borsuk and Szmielew (1960, 7–8) on this topic.

30 A suitable "ambient set theory," a system of axioms for the existence of sets, where the points will now be *urelements* or *atoms* (i.e., not sets or classes), and where comprehension, separation, and replacement schemes can be applied to any urelement predicate (e.g., `Bet` and so on), is given in Ketland (2021). The ambient set theory is called $\mathsf{ZFU}_{V(T)}$ in Field (1980, 17).

doi: 10.48106/dial.v77.i2

arithmetic PA_2, although, as a matter of fact, the categoricity of PA_2 can be "internalized" as a proof *inside* PA_2 itself (see Simpson and Yokoyama 2013).

The three Galilean axioms Gal3, Gal4, and Gal5 are the glue that holds together the betweenness axioms BG(4) and the Euclidean axioms $EG(3)^{\sim}$. The content of Gal3 and Gal4 seems evident. The final axiom Gal5 is the sole axiom that needs some further explanation.[31] This axiom expresses the *translation invariance* of the $\equiv^{\sim}$ relation and may be expressed using vector notation as follows:

$$pq \equiv^{\sim} rs \rightarrow (p+\mathbf{v})(q+\mathbf{v}) \equiv^{\sim} (r+\mathbf{v})(s+\mathbf{v}). \tag{22}$$

In other words, if the (simultaneous) segments pq and rs have the same length, then the (simultaneous) segments $(p+\mathbf{v})(q+\mathbf{v})$ and $(r+\mathbf{v})(s+\mathbf{v})$ have the same length for any vector $\mathbf{v}$.[32]

An equivalent axiom can be expressed solely using the primitives $\mathtt{Bet}$, $\sim$, and $\equiv^{\sim}$ and quantifying over points. Roughly, the axiom Gal5 is equivalent to the following rather long-winded claim:

> If p, q, r, s, and p', q', r', s' are points such that the vectors $\mathbf{v}_{p,p'}$, $\mathbf{v}_{q,q'}$, $\mathbf{v}_{r,r'}$, $\mathbf{v}_{s,s'}$ are all equal and $pq \equiv^{\sim} rs$, then $p'q' \equiv^{\sim} r's'$.

Note that the equality clause "$\mathbf{v}_{p,p'} = \mathbf{v}_{q,q'}$" means "$p, q, p', q'$ is a *parallelogram*," and the 4-place predicate "p_1, p_2, p_3, p_4 is a parallelogram" can be defined using $\mathtt{Bet}$ (see definition 15).

The second-order theories BG(4) and EG(3), with their point *set* variables, contain the second-order Continuity Axiom (Tarski 1959, 18):

$$\begin{aligned} &[\exists r\, (\forall p \in X_1)\, (\forall q \in X_2) \mathtt{Bet}(r, p, q)] \\ &\rightarrow [\exists s\, (\forall p \in X_1)\, (\forall q \in X_2)\, \mathtt{Bet}(p, s, q)]. \end{aligned} \tag{23}$$

This geometrical continuity axiom, it may be noted, is closely analogous to the "Dedekind Cut Axiom," which may be used as an axiom in the formaliza-

31 The axiom Gal5 is so obvious that it occurred to me that it might indeed be provable from the remainder. However, I've not found a proof of this. So, I retain it. It is needed to show that the vector translation of a Galilean 4-frame is also a Galilean 4-frame (lemma 53 below).

32 The fact that if the points p, q, r, s are simultaneous, then the points $p+\mathbf{v}$, $q+\mathbf{v}$, $r+\mathbf{v}$, and $s+\mathbf{v}$ are also simultaneous is given in lemma 45 below.

tion of the second-order theory ALG of real numbers:[33]

$$
\begin{aligned}
&(\forall \mathrm{X}_1 \subseteq \mathbb{R})\,(\forall \mathrm{X}_2 \subseteq \mathbb{R})\,(\mathrm{X}_1 \neq \varnothing \wedge \mathrm{X}_2 \neq \varnothing \wedge \overbrace{(\forall x \in \mathrm{X}_1)\,(\forall y \in \mathrm{X}_2)\,(x \leq y)}^{\mathrm{X}_1 \text{ “precedes” } \mathrm{X}_2} \\
&\qquad \rightarrow \exists s\, \overbrace{(\forall x \in \mathrm{X}_1)\,(\forall y \in \mathrm{X}_2)\,(x \leq s \wedge s \leq y)}^{\text{the point } s \text{ “cuts” } \mathrm{X}_1 \text{ and } \mathrm{X}_2}).
\end{aligned}
\tag{24}
$$

The second-order theories BG(4) and EG(3) are, foundationally speaking, strong, and both interpret ALG. They have *first-order* versions—their "little brothers," so to speak, which I shall call $\mathrm{BG}_0(4)$ and $\mathrm{EG}_0(3)$—obtained by replacing the single Continuity Axiom by infinitely many instances of the Continuity *axiom scheme*: in these instances, there are only point variables.

The little brothers, $\mathrm{BG}_0(4)$ and $\mathrm{EG}_0(3)$, are meta-mathematically somewhat different from their big brothers. In particular, they are, in fact, *complete* (and, since they are recursively axiomatized, *decidable*), as established by a celebrated theorem of Alfred Tarski (1951). But the big brothers are *incomplete* because they interpret Peano arithmetic (PA), and then Gödel's incompleteness results apply. This observation leads to an important difficulty faced by Field's nominalism:

Remark 14. The *second-order* nature of BG(4)—i.e., its point variables range over points, and its set variables range over *sets of points*—is what lies at the root of the technical problem for Hartry Field's nominalist program (1980) highlighted, first informally by John Burgess, Saul Kripke, and Yiannis Moschovakis, and then, in detail, by Stewart Shapiro in (1983), and also mentioned in Burgess (1984, last section). The required *representation theorem* indeed holds for BG(4) with respect to full models (and from this, the other representation theorems can be built up, just as we do below). This is theorem 62 below. But, unfortunately, adding additional *set theory* axioms to BG(4) is *non-conservative*. This is because BG(4) interprets Peano arithmetic. And then, by Gödel's incompleteness results (Gödel 1931; Raatikainen 2015), there is a consistency sentence Con(BG(4)) in the language of BG(4)

33 I follow Burgess (1984) in calling this theory ALG. A standard axiomatization of ALG is given in Apostol (1967, 18, 20, 25). An equivalent axiomatization appears in Rudolf Carnap's neglected textbook Carnap (1958, paras. 45, 183–185). See also Tarski (1946, 215) for a similar and equivalent formulation to (24), but Tarski uses the notion "the set X strictly precedes the set Y" (using $<$ instead of $\leq$) and "s separates the sets X and Y" (again using $<$ instead of $\leq$). But these Continuity axioms are equivalent. And both are equivalent to the usual Dedekind cut axiom given in an analysis textbook: "any non-empty bounded subset of $\mathbb{R}$ has a supremum" (e.g., Apostol 1967, 25).

doi: 10.48106/dial.v77.i2

itself such that BG(4) does not prove Con(BG(4)). Con(BG(4)) is indeed true in the standard coordinate structure since BG(4) *is* consistent (for it has a model). This sentence becomes *provable* when further set axioms are added. On the other hand, BG(4) has a little brother, $BG_0(4)$, which is a *first-order* theory (we replace the Continuity Axiom by infinitely many instances of the Continuity axiom scheme). Then, conservativeness holds for $BG_0(4)$ because it is *complete*! (As we know from the aforementioned celebrated result by Tarski 1951.) But now the required representation theorem does *not* hold for the little brother $BG_0(4)$. Instead, a rather different representation theorem holds, replacing $\mathbb{R}^n$ by F^n for "some real-closed field F." This is a revision of theoretical physics, for physics works with a *manifold*, a point set equipped with a system of charts, which are maps into $\mathbb{R}^n$. Field's program required both conservativeness (to vindicate the claimed "instrumentalist nature" of mathematics) and representation (to vindicate the claimed "purely representational" feature of applied mathematics). But the technical snag is that we cannot have *both* conservativeness and the representation theorem.

4 Main Results About Gal(1, 3)

4.1 *Definitions: Betweenness Geometry*

Definition 15. The formula $\mathtt{Bet}(p,q,r) \vee \mathtt{Bet}(q,r,p) \vee \mathtt{Bet}(r,p,q)$ expresses that points p, q, r are *collinear*. Assuming $p \neq q$, we use $\ell(p,q)$ to mean the set of points collinear with p and q, i.e., the line through p, q. It can be proved in BG(4) that each line is determined by exactly two points. We may express notions of *coplanarity*, *cohyperplanarity*, and so on through all positive integer dimensions using formulas that I write as $\mathtt{co}_n(p_1, \dots, p_{n+2})$.[34] So, $\mathtt{co}_1(p,q,r)$ means that p, q, r are collinear; $\mathtt{co}_2(p,q,r,s)$ means that p, q, r, s are coplanar; and so on through higher dimensions. Lines $\ell(p,q)$ and $\ell(r,s)$ are *parallel* if and only if $\mathtt{co}_2(p,q,r,s)$ and either $\ell(p,q) = \ell(r,s)$, or $\ell(p,q)$ and $\ell(r,s)$ do not intersect (i.e., have no point in common). For this, we write $\ell(p,q) \parallel \ell(r,s)$. Four distinct points p, q, r, s form a *parallelogram* just if $\ell(p,q) \parallel \ell(r,s)$ and $\ell(p,s) \parallel \ell(q,r)$ (see Bennett 1995, 49). The notion of what I call a *4-frame* is given below (definition 58, in appendix B): an ordered quintuple O, X, Y, Z, I that do not lie in the same 3-dimensional space.

34 The precise definitions of the predicates $\mathtt{co}_n$ are given in Szczerba and Tarski (1979, 190). (Szczerba and Tarski call these predicates $\mathbf{L}_n$.) The definition is recursive: for $n > 1$, each $\mathtt{co}_n$ is defined in terms of the previous ones. These definitions are due to Kordos (1969).

The theory BG(4) proves the existence of a 4-frame: this is simply the Lower Dimension Axiom (the axioms are listed in appendix A). It can be proved in BG(4) that, given a line ℓ and a point p, there is a unique line ℓ' parallel to ℓ and containing p (this is called Playfair's Axiom and is an equivalent of Euclid's Parallel Postulate). From Playfair's Axiom, it can be proved in BG(4) that $\parallel$ is an equivalence relation. A number of other theorems from plane and solid geometry can be established, including Desargues's Theorem and Pappus's Theorem. See Bennett (1995) for an explanation of these theorems. It can be proved that there is a bijection between any pair of lines. The claims mentioned so far are sufficient (the assumptions required include Desargues's Theorem and Pappus's Theorem) to establish that, given distinct parameters p, q, the line $\ell(p, q)$ is isomorphic to an ordered field.[35] The Continuity Axiom of BG(4) then ensures that this field is order-complete. From this, we conclude that there is a (unique) isomorphism $\varphi_{p,q} : \ell(p, q) \to \mathbb{R}$, i.e., $\varphi_{p,q}(p) = 0$ and $\varphi_{p,q}(q) = 1$. See also the proof sketch for theorem 62 below.

4.2 *Definitions: Galilean Geometry*

Turning to the system Gal(1, 3), we need separate definitions of notions pertaining to simultaneity ($\sim$) and sim-congruence ($\equiv^{\sim}$).

Definition 16. A *time axis* T is a line $\ell(p, q)$, where $p \nsim q$.

Definition 17. A *simultaneity hypersurface* Σ_p is the set $\{q \mid q \sim p\}$ of points simultaneous with p.

Beyond the notion of a 4-frame, we need a few more specialized notions of "frame" for Galilean spacetime.

Definition 18 (sim 4-frame)**.** A *sim 4-frame* is a sequence of five points O, X, Y, Z, I such that O, X, Y, Z are simultaneous and not coplanar, and I is not simultaneous with O. A sim 4-frame is automatically a 4-frame.

Definition 19 (Euclidean sim 3-frame)**.** A *Euclidean sim 3-frame* is a sequence of four points O, X, Y, Z that are simultaneous, are not co_2, and OX, OY, OZ have the same length and are mutually perpendicular. That

35 The required definitions of geometrical addition $+$ and geometrical multiplication $\times$ (which go back to Hilbert 1899) are given in Bennett (1995). The definition of the order on a fixed line in terms of `Bet` is given in Tarski (1959, proof of theorem 1).

is, $OX \equiv^{\sim} OY$, $OX \equiv^{\sim} OZ$, and $OY \equiv^{\sim} OZ$; and $OX \perp^{\sim} OY$, $OX \perp^{\sim} OZ$, and $OY \perp^{\sim} OZ$.[36]

Definition 20 (Galilean 4-frame)**.** A *Galilean 4-frame* is a sequence of five points O, X, Y, Z, I that are a sim 4-frame and such that the four points O, X, Y, Z are a Euclidean sim 3-frame. Note that $O \nsim I$, and then the line $\ell(O, I)$ is called the *time axis* of the Galilean 4-frame. A Galilean 4-frame is automatically a 4-frame. We shall simply call it a Galilean frame.

4.3 *Soundness*

It is straightforward to demonstrate that Gal(1, 3) is true in the coordinate structure $\mathbb{G}^{(1,3)}$ by verifying that each axiom of Gal(1, 3) is true in $\mathbb{G}^{(1,3)}$.

Lemma 21 (Soundness Lemma)**.** $\mathbb{G}^{(1,3)} \vDash_2 \mathrm{Gal}(1,3)$.

4.4 *Lemmas*

Lemma 22. *Given a point p and a time axis T, there is a unique line $\ell' \parallel T$ st $p \in \ell'$. (This is Playfair's Axiom, a theorem of* BG(4)*, and an equivalent of Euclid's parallel postulate.)*

Lemma 23. *Any five simultaneous points are* co_3 *(i.e., cohyperplanar$_3$).*

Proof. This follows from the Upper Dimension Axiom in $\mathrm{EG}(3)^{\sim}$. This asserts that, for a fixed simultaneity hypersurface Σ_z, any five points in Σ_z are co_3. Hence, any five simultaneous points are co_3.

□

Lemma 24 (Non-Triviality)**.** *There are at least two non-simultaneous points.*

Proof. By the Lower Dimension axiom in BG(4), there is a 4-frame of five points, O, X, Y, Z, I, which are not co_3. By lemma 23, any five simultaneous points are co_3. If $O \sim X \sim Y \sim Z \sim I$, they'd be co_3, a contradiction. So, there are at least two non-simultaneous points.

□

36 Perpendicularity $OX \perp^{\sim} OY$, for three distinct simultaneous points O, X, Y, is defined just as in definition 59 in appendix B, but replacing the ordinary congruence predicate $\equiv$ by the sim-congruence predicate $\equiv^{\sim}$.

Lemma 25 (Galilean Frame Lemma)**.** *There is a Galilean frame O, X, Y, Z, I.*

Proof. By lemma 24, let O, I be two non-simultaneous points. By $\mathsf{EG}(3)^{\sim}$, Euclidean three-dimensional geometry holds on simultaneity hypersurface Σ_O. So, there exists O, X, Y, Z, a Euclidean sim 3-frame in Σ_O. Since O and I are not simultaneous, O, X, Y, Z, I form a Galilean frame (whose time axis is $\ell(O, I)$).

□

4.5 *Vector Methods*

In the first part of this section, we first assume that we are considering a full model $M \vDash_2 \mathsf{BG}(4)$, with $M = (\mathbb{P}, B)$ (i.e., $B \subseteq \mathbb{P}^3$ is the interpretation in M of the predicate `Bet`). And then, we further assume we are considering a full model $M \vDash_2 \mathsf{Gal}(1,3)$, with $M = (\mathbb{P}, B, \sim, \equiv^{\sim})$. We assume the material in appendix D, which introduces the new sorts: *reals* and *vectors*.[37] The vector displacement from p to q is written: $\mathbf{v}_{p,q}$.[38] In particular, recall that, by theorem 68, the vector space $\mathbb{V}$ of displacements is isomorphic to $\mathbb{R}^4$ (as a vector space).[39]

Since $M \vDash_2 \mathsf{BG}(4)$, we know, by theorem 62, that there exists a coordinate system $\Phi : \mathbb{P} \to \mathbb{R}^4$ on M, i.e., an isomorphism $\Phi : (\mathbb{P}, B) \to (\mathbb{R}^4, B_{\mathbb{R}^4})$.

Definition 26. Let O, X, Y, Z, I be a 4-frame in M. Define the four vectors:

$$\mathbf{e}_1 := \mathbf{v}_{O,X}; \quad \mathbf{e}_2 := \mathbf{v}_{O,Y}; \quad \mathbf{e}_3 := \mathbf{v}_{O,Z}; \quad \mathbf{e}_4 := \mathbf{v}_{O,I}. \tag{25}$$

Lemma 27. *$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ is a basis for $\mathbb{V}$.*

This is established inside the detailed proof of theorem 68 below.

Definition 28. Given a coordinate system Φ on M, we can define the *associated 4-frame*, O, X, Y, Z, I of points in M:

$$\begin{gathered} O := \Phi^{-1}(\mathbf{O}), \quad X := \Phi^{-1}(\mathbf{X}), \quad Y := \Phi^{-1}(\mathbf{Y}), \\ Z := \Phi^{-1}(\mathbf{Z}), \quad I := \Phi^{-1}(\mathbf{I}). \end{gathered} \tag{26}$$

37 See also Malament (2009) for a nice exposition of these ideas.

38 Some geometry texts (e.g., Coxeter 1969, 213) will write $\overrightarrow{pq}$. E.g., Chasles's Relation then becomes $\overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}$.

39 I am grateful to a referee for bringing to my attention Saunders (2013), whose discussion of Galilean spacetime uses similar vector methods and the notion of affine space.

doi: 10.48106/dial.v77.i2

Definition 29. Given a coordinate system Φ, we define four basis vectors:

$$\mathbf{e}_1^\Phi := \mathbf{v}_{O,X}; \quad \mathbf{e}_2^\Phi := \mathbf{v}_{O,Y}; \quad \mathbf{e}_3^\Phi := \mathbf{v}_{O,Z}; \quad \mathbf{e}_4^\Phi := \mathbf{v}_{O,I}. \tag{27}$$

Lemma 30. $\{\mathbf{e}_1^\Phi, \mathbf{e}_2^\Phi, \mathbf{e}_3^\Phi, \mathbf{e}_4^\Phi\}$ *is a basis for* $\mathbb{V}$.

This is a corollary of lemma 27.

Given a coordinate system Φ and a point p, the four components of $\Phi(p)$ are written as follows:

$$\Phi(p) = \begin{pmatrix} \Phi^1(p) \\ \Phi^2(p) \\ \Phi^3(p) \\ \Phi^4(p) \end{pmatrix}. \tag{28}$$

Lemma 31. *For any point p, we have:*

$$\mathbf{v}_{O,p} = \sum_{a=1}^{4} \Phi^a(p)\, \mathbf{e}_a^\Phi. \tag{29}$$

Proof. Consider some of the details of the proof of the Representation Theorem for BG(4) (see Theorem 62 below). Examining the vector $\mathbf{v}_{O,p}$ from the origin O to p, one can see that:

$$\mathbf{v}_{O,p} = \mathbf{v}_{O,p_X} + \mathbf{v}_{O,p_Y} + \mathbf{v}_{O,p_Z} + \mathbf{v}_{O,p_I}, \tag{30}$$

where p_X, p_Y, p_Z, and p_I are the "ordinates" on the four axes. Note first that $\mathbf{v}_{O,p_X} = \varphi_{O,X}(p_X)\mathbf{v}_{O,X} = \varphi_{O,X}(p_X)\mathbf{e}_1^\Phi$, and similarly for the other three vectors. So:

$$\mathbf{v}_{O,p} = \varphi_{O,X}(p_X)\mathbf{e}_1^\Phi + \varphi_{O,Y}(p_Y)\mathbf{e}_2^\Phi + \varphi_{O,Z}(p_Z)\mathbf{e}_3^\Phi + \varphi_{O,I}(p_I)\mathbf{e}_4^\Phi. \tag{31}$$

Note second that $\Phi^1(p)$ is defined to be $\varphi_{O,X}(p_X)$, and $\Phi^2(p)$ is defined to be $\varphi_{O,Y}(p_Y)$, and similarly for Z and I. Hence:

$$\mathbf{v}_{O,p} = \Phi^1(p)\mathbf{e}_1^\Phi + \Phi^2(p)\mathbf{e}_2^\Phi + \Phi^3(p)\mathbf{e}_3^\Phi + \Phi^4(p)\mathbf{e}_4^\Phi. \tag{32}$$

□

Lemma 32. $\mathbf{v}_{p,q} = \sum_{a=1}^4 (\Phi^a(q) - \Phi^a(p))\, \mathbf{e}_a^\Phi$.

Proof. This is verified as follows:

$$\begin{aligned} \mathbf{v}_{p,q} &= \mathbf{v}_{p,O} + \mathbf{v}_{O,q} = (-\mathbf{v}_{O,p}) + \mathbf{v}_{O,q} = \mathbf{v}_{O,q} - \mathbf{v}_{O,p} \\ &= \sum_{a=1}^{4} \Phi^a(q)\, \mathbf{e}_a^\Phi - \sum_{a=1}^{4} \Phi^a(p)\, \mathbf{e}_a^\Phi \\ &= \sum_{a=1}^{4} (\Phi^a(q) - \Phi^a(p))\, \mathbf{e}_a^\Phi, \end{aligned} \tag{33}$$

where we used Chasles's Relation (i.e., $\mathbf{v}_{p,q} + \mathbf{v}_{q,r} = \mathbf{v}_{p,r}$), some properties of vectors, and then lemma 31 to expand $\mathbf{v}_{O,q}$ and $\mathbf{v}_{O,p}$ into their components in the Φ-basis.

□

Note that the vector $\mathbf{v}_{p,q}$ from p to q is entirely coordinate-independent.

Let us now assume we are considering a full model $M \vDash_2 \mathsf{Gal}(1,3)$, with $M = (\mathbb{P}, B, \sim, \equiv^\sim)$.

Lemma 33. *Any simultaneity hypersurface Σ in M is a three-dimensional affine space.*

Proof. If Σ_p is a simultaneity hypersurface, then, by $\mathsf{EG}(3)^\sim$, the restriction $(\Sigma_p, B \restriction_{\Sigma_p}, (\equiv^\sim) \restriction_{\Sigma_p})$ is a Euclidean three-space isomorphic to $(\mathbb{R}^3, B_{\mathbb{R}^3}, \equiv_{\mathbb{R}^3})$ by theorem 63. Since the reduct $(\Sigma_p, B \restriction_{\Sigma_p})$ (i.e., forgetting the congruence relation) of a Euclidean 3-space is an affine 3-space, Σ_p is an affine three-space and indeed isomorphic to $(\mathbb{R}^3, B_{\mathbb{R}^3})$.

□

Definition 34. We define the *horizontal, or simultaneity, vector subspace* $\mathbb{V}^\sim$ as follows:

$$\mathbb{V}^\sim := \{\mathbf{v}_{p,q} \in \mathbb{V} \mid p \sim q\}. \tag{34}$$

Definition 34 yields:

Lemma 35. *$p \sim q$ iff $\mathbf{v}_{p,q} \in \mathbb{V}^\sim$.*

From lemma 33, we obtain:

Lemma 36. *$\mathbb{V}^\sim$ is a three-dimensional linear subspace of $\mathbb{V}$.*

Definition 37. We define $p + \mathbb{V}^\sim := \{q \in \mathbb{P} \mid \mathbf{v}_{p,q} \in \mathbb{V}^\sim\}$.

Lemma 38. *$q \in p + \mathbb{V}^{\sim}$ if and only if $p \sim q$.*

Proof. This is immediate from definition 37 and lemma 35.

□

Lemma 39. $\Sigma_p = p + \mathbb{V}^{\sim}$.

Proof. $q \in \Sigma_p$, if and only if $p \sim q$, if and only if (lemma 38) $q \in p + \mathbb{V}^{\sim}$.

□

Lemma 40. *Let a Galilean frame O, X, Y, Z, I be given, and let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ be defined as in definition 26. Then the subset $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is a basis for $\mathbb{V}^{\sim}$.*

Proof. The proof is that the vectors $\mathbf{v}_{O,X}$, $\mathbf{v}_{O,Y}$, and $\mathbf{v}_{O,Z}$ each lie in $\mathbb{V}^{\sim}$, and, moreover, given any point $q \in \Sigma_O$, the vector $\mathbf{v}_{O,q}$ is a linear combination of $\mathbf{v}_{O,X}$, $\mathbf{v}_{O,Y}$, and $\mathbf{v}_{O,Z}$.

□

Lemma 41. *Given a coordinate system Φ, the set $\{\mathbf{e}_1^{\Phi}, \mathbf{e}_2^{\Phi}, \mathbf{e}_3^{\Phi}\}$ is a basis for $\mathbb{V}^{\sim}$.*

This is a corollary of the previous lemma.

Lemma 42. *Let a Galilean frame O, X, Y, Z, I be given, and let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ be defined as in definition 26 above. Let $\mathbf{v} \in \mathbb{V}$ with $\mathbf{v} = \sum_{i=1}^{4} v^i \mathbf{e}_i$. Then*

$$\mathbf{v} \in \mathbb{V}^{\sim} \leftrightarrow v^4 = 0. \tag{35}$$

Proof. Let p be any point, and consider:

$$p' = p + \mathbf{v} = p + \sum_{i=1}^{3} v^3 \mathbf{e}_i + \sum_{i=1}^{4} v^4 \mathbf{v}_{O,I}. \tag{36}$$

So, $\mathbf{v} = \mathbf{v}_{p,p'}$. If $\mathbf{v}_{p,p'} \in \mathbb{V}^{\sim}$, we infer that: $\mathbf{v}_{p,p'} = \alpha^1 \mathbf{e}_1 + \alpha^2 \mathbf{e}_2 + \alpha^3 \mathbf{e}_3$ (for some coefficients $\alpha^i \in \mathbb{R}$) by lemma 36. Equating coefficients, we conclude that $\alpha^i = v^i$ (for $i = 1, 2, 3$) and $v^4 = 0$, as claimed. Conversely, if $v^4 = 0$, we infer: $\mathbf{v}_{p,p'} = \sum_{i=1}^{3} v^i \mathbf{e}_i + \sum_{i=1}^{4} 0.\mathbf{v}_{O,I} = \sum_{i=1}^{3} v^i \mathbf{e}_i$. And thus, $\mathbf{v}_{p,p'} \in \mathbb{V}^{\sim}$. This implies that $\mathbf{v} \in \mathbb{V}^{\sim}$.

□

Definition 43. Let Σ_p and Σ_q be simultaneity hypersurfaces. We say that Σ_p is *parallel* to Σ_q if and only if either $\Sigma_p = \Sigma_q$ or there is no intersection of Σ_p and Σ_q. This is written: $\Sigma_p \parallel \Sigma_q$.

Lemma 44. *All simultaneity hypersurfaces are parallel.*

Proof. Let Σ_p and Σ_q be simultaneity hypersurfaces. For a contradiction, suppose $\Sigma_p \nparallel \Sigma_q$. So, $\Sigma_p \neq \Sigma_q$, and there is an intersection $r \in \Sigma_p \cap \Sigma_q$. So, $r \sim p$ and $r \sim q$. Hence, $p \sim q$. Hence, $\Sigma_p = \Sigma_q$, a contradiction.

□

Lemma 45 (Translation Invariance of Simultaneity). *If $p \sim q$, then $(p + \mathbf{v}) \sim (q + \mathbf{v})$.*

Proof. Suppose $p \sim q$. So, we have: $\mathbf{v}_{p,q} \in \mathbb{V}^{\sim}$. Consider $p' = p + \mathbf{v}$ and $q' = q + \mathbf{v}$. Let $\mathbf{w} = \mathbf{v}_{p,q}$. Since $q = p + \mathbf{w}$, we have $q + \mathbf{v} = (p + \mathbf{w}) + \mathbf{v}$, which implies (using some properties of vector addition and the action) $q' = p' + \mathbf{w}$. Hence, $\mathbf{w} = \mathbf{v}_{p',q'}$. So, $\mathbf{v}_{p',q'} = \mathbf{v}_{p,q}$. Since $\mathbf{v}_{p,q} \in \mathbb{V}^{\sim}$, we infer: $\mathbf{v}_{p',q'} \in \mathbb{V}^{\sim}$. From this, it follows that $p' \sim q'$.

□

Lemma 46. *Given a simultaneity hypersurface Σ and time axis T, there is a unique intersection point lying in both Σ and T.*

Proof. Let hypersurface Σ and time axis T be given. There cannot be two distinct intersections, say q and q', for then we should have $q \sim q'$, contradicting the assumption that T is a time axis. To establish the existence of at least one intersection, let us fix a Galilean frame O, X, Y, Z, I with $O, I \in T$, i.e., $T = \ell(O, I)$. For any point p, we have that there exist unique coefficients v^i and v^4 such that:

$$p = O + \sum_{i=1}^{3} v^i \mathbf{e}_i + v^4 \mathbf{v}_{O,I}. \tag{37}$$

Pick any point $p \in \Sigma$ (so $\Sigma = \Sigma_p$). Next, define the point q:

$$q := O + v^4 \mathbf{v}_{O,I}. \tag{38}$$

Then, we infer $\mathbf{v}_{O,q} = v^4 \mathbf{v}_{O,I}$, which implies that $q \in T$. Next, consider $\mathbf{v}_{q,p}$:

$$\mathbf{v}_{q,p} = \mathbf{v}_{q,O} + \mathbf{v}_{O,p} = -v^4 \mathbf{v}_{O,I} + \sum_{i=1}^{3} v^i \mathbf{e}_i + v^4 \mathbf{v}_{O,I} = \sum_{i=1}^{3} v^i \mathbf{e}_i. \tag{39}$$

Since $\mathbf{v}_{q,p} = \sum_{i=1}^{3} v^i \mathbf{e}_i$ and $\sum_{i=1}^{3} v^i \mathbf{e}_i \in \mathbb{V}^{\sim}$, it follows that $q \sim p$. This implies that $q \in \Sigma_p$, and therefore $q \in \Sigma$. The defined point q is, therefore, the required intersection of T and Σ.

□

Definition 47. Let $\ell = \ell(p, q)$ (with $p \neq q$) be a line, and let Σ be a simultaneity hypersurface. We say that ℓ is *parallel* to Σ if and only if either $\ell \subseteq \Sigma$ or there is no intersection $r \in T \cap \Sigma$. This is written: $\ell \parallel \Sigma$.

Lemma 48. *No time axis is parallel to a simultaneity hypersurface.*

Proof. Let $T = \ell(p, q)$ be a time axis (i.e., $p \nsim q$). Let Σ be a simultaneity hypersurface. For a contradiction, suppose $T \parallel \Sigma$. So, either $\ell(p, q) \subseteq \Sigma$ or there is no intersection $r \in T \cap \Sigma$. But, by lemma 46, there is a unique intersection $r \in T \cap \Sigma$. So, we must have: $\ell(p, q) \subseteq \Sigma$. Then, since $p, q \in \ell(p, q)$, we have $p, q \in \Sigma$. Hence, $p \sim q$, a contradiction. Therefore, $T \nparallel \Sigma$. □

Lemma 49. *Let lines $\ell(p, q)$ and $\ell(r, s)$ be parallel. Then, for some $\alpha \neq 0$, $\mathbf{v}_{p,q} = \alpha \mathbf{v}_{r,s}$.*

Proof. This follows from the detailed construction of $\mathbb{V}$ (based on parallelograms and equipollence), which yields theorem 68. □

Lemma 50. *Any line parallel to a time axis is a time axis.*

Proof. Suppose line $\ell(p, q)$ is parallel to a time axis $T = \ell(O, I)$, with $O \nsim I$. Then, by lemma 49, $\mathbf{v}_{p,q} = \alpha \mathbf{v}_{O,I}$, with $\alpha \neq 0$. Since $O \nsim I$, we have $\mathbf{v}_{O,I} \notin \mathbb{V}^{\sim}$. In general, for any $\alpha \neq 0$, $\mathbf{v} \in \mathbb{V}^{\sim}$ if and only if $\alpha \mathbf{v} \in \mathbb{V}^{\sim}$. So, it follows that $\mathbf{v}_{p,q} \notin \mathbb{V}^{\sim}$. Hence, $p \nsim q$. Thus, $\ell(p, q)$ is a time axis. □

4.6 *Representation*

Definition 51. Let $M = (\mathbb{P}, B, \sim, \equiv^{\sim})$ be a σ_{Gal}-structure (i.e., B interprets `Bet`, $\sim$ interprets $\sim$, and $\equiv^{\sim}$ interprets $\equiv^{\sim}$). Suppose that $M \vDash_2 \mathrm{Gal}(1, 3)$. Let $\Phi : \mathbb{P} \to \mathbb{R}^4$ be a function. We say:

$B_{\mathbb{R}^4}$ *represents* B wrt Φ	iff	for all $p, q, r \in \mathbb{P}$: $B(p, q, r) \leftrightarrow (\Phi(p), \Phi(q), \Phi(r)) \in B_{\mathbb{R}^4}$.
$\sim_{\mathbb{R}^4}$ *represents* $\sim$ wrt Φ	iff	for all $p, q \in \mathbb{P}$: $p \sim q \leftrightarrow \Phi(p) \sim_{\mathbb{R}^4} \Phi(q)$.
$\equiv^{\sim}_{\mathbb{R}^4}$ *represents* $\equiv^{\sim}$ wrt Φ	iff	for all $p, q, r, s \in \mathbb{P}$: $pq \equiv^{\sim} rs \leftrightarrow \Phi(p)\Phi(q) \equiv^{\sim}_{\mathbb{R}^4} \Phi(r)\Phi(s)$

If Φ is a bijection and each of the three above representation conditions holds, then Φ is an *isomorphism* from M to $\mathbb{G}^{(1,3)}$.

In order to prove the Representation Theorem for Gal(1, 3), we need to establish three main lemmas. I call these the Chronology Lemma, the Galilean Frame Translation Invariance Lemma, and the Congruence Lemma.

4.7 *The Chronology Lemma*

Lemma 52 (Chronology). *Let* $M = (\mathbb{P}, B, \sim, \equiv^{\sim})$ *be a* σ_{Gal}*-structure, with* $M \vDash_2$ Gal(1, 3)*. Let* O, X, Y, Z, I *be a* sim *4-frame in* M*. Since* $(\mathbb{P}, B) \vDash_2$ BG(4)*, let* $\Phi : (\mathbb{P}, B) \to (\mathbb{R}^4, B_{\mathbb{R}^4})$ *be an isomorphism matching* O, X, Y, Z, I*. Then* $\sim_{\mathbb{R}^4}$ *represents* $\sim$ *wrt* Φ*.*

Proof. Since O, X, Y, Z, I is a *sim* 4-frame, the points O, X, Y, Z are simultaneous, not coplanar, and $O \nsim I$. Given that Φ matches O, X, Y, Z, I, with O, X, Y, Z simultaneous, the associated basis $\{\mathbf{e}_1^{\Phi}, \mathbf{e}_2^{\Phi}, \mathbf{e}_3^{\Phi}\}$ is a basis for the simultaneity vector space $\mathbb{V}^{\sim}$, by lemma 41. Since a sim 4-frame is a 4-frame, $\{\mathbf{e}_1^{\Phi}, \mathbf{e}_2^{\Phi}, \mathbf{e}_3^{\Phi}, \mathbf{e}_4^{\Phi}\}$ is a basis for $\mathbb{V}$. Let points p, q be given. We claim:

$$p \sim q \quad \leftrightarrow \quad \Phi^4(p) = \Phi^4(q). \tag{40}$$

From lemma 35, we have that $p \sim q$ holds if and only if $\mathbf{v}_{p,q} \in \mathbb{V}^{\sim}$. Using lemma 32, we next expand $\mathbf{v}_{p,q}$ in the basis $\{\mathbf{e}_a^{\Phi}\}$ determined by Φ:

$$\mathbf{v}_{p,q} = \sum_{a=1}^{4} (\Phi^a(q) - \Phi^a(p))\mathbf{e}_a^{\Phi}. \tag{41}$$

From lemma 42, we conclude that $\mathbf{v}_{p,q} \in \mathbb{V}^{\sim}$ iff $(\mathbf{v}_{p,q})^4 = 0$. That is, $p \sim q$ iff $\Phi^4(q) - \Phi^4(p) = 0$. And therefore, $p \sim q$ iff $\Phi^4(q) = \Phi^4(p)$, as claimed. □

4.8 *The Galilean Frame Translation Invariance Lemma*

Lemma 53 (Galilean Frame Translation Invariance). *Let* $M = (\mathbb{P}, B, \sim, \equiv^{\sim})$ *be a* σ_{Gal}*-structure, with* $M \vDash_2$ Gal(1, 3)*. Let* O, X, Y, Z, I *be a Galilean 4-frame in* M*. Since* $(\mathbb{P}, B) \vDash_2$ BG(4)*, let* $\Phi : (\mathbb{P}, B) \to (\mathbb{R}^4, B_{\mathbb{R}^4})$ *be an isomorphism matching* O, X, Y, Z, I*. Let* $\mathbf{v} \in \mathbb{V}$*. Let* $O' = O + \mathbf{v}$*,* $X' = X + \mathbf{v}$*,* $Y' = Y + \mathbf{v}$*,* $Z' = Z + \mathbf{v}$*,* $I' = I + \mathbf{v}$*. Then* O', X', Y', Z', I' *is a Galilean 4-frame.*

That is, leaving the assumptions as stated, when we apply a translation (given by a vector $\mathbf{v}$) to a Galilean frame, so $O' = O + \mathbf{v}$, etc., the result is also a Galilean frame:

$$O, X, Y, Z, I \text{ is a Galilean 4-frame iff } O', X', Y', Z' \text{ is a Galilean 4-frame.} \tag{42}$$

Proof. Without loss of generality, we may suppose that $\mathbf{v}$ does not lie in the simultaneity hypersurface Σ_O. For if it does, the vector will simply translate the frame “horizontally,” along within Σ_O and the Euclidean axioms, along with the fact that the temporal benchmark point I also moves “horizontally” too within the hypersurface Σ_I, guarantee that O', X', Y', Z', I' is a Galilean 4-frame.

I will sketch how the proof goes. It is best illustrated by figure 2.

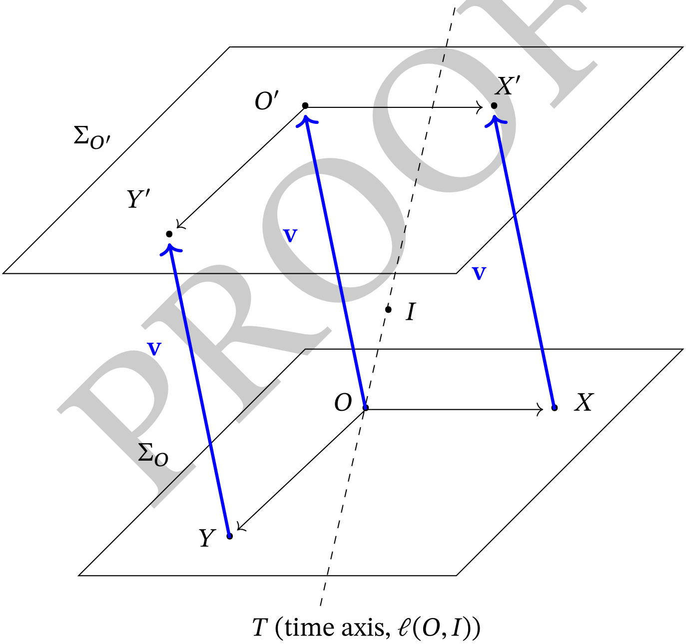

Figure 2: “Transformed Galliean frame” on $\Sigma_{O'}$ (axis $\ell(O, Z)$ and point I' suppressed).

This is the sole part of our analysis appealing to the axiom Gal5, stating the translation invariance of $\equiv^{\sim}$.

The five points O, X, Y, Z, I form a Galilean frame, and thus the four points O, X, Y, Z form a Euclidean sim 3-frame. So, in the lower simultaneity hypersurface, Σ_O, we have a Euclidean sim 3-frame O, X, Y, Z: the three legs OX, OY, and OZ are perpendicular and of equal length. (The point Z and the axis $\ell(O, Z)$ are suppressed in figure 2.)

Consider the hypersurface $\Sigma_{O'}$. By assumption, each of the points O', X', Y', Z' is obtained by adding the *same* displacement vector: $\mathbf{v} = \mathbf{v}_{O,O'}$:

$$\begin{aligned} O' &= O + \mathbf{v}, \quad X' = X + \mathbf{v}, \\ Y' &= Y + \mathbf{v}, \quad Z' = Z + \mathbf{v}. \end{aligned} \tag{43}$$

Since O, X, Y, Z are simultaneous, it follows, using lemma 45, that O', X', Y', Z' are simultaneous. So, all four points lie in $\Sigma_{O'}$.

Next, we use the Translation Invariance axiom Gal5 of Gal(1, 3): $\equiv^{\sim}$ is translation invariant. Since O, X, Y, Z form a Euclidean sim 3-frame, we may conclude, from the translation invariance of $\equiv^{\sim}$, that O', X', Y', Z' is also a Euclidean sim 3-frame. Since $\mathbf{v}$ does not lie parallel to Σ_O, I' is not simultaneous with O', X', Y', Z'. And, so, O', X', Y', Z', I' is a Galilean 4-frame.

□

4.9 *The Congruence Lemma*

Lemma 54 (Congruence). *Let $M = (\mathbb{P}, B, \sim, \equiv^{\sim})$ be a σ_{Gal}-structure, with $M \models_2 \text{Gal}(1,3)$. Let O, X, Y, Z, I be a Galilean 4-frame in M. By the Chronology Lemma (lemma 52), there is an isomorphism $\Phi : (\mathbb{P}, B, \sim) \to (\mathbb{R}^4, B_{\mathbb{R}^4}, \sim_{\mathbb{R}^4})$ matching O, X, Y, Z, I. Then, $\equiv^{\sim}_{\mathbb{R}^4}$ represents $\equiv^{\sim}$ with respect to Φ.*

Proof. We are given a structure $M = (\mathbb{P}, B, \sim, \equiv^{\sim})$, a Galilean frame, O, X, Y, Z, I in M, and an isomorphism $\Phi : (\mathbb{P}, B, \sim) \to (\mathbb{R}^4, B_{\mathbb{R}^4}, \sim_{\mathbb{R}^4})$, matching O, X, Y, Z, I. We shall call Φ the “global isomorphism.” We claim that $\equiv^{\sim}_{\mathbb{R}^4}$ represents $\equiv^{\sim}$ with respect to Φ; that is, for simultaneous points p, q, r, s, we have:[40]

$$pq \equiv^{\sim} rs \leftrightarrow \Delta_3(\vec{\Phi}(p), \vec{\Phi}(q)) = \Delta_3(\vec{\Phi}(r), \vec{\Phi}(s)). \tag{44}$$

Consider figure 3:

40 Where $\vec{\Phi}(p)$ is the triple $(\Phi^1(p), \Phi^2(p), \Phi^3(p)) \in \mathbb{R}^3$, and Δ_3 is the metric function on $\mathbb{R}^3$.

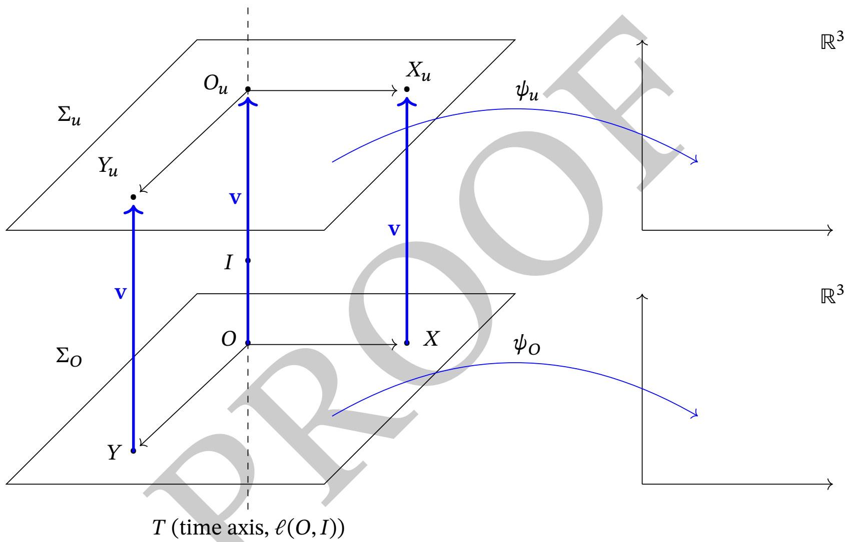

Figure 3: "Lifted Euclidean frame" on Σ_u (axis $\ell(O, Z)$ suppressed).

By hypothesis, the five points O, X, Y, Z, I form a Galilean frame, and thus the four points O, X, Y, Z form a Euclidean sim 3-frame. For points in the lower simultaneity hypersurface, Σ_O, we have, from the Euclidean axiom group EG(3)$^{\sim}$ in Gal(1, 3) and the Representation Theorem for Euclidean geometry (theorem 63), the existence of an isomorphism (i.e., coordinate system on Σ_O),

$$\psi_O : (\Sigma_O, B \restriction_{\Sigma_O}, (\equiv^{\sim}) \restriction_{\Sigma_O}) \to (\mathbb{R}^3, B_{\mathbb{R}^3}, \equiv_{\mathbb{R}^3}), \tag{45}$$

that matches this Euclidean sim 3-frame O, X, Y, Z. So, in the hypersurface Σ_O, a "mini-representation theorem" holds. For any $p, q, r, s \in \Sigma_O$,

$$pq \equiv^{\sim} rs \leftrightarrow \vec{\psi}_O(p)\vec{\psi}_O(q) \equiv_{\mathbb{R}^3} \vec{\psi}_O(r), \vec{\psi}_O(s). \tag{46}$$

Let Φ_O be $\Phi \restriction_{\Sigma_O}$: the restriction of the global isomorphism Φ to the hypersurface Σ_O. We are also given that Φ_O also matches O, X, Y, Z. By the uniqueness of coordinate systems that match the same frame (lemma 66), we conclude:

$$\psi_O = \Phi_O. \tag{47}$$

Thus, by (46) and (47), Φ_O satisfies:

$$pq \equiv^{\sim} rs \leftrightarrow \vec{\Phi}_O(p)\vec{\Phi}_O(q) \equiv_{\mathbb{R}^3} \vec{\Phi}_O(r), \vec{\Phi}_O(s). \tag{48}$$

We now repeat the same argument for an arbitrary simultaneity surface, Σ_u.

Given any point u, we consider the hypersurface Σ_u. By lemma 46, the time axis $\ell(O, I)$ intersects Σ_u at the corresponding "origin," O_u. By lemma 22, there are unique lines through X, Y, and Z, each parallel to $\ell(O, O_u)$. By lemma 46 again, these intersect Σ_u at points X_u, Y_u, Z_u. By lemma 44, the hypersurfaces Σ_O and Σ_u are parallel; this guarantees that each of the points O_u, X_u, Y_u, Z_u is obtained by adding the *same* displacement vector: $\mathbf{v} = \mathbf{v}_{O,O_u}$:

$$\begin{aligned} O_u &= O + \mathbf{v}, \quad X_u = X + \mathbf{v}, \\ Y_u &= Y + \mathbf{v}, \quad Z_u = Z + \mathbf{v}. \end{aligned} \tag{49}$$

By the Translation Invariance of Galilean frames, lemma 53, since O, X, Y, Z, I form a Galilean 4-frame, we may conclude that O_u, X_u, Y_u, Z_u, I_u (where $I_u = I + \mathbf{v}$) also form a Galilean 4-frame. And thus, O_u, X_u, Y_u, Z_u

form a Euclidean sim 3-frame. By the Representation Theorem for Euclidean geometry, there is an isomorphism ψ_u, which matches O_u, X_u, Y_u, Z_u. By similar reasoning to the case of Σ_O, we define the restriction Φ_u to be $\Phi \restriction_{\Sigma_u}$—i.e., the restriction of the global isomorphism Φ to the hypersurface Σ_u. We can conclude:

$$\psi_u = \Phi_u. \tag{50}$$

Thus, Φ_u satisfies the following: for any points $p, q, r, s \in \Sigma_u$,

$$pq \equiv^{\sim} rs \quad \leftrightarrow \quad \vec{\Phi}_u(p)\,\vec{\Phi}_u(q) \equiv_{\mathbb{R}^3} \vec{\Phi}_u(r)\,\vec{\Phi}_u(s). \tag{51}$$

This is equivalent to (44).

□

5 Representation Theorem for Gal(1, 3)

Our main theorem is then the following:

Theorem 55 (Representation Theorem for Galilean Spacetime). *Let* $M = (\mathbb{P}, B, \sim, \equiv^{\sim})$ *be a full* σ_{Gal}*-structure. Then*

$$M \vDash_2 \mathrm{Gal}(1,3) \quad \textit{if and only if} \quad \textit{there is an isomorphism } \Phi : M \to \mathbb{G}^{(1,3)}. \tag{52}$$

Proof. For the right-to-left direction, suppose there is an isomorphism $\Phi : M \to \mathbb{G}^{(1,3)}$. So, $M \cong \mathbb{G}^{(1,3)}$. By the Soundness Lemma (lemma 21), $\mathbb{G}^{(1,3)} \vDash_2 \mathrm{Gal}(1,3)$. Since isomorphic structures satisfy the same sentences, it follows that $M \vDash_2 \mathrm{Gal}(1,3)$.

For the converse, let $M \vDash_2 \mathrm{Gal}(1,3)$. From the Galilean Frame Lemma (lemma 25), a Galilean frame O, X, Y, Z, I exists. This is a 4-frame. By the Representation Theorem for BG(4) (theorem 62), we conclude that there is a global isomorphism:

$$\Phi : \mathbb{P} \to \mathbb{R}^4 \tag{53}$$

such that Φ matches the frame O, X, Y, Z, I, and $\Phi : (\mathbb{P}, B) \to (\mathbb{R}^4, B_{\mathbb{R}^4})$ is an isomorphism. So, $B_{\mathbb{R}^4}$ represents the betweenness relation B of M with respect to Φ. Recall that the global isomorphism Φ matches a *Galilean* frame O, X, Y, Z, I. Since O, X, Y, Z, I is a Galilean frame, O, X, Y, Z, I is a *sim* frame. By the Chronology Lemma (lemma 52), we conclude that the relation $\sim_{\mathbb{R}^4}$ represents the simultaneity relation $\sim$ of M with respect to Φ. What is more, again, since O, X, Y, Z, I is a Galilean frame, we can appeal to the Congruence Lemma (lemma 54) and conclude that $\equiv^{\sim}_{\mathbb{R}^4}$ represents the sim-congruence relation $\equiv^{\sim}$ of M with respect to Φ.

Assembling this, $\Phi : M \to \mathbb{G}^{(1,3)}$ is an isomorphism, as claimed.

□

Such isomorphisms $\Phi : M \to \mathbb{G}^{(1,3)}$ are *inertial charts* on Galilean spacetime. They correspond, one-to-one, with Galilean frames. As we have seen, the transformation group between these isomorphisms (or, if you wish, between the Galilean frames) is precisely $\mathcal{G}^e(1,3)$—the extended Galilean group.

Appendices

Appendix A: Axioms

Definition 56. The non-logical axioms of BG(4) in $L(\sigma_{\mathrm{Gal},\in})$ are the following nine:[41]

Table 4: Order axioms for betweenness.

B1	Bet-Identity	$\mathtt{Bet}(p,q,p) \to p = q$.
B2	Bet-Transitivity	$\mathtt{Bet}(p,q,r) \wedge \mathtt{Bet}(q,r,s) \wedge q \neq r \to \mathtt{Bet}(p,q,s)$.
B3	Bet-Connectivity	$\mathtt{Bet}(p,q,r) \wedge \mathtt{Bet}(p,q,r') \wedge p \neq q \to (\mathtt{Bet}(p,r,r') \vee \mathtt{Bet}(p,r',r))$.
B4	Bet-Extension	$\exists p\,(\mathtt{Bet}(p,q,r) \wedge p \neq q)$.
B5	Pasch	$\mathtt{Bet}(p,q,r) \wedge \mathtt{Bet}(s,u,q) \to \exists x\,(\mathtt{Bet}(r,x,s) \wedge \mathtt{Bet}(p,u,x))$.
B6	Euclid	$\mathtt{Bet}(a,d,t) \wedge \mathtt{Bet}(b,d,c) \wedge a \neq d \to \exists x\,\exists y\,(\mathtt{Bet}(a,b,x) \wedge \mathtt{Bet}(a,c,y) \wedge \mathtt{Bet}(x,t,y))$.
B7	Lower Dimension	There exist five points which are not $\mathtt{co}_3$.

41 These axioms are given originally in Szczerba and Tarski (1965, 1979). See also Goldblatt (1987, 165) for the corresponding first-order theory, which we have called $\mathrm{BG}_0(4)$. Goldblatt calls this "the first-order theory of ordered affine fourfolds over real-closed fields."

doi: 10.48106/dial.v77.i2

B8	Upper Dimension	Any six points are co_4.
B9	Continuity Axiom	$[\exists r\,(\forall p \in \mathrm{X}_1)(\forall q \in \mathrm{X}_2)\,\mathtt{Bet}(r,p,q)] \to \exists s\,(\forall p \in \mathrm{X}_1)(\forall q \in \mathrm{X}_2)\,\mathtt{Bet}(p,s,q)$.

See Szczerba and Tarski (1979, 159–160) for the first-order two-dimensional theory GA_2 (for "neutral" or "absolute geometry"), which lacks the Euclid Parallel axiom (which is called (E) in Szczerba and Tarski 1979 and is called (Euclid) above). Their system includes Desargues's Theorem. But, for us, this axiom is no longer required, as it is provable from the remaining axioms in dimensions above two (Szczerba and Tarski 1979, 190). The above axiom system is the second-order, four-dimensional theory and contains (E), i.e., (Euclid). The relevant representation theorem follows from theorem 5.12 of Szczerba and Tarski (1979, 185, see also example 6.1). The same theorem is stated, somewhat indirectly, in Borsuk and Szmielew (1960, 196–197). The representation theorem itself goes back to Veblen (1904).

Definition 57. The non-logical axioms of $\mathsf{EG}(3)$ in $L(\sigma_{\mathtt{Bet},\equiv,\in})$ are the following eleven:

Table 5: The axioms of Euclidean Geometry for three dimensions.

E1	$\mathtt{Bet}$-Identity	$\mathtt{Bet}(p,q,p) \to p = q$.
E2	$\equiv$-Identity	$pq \equiv rr \to p = q$.
E3	$\equiv$-Transitivity	$pq \equiv rs \wedge pq \equiv tu \to rs \equiv tu$.
E4	$\equiv$-Reflexivity	$pq \equiv qp$.
E5	$\equiv$-Extension	$\exists r\,(\mathtt{Bet}(p,q,r) \wedge qr \equiv su)$.
E6	Pasch	$\mathtt{Bet}(p,q,r) \wedge \mathtt{Bet}(s,u,r) \to \exists x\,(\mathtt{Bet}(q,x,s) \wedge \mathtt{Bet}(u,x,p))$.
E7	Euclid	$\mathtt{Bet}(a,d,t) \wedge \mathtt{Bet}(b,d,c) \wedge a \neq d \to \exists x\,\exists y\,(\mathtt{Bet}(a,b,x) \wedge \mathtt{Bet}(a,c,y) \wedge \mathtt{Bet}(x,t,y))$.
E8	5-Segment	$(p \neq q \wedge \mathtt{Bet}(p,q,r) \wedge \mathtt{Bet}(p',q',r') \wedge pq \equiv p'q' \wedge qr \equiv q'r' \wedge ps \equiv p's' \wedge qs \equiv q's') \to rs \equiv r's'$.

E9	Lower Dimension	There exist four points which are not $\mathtt{co}_2$.
E10	Upper Dimension	Any five points are $\mathtt{co}_3$.
E11	Continuity Axiom	$[\exists r\, (\forall p \in \mathrm{X}_1)\, (\forall q \in \mathrm{X}_2)\, \mathtt{Bet}(r, p, q)] \rightarrow \exists s\, (\forall p \in \mathrm{X}_1)\, (\forall q \in \mathrm{X}_2)\, \mathtt{Bet}(p, s, q)$.

The original source of this axiomatization is Tarski (1959) and Tarski and Givant (1999). See Tarski (1959, 19–20) for a formulation of the first-order two-dimensional theory, with twelve axioms and one axiom scheme (for continuity); and Tarski and Givant (1999) for a simplification down to ten axioms and one axiom scheme (for continuity). The above axiom system is the second-order, four-dimensional theory (i.e., the single Continuity Axiom is the second-order one).

Appendix B: Representation Theorems

Definition 58 (4-frame)**.** For betweenness geometry, a *4-frame* is an ordered tuple of five points O, X, Y, Z, I, which are not $\mathtt{co}_3$.[42]

Definition 59 (Perpendicularity)**.** In Euclidean geometry, perpendicularity $OX \perp OY$ for three distinct points O, X, Y is defined as follows: $OX \perp OY$ holds iff $XY \equiv (-X)Y$, where $(-X)$ is the unique point p on $\ell(O, X)$ such that $p \neq X$ and $Op \equiv OX$.

Definition 60 (Euclidean 3-frame)**.** For Euclidean geometry, a Euclidean 3-frame is an ordered quadruple O, X, Y, Z of points that are not $\mathtt{co}_2$ (i.e., not coplanar) and such that the segments OX, OY, OZ are mutually perpendicular and of equal length.

42 Burgess refers to such systems of points as "benchmarks": Burgess and Rosen (1997, 107). For example, in the two-dimensional case, one may imagine marking three non-collinear points O, X, Y on a bench. This will be a "2-frame" and will determine a two-dimensional coordinate system, with O at the origin, $\ell(O, X)$ the "x-axis," and $\ell(O, Y)$ the "y-axis."

doi: 10.48106/dial.v77.i2

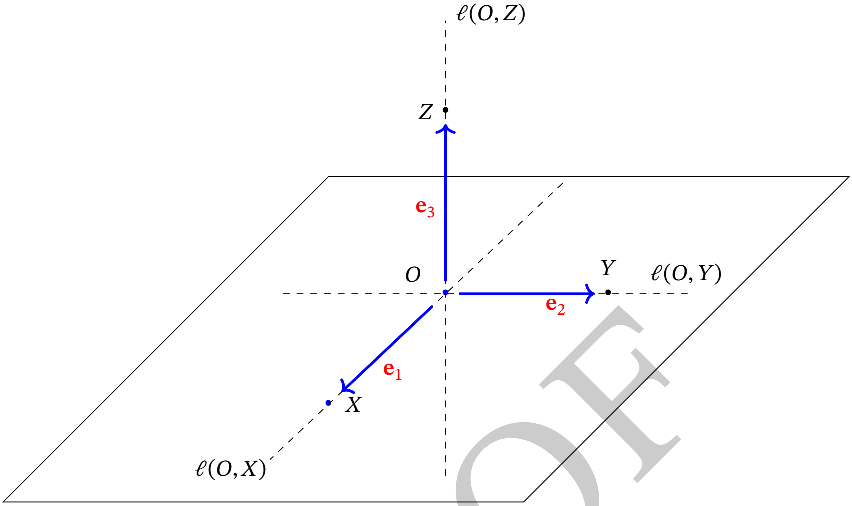

Figure 4: Euclidean 3-Frame.

Definition 61 (Matching)**.** Suppose that $M = (\mathbb{P}, B)$ is a $\sigma_{\mathtt{Bet}}$-structure with $M \vDash_2 \mathrm{BG}(4)$, and suppose that O, X, Y, Z, I is a 4-frame in M. Suppose that $\Phi : \mathbb{P} \to \mathbb{R}^4$ is a function. We say that Φ *matches* O, X, Y, Z, I just if:[43]

$$\Phi(O) = \mathbf{O}, \quad \Phi(X) = \mathbf{X}, \quad \Phi(Y) = \mathbf{Y}, \quad \Phi(Z) = \mathbf{Z}, \quad \Phi(I) = \mathbf{I}. \tag{54}$$

The following two theorems are primarily due to Hilbert (1899), Veblen (1904), and Tarski (1959):[44]

Theorem 62 (Representation Theorem for BG(4))**.** *Let $M = (\mathbb{P}, B)$ be a $\sigma_{\mathtt{Bet}}$-structure. Assume that $M \vDash_2 \mathrm{BG}(4)$. Suppose O, X, Y, Z, I is a 4-frame in M. Then there exists a bijection $\Phi : \mathbb{P} \to \mathbb{R}^4$ such that:*

(a) Φ matches O, X, Y, Z, I.
(b) For all $p, q, r \in \mathbb{P}$: $(p, q, r) \in B \leftrightarrow B_{\mathbb{R}^4}(\Phi(p), \Phi(q), \Phi(r))$.

43 A similar definition, *mutatis mutandis*, can be applied to BG(n) in general and to EG(n) in general.
44 See also Borsuk and Szmielew (1960) and Szczerba and Tarski (1965, 1979).

Proof. I give a brief sketch. Given a 4-frame O, X, Y, Z, I in M, we first define four lines $\ell(O,X)$, $\ell(O,Y)$, $\ell(O,Z)$, and $\ell(O,I)$: these are the "x-axis," "y-axis," "z-axis," and "t-axis" of the 4-frame. One can define (as in Hilbert 1899) geometrical operations $+$, $\times$, and a linear order $\leq$ on each axis (relative to the two fixed parameters that determined that axis). These definitions are explained very clearly in Bennett (1995): for $+$ at p. 48 and for $\times$ at p. 62. Also, see Goldblatt (1987, 23–27). The definition of $\leq$ is given in Tarski (1959, proof of theorem 1). Then, using the betweenness axioms, one shows that, on each axis, $\ell(O,X)$, $\ell(O,Y)$, $\ell(O,Z)$, and $\ell(O,I)$, these definitions specify an ordered field. For details (ignoring the order aspect), see Bennett (1995, 48–72, especially theorem 1, p. 72). What is more, the Continuity Axiom guarantees that this ordered field is a *complete ordered field*. Up to isomorphism, there is exactly one complete ordered field, and this is also rigid. Consequently, there is a (unique) isomorphism $\varphi_{O,X} : \ell(O,X) \to \mathbb{R}$ (and similarly on each axis):

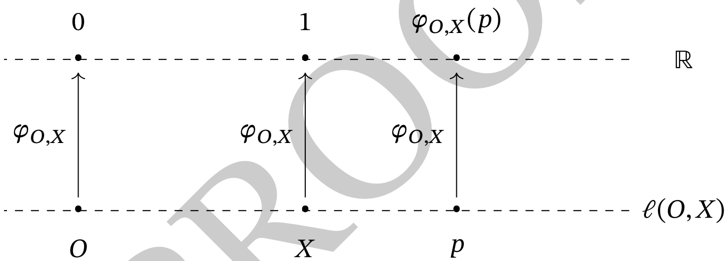

Figure 5: Isomorphism from $\ell(O,X)$ to $\mathbb{R}$.

Given any point p, one then constructs four "ordinates" p_X, p_Y, p_Z, p_I on the four axes $\ell(O,X)$, $\ell(O,Y)$, $\ell(O,Z)$, $\ell(O,I)$ by certain parallel lines to these axes. Then, one defines the coordinate system Φ as follows. Given any point $p \in \mathbb{P}$, define:

$$\Phi(p) := \begin{pmatrix} \varphi_{O,X}(p_X) \\ \varphi_{O,Y}(p_Y) \\ \varphi_{O,Z}(p_Z) \\ \varphi_{O,I}(p_I) \end{pmatrix}. \tag{55}$$

It is clear that Φ matches O, X, Y, Z, I. Finally, one shows that Φ is a bijection and that it satisfies the required isomorphism condition. Namely, for $p, q, r \in \mathbb{P}$: $B(p,q,r)$ iff $B_{\mathbb{R}^4}(\Phi(p), \Phi(q), \Phi(r))$.

doi: 10.48106/dial.v77.i2

$\square$

Theorem 63 (Representation Theorem for EG(3))**.** *Let $M = (\mathbb{P}, B, \equiv)$ be a $\sigma_{\mathtt{Bet},\equiv}$-structure. Assume that $M \vDash_2$ EG(3). Suppose O, X, Y, Z is a Euclidean 3-frame in M. Then there exists a bijection $\Phi : \mathbb{P} \to \mathbb{R}^3$ such that:*

(a) Φ matches O, X, Y, Z.
(b) For all $p, q, r \in \mathbb{P}$: $(p, q, r) \in B \leftrightarrow B_{\mathbb{R}^3}(\Phi(p), \Phi(q), \Phi(r))$.
(c) For all $p, q, r, s \in \mathbb{P}$: $pq \equiv rs \leftrightarrow \Phi(p)\Phi(q) \equiv_{\mathbb{R}^3} \Phi(r)\Phi(s)$.

Roughly, this corresponds to theorem 1 of Tarski (1959), and a sketch of the proof is given there. The difference is that Tarski considers the two-dimensional first-order theory, whose axioms are what we've called $\mathsf{EG}_0(2)$, with the first-order continuity axiom scheme. The Representation Theorem in Tarski (1959) asserts that, given a model $M \vDash \mathsf{EG}_0(2)$ and a Euclidean frame, there is a real-closed field F such that the conditions (a), (b), (c) hold, with $\mathbb{R}$ replaced by that field and "3" replaced by "2." When we strengthen to the second-order Continuity axiom, it follows that this field is in fact $\mathbb{R}$.

Appendix C: Automorphisms and Coordinate Systems

Theorem 64. *The automorphism (symmetry) groups of the structures defined in definitions 1 and 4 are characterized in table 6.*

Table 6: The automorphism groups of standard Euclidean metric space, where $h : \mathbb{R}^n \to \mathbb{R}^n$.

Aut. Group	Condition
$h \in \mathrm{Aut}(\mathbb{BG}^n)$	$(\exists A \in GL(n))\,(\exists \mathbf{d} \in \mathbb{R}^n)\,(\forall \mathbf{x} \in \mathbb{R}^n)\,[h(\mathbf{x}) = A\mathbf{x} + \mathbf{d}]$
$h \in \mathrm{Aut}(\mathbb{EG}^n)$	$(\exists R \in O(n))\,(\exists \lambda \in \mathbb{R} - \{0\})(\exists \mathbf{d} \in \mathbb{R}^n)\,(\forall \mathbf{x} \in \mathbb{R}^n)\,[h(\mathbf{x}) = \lambda R\mathbf{x} + \mathbf{d}]$
$h \in \mathrm{Aut}(\mathbb{EG}^n_{\mathrm{metric}})$	$(\exists R \in O(n))\,(\exists \mathbf{d} \in \mathbb{R}^n)\,(\forall \mathbf{x} \in \mathbb{R}^n)\,[h(\mathbf{x}) = R\mathbf{x} + \mathbf{d}]$

Proof. I give a brief summary. For the first, the proof relies on the requirement that straight lines get mapped to straight lines and parallel lines get mapped to parallel lines. The outcome is that any such mapping h must be an affine transformation generated by a $GL(n)$ matrix A and a translation $\mathbf{d}$. So, the automorphism group is what is usually called Aff(n), the *affine group* in n dimensions. For the third, the symmetry group is the isometry group of the metric space $\mathbb{EG}^n_{\text{metric}}$—thus, what's usually called the *Euclidean group* $E(n)$: rotations, inversions, reflections, and translations (reflections and inversions are $O(n)$ matrices with determinant -1). For the second, which is less familiar, the symmetries include rotations, inversions, reflections, and translations again, but also include *scalings* too:

$$\mathbf{x} \mapsto \lambda\mathbf{x}. \tag{56}$$

The latter are sometimes called *similitudes* or *dilations* (the non-zero factor λ represents this scaling). Although the metric distance between two points is not invariant, nonetheless *metric equalities* are invariant. Imagine a rubber sheet pinned at some central point, say, O, and imagine "stretching" it uniformly and radially from O by some factor. The *distance* between two points on the sheet is not invariant under the stretching: $\Delta(\mathbf{x},\mathbf{y}) \mapsto |\lambda|\Delta(\mathbf{x},\mathbf{y})$, but equality between distances of points (i.e., congruence) is invariant.

□

Lemma 65 (Coordinate Transformations)**.** *Given two coordinate systems $\Phi, \Psi : \mathbb{P} \to \mathbb{R}^4$, on a full model $M = (\mathbb{P}, B)$ of* BG(4)*, they are related as follows: there is a $GL(4)$ matrix A and a translation $\mathbf{d} \in \mathbb{R}^4$ such that, for any point $p \in \mathbb{P}$, we have:*

$$\Psi(p) = A\Phi(p) + \mathbf{d}. \tag{57}$$

This follows from two facts. First, if $\Phi, \Psi : M \to (\mathbb{R}^4, B_{\mathbb{R}^4})$ are isomorphisms, then $\Psi \circ \Phi^{-1} \in \text{Aut}((\mathbb{R}^4, B_{\mathbb{R}^4}))$. Second, we have $\text{Aut}((\mathbb{R}^4, B_{\mathbb{R}^4})) = \text{Aff}(4)$. (This is the result given in theorem 64 for the automorphisms of the standard coordinate structure $(\mathbb{R}^4, B_{\mathbb{R}^4})$ for BG(4).)

Lemma 66. *Given a 4-frame O, X, Y, Z, I and two coordinate systems, Φ, Ψ, on a model M of* BG(4)*, both of which match the frame O, X, Y, Z, I, we have:*

$$\Psi = \Phi. \tag{58}$$

The proof applies the coordinate transformation equation (57) to the five points, O, X, Y, Z, I, which gives five specific instances. The first of these implies that $\mathbf{d} = \mathbf{o}$. The remaining four imply that the $GL(4)$ matrix A is the identity matrix. Similar reasoning applies in any dimension and also to the Euclidean case.

doi: 10.48106/dial.v77.i2

Appendix D: Reals and Vectors

Given a model $(\mathbb{P}, B) \models_2 \mathrm{BG}(4)$, we know, by theorem 62, that it is isomorphic to the standard coordinate structure $(\mathbb{R}^4, B_{\mathbb{R}^4})$.

Using abstraction (or, equivalently, a quotient construction), we can extend $(\mathbb{P}, B)$ with a new sort (or "universe" or carrier set) $\mathfrak{R}$ (of ratios) and operations $0, 1, +, \times, \leq$ to a two-sorted structure $(\mathbb{P}, \mathfrak{R}; B; 0, 1, +, \times, \leq)$ where the reduct $(\mathfrak{R}; 0, 1, +, \times, \leq)$ is isomorphic to $\mathbb{R}$ (as an ordered field).[45] Call a triple p, q, r of points a *configuration* just if $p \neq q$ and p, q, r are collinear. This abstraction proceeds by the equivalence relation on configurations (p, q, r) of *proportionateness*. In geometrical terms, there are three basic cases of proportionateness:[46]

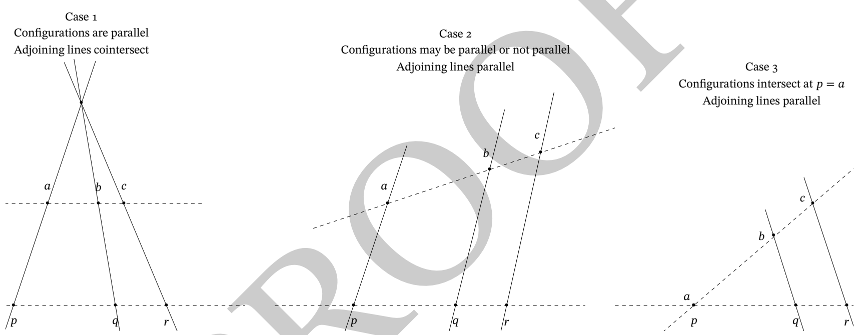

Figure 6: Proportionate Configurations

A real, or ratio, is then an equivalence class $[(p, q, r)]$ with respect to proportionateness, and $\mathfrak{R}$ is the set of these equivalence classes. One may define a zero 0 as $[(p, q, p)]$ and a unit 1 as $[(p, q, q)]$. One defines field operations $+, \times, \leq$ in terms of the corresponding operations on a fixed line (see Bennett

45 I included this appendix, in part, because I had difficulty locating the material elsewhere. One important textbook, Bennett (1995), where the definitions of addition $+$ and multiplication $\times$ on a line $\ell(p, q)$, and the proof that these induce a division ring (given Desargues' Theorem) or a field (given Pappus's Theorem), are explained very clearly, is out of print. Also, because I need, in the main part of the article, to refer to a couple of the summary theorems at the end of this appendix.

46 Burgess and Rosen (1997, 110) list two basic cases, our Case 1 and Case 3. In a sense, Case 3 is a limiting case of Case 2 by "sliding" the configuration *abc* parallel to the three parallel lines until *a* now coincides with *p*.

1995). One readily checks that the result is that $\mathfrak{R}$, with these operations, is a complete ordered field (and can then be identified with $\mathbb{R}$). Although we described this model theoretically, this construction can be "internalized" within BG(4) by adding suitable abstraction axioms (a "definition by abstraction") for a new sort, with variables ξ_i and a 3-place function symbol $\xi(p,q,r)$, and then explicitly defining $0, 1, +, \times$, and $\leq$ on these new objects, and then proving that the resulting abstracta, i.e., the $\xi(p,q,r)$ for any configuration p,q,r, satisfy the second-order axioms for a complete ordered field.[47]

We may further extend, with a new universe $\mathbb{V}$ (of displacements, or vectors) and operations $\mathbf{o}, +, \cdot$, to a three-sorted structure $(\mathbb{P}, \mathfrak{R}, \mathbb{V}; B, 0, 1, +, \times, \leq; \mathbf{o}, +, \cdot)$, where the reduct $(\mathbb{V}, \mathfrak{R}; 0, 1, +, \times; \mathbf{o}, +, \cdot)$ is isomorphic to $\mathbb{R}^4$ (as a vector space).[48] This abstraction proceeds by the equivalence relation on ordered pairs (p,q) of *equipollence*: (p,q) is equipollent to (r,s) just if p,q,s,r is a parallelogram:

A displacement, or vector, is then an equivalence class $[(p,q)]$ with respect to equipollence, and $\mathbb{V}$ is the set of these equivalence classes. An equivalence class $[(p,q)]$ is written $\mathbf{v}_{p,q}$. One may define the zero vector $\mathbf{o}$ as $\mathbf{v}_{p,p}$. One defines vector addition $+$ so that $\mathbf{v}_{p,q} + \mathbf{v}_{q,r} = \mathbf{v}_{p,r}$ holds (usually called *Chasles's Relation*). One may define the scalar multiplication $\cdot$ so that, when $p \neq q$, $\alpha \cdot \mathbf{v}_{p,q} = \mathbf{v}_{p,r}$ just if $\alpha = [(p,q,r)]$; and, otherwise, $\alpha \cdot \mathbf{o} = \mathbf{o}$. One checks that the vector space axioms are true and that $\mathbb{V}$ is 4-dimensional.

Finally, by an explicit definition of an action $+ : \mathbb{P} \times \mathbb{V} \to \mathbb{P}$, we can further extend to $(\mathbb{P}, \mathfrak{R}, \mathbb{V}; B; 0, 1, +, \times, \leq; \mathbf{o}, +, \cdot; +)$ such that $(\mathbb{P}, \mathbb{V}, +)$ is isomorphic to the affine space $\mathbb{A}^4$.[49] The definition of the action $(p, \mathbf{v}) \mapsto p + \mathbf{v}$ is: $q = p + \mathbf{v}$ iff $\mathbf{v} = \mathbf{v}_{p,q}$. One may then show that $+$ is a free and transitive action of $\mathbb{V}$ on

47 The details are given in Burgess (1984). What we've called "configurations," Burgess calls "suitable configurations." For the simple case of "extension by abstraction," with a formula $\varphi(x,y)$ that can be shown to be an equivalence relation in the basic theory T, an extension of T by abstraction is obtained by abstraction axioms (i): $\xi(x) = \xi(y)$ iff $\varphi(x,y)$; and (ii): $\forall \xi\, \exists x\, (\xi = \xi(x))$, where ξ is a new variable sort, and $\xi(x)$ is a function symbol (which Burgess writes as "$[x]$"). See Burgess (1984, 381). Burgess shows (theorem 1.3) that this (indeed any) "extension by abstraction" is a *conservative extension* of the original theory T and may be *interpreted* into the original theory. For the geometrical case, the abstraction axioms are (i): $\xi(p,q,r) = \xi(p',q',r')$ iff the configurations p,q,r and p',q',r' are proportionate; and (ii): $\forall \xi\, \exists p,q,r\, (p \neq q \wedge \mathrm{co}_1(p,q,r) \wedge \xi = \xi(p,q,r))$. See Burgess (1984, 387, axioms (1) and (2)).

48 The two pluses $(+)$ here have been overloaded: the first is the *field* addition, and the second is the *vector* addition.

49 The three pluses $(+)$ here are overloaded: the first is the field addition; the second is the vector addition; and the third is the action, $(p, \mathbf{v}) \mapsto p + \mathbf{v}$.

doi: 10.48106/dial.v77.i2

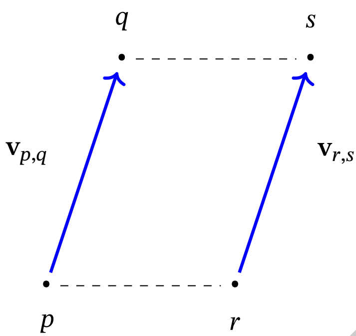

Figure 7: Equipollence.

$\mathbb{P}$. The affine space obtained in this way (basically, from the *vector space* $\mathbb{R}^4$, by "forgetting the origin") is called $\mathbb{A}^4$.

The discussion and constructions above may be summarized in the following three theorems (I follow the usual practice of conflating the name of a structure with the name of its carrier set):

Theorem 67. $\mathfrak{R}$ *is isomorphic to the complete ordered field* $\mathbb{R}$.

Theorem 68. $\mathbb{V}$ *is isomorphic to the vector space* $\mathbb{R}^4$.

Theorem 69. $(\mathbb{P}, \mathbb{V}, +)$ *is isomorphic to the affine space* $\mathbb{A}^4$.

*

Jeffrey Ketland
0000-0002-5128-4387
Institute of Philosophy, University of Warsaw
jeffreyketland@gmail.com

* I am grateful to two anonymous referees for their helpful comments. I am grateful to Professor Victor Pambuccian and Professor Robert Goldblatt for helping me clear up some confused thoughts of mine. I am grateful to Professor David Malament for detailed comments on an earlier draft of this article and for the suggestion of using vectors to simplify things. I am grateful to Joshua Babic and Lorenzo Cocco for some valuable advice. This work was supported by a research grant from the Polish National Science Center (Narodowe Centrum Nauki w Krakowie (NCN), Kraków, Poland), grant number 2020/39/B/HS1/02020.

References

Apostol, Tom M. 1967. *Calculus Volume 1: One-Variable Calculus, with an Introduction to Linear Algebra*. 2nd ed. Hoboken, New Jersey: John Wiley and Sons, Inc.

Arnold, Vladmir I. 1979. *Matematicheskie metody klassicheskoĭ mekhaniki*. 2nd ed. Moscow: Nauka pub.

—. 1989. *Mathematical Methods of Classical Mechanics*. 2nd ed. Graduate Texts in Mathematics n. 60. New York: Springer. Translation of Arnold (1979) by Karen Vogtmann and Alan D. Weinstein, doi:10.1007/978-1-4757-2063-1.

Ax, James. 1978. "The Elementary Foundations of Spacetime." *Foundations of Physics* 8(7/8): 507–546, doi:10.1007/bf00717578.

Bennett, Mary Katherine. 1995. *Affine and Projective Geometry*. Hoboken, New Jersey: John Wiley and Sons, Inc., doi:10.1002/9781118032565.

Boolos, George. 1987. "A Curious Inference." *Journal of Philosophical Logic* 16(1): 1–12. Reprinted in Boolos (1998, 376–382), doi:10.1007/bf00250612.

—. 1998. *Logic, Logic, and Logic*. Cambridge, Massachusetts: Harvard University Press. Introductions and afterword by John P. Burgess; edited by Richard Jeffrey.

Borsuk, Karol and Szmielew, Wanda. 1960. *Foundations of Geometry: Euclidean and Bolyai-Lobachevskian Geometry; Projective Geometry*. Amsterdam: North-Holland Publishing Co. Revised English translation by Erwin Marquit; reissued as Borsuk and Szmielew (2018).

—. 2018. *Foundations of Geometry: Euclidean, Bolyai-Lobachevskian, and Projective Geometry*. Mineola, New York: Dover Publications. Reissue of Borsuk and Szmielew (1960).

Burgess, John P. 1984. "Synthetic Mechanics." *Journal of Philosophical Logic* 13(4): 379–395, doi:10.1007/bf00247712.

Burgess, John P. and Rosen, Gideon. 1997. *A Subject with No Object: Strategies for Nominalistic Interpretation of Mathematics*. Oxford: Oxford University Press, doi:10.1093/0198250126.001.0001.

Cantor, Georg. 1897. "Beiträge zur Begründung der transfiniten Mengenlehre, 2." *Mathematische Annalen* 49: 207–246. Translation by Philip Jourdain in Cantor (1915, 137–201), reprinted in Cantor (1932, 312–356), doi:10.1007/bf01444205.

—, ed. 1915. *Contributions to the Founding of the Theory of Transfinite Numbers*. LaSalle, Illinois: Open Court Publishing Co. Translated with an introduction and notes by Philip Edward Bertrand Jourdain.

—. 1932. *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*. Berlin: Julius Springer. Edited by Ernst Zermelo. Reprinted Hildesheim: Olms, 1966, doi:10.1007/978-3-662-00274-2.

Carnap, Rudolf. 1954. *Einführung in die symbolische Logik, mit besonderer Berücksichtigung ihrer Anwendungen*. Wien: Springer. Third edition: Carnap (1968), doi:10.1007/978-3-7091-3534-1.

—. 1958. *Introduction to Symbolic Logic and Its Applications*. Mineola, New York: Dover Publications. Translation of Carnap (1954) by William H. Meyer and John Wilkinson.

—. 1968. *Einführung in die symbolische Logik, mit besonderer Berücksichtigung ihrer Anwendungen*. Wien: Springer. First edition: Carnap (1954), doi:10.1007/978-3-7091-3140-4.

Cocco, Lorenzo and Babic, Joshua. 2021. "A System of Axioms for Minkowski Spacetime." *Journal of Philosophical Logic* 50(1): 149–185, doi:10.1007/s10992-020-09565-6.

Coxeter, Harold Scott MacDonald. 1969. *Introduction to Geometry*. 2nd ed. Hoboken, New Jersey: John Wiley and Sons, Inc.

Dedekind, Richard. 1872. *Stetigkeit und irrationale Zahlen*. Braunschweig: Vieweg. Translated as "Continuity and irrational numbers" by Wooster W. Beman in Dedekind (1963, 1–27), reprinted, with corrections, in Ewald (1996, 765–779); original text reprinted in Müller-Stach (2017, 111–132), doi:10.1007/978-3-322-98548-4.

—. 1888. *Was sind und was sollen die Zahlen?* Braunschweig: Vieweg. Second edition: Dedekind (1893); English translation by Wooster W. Beman in Dedekind (1901); reprinted, with corrections, in Ewald (1996, 787–833).

—. 1893. *Was sind und was sollen die Zahlen?* 2nd ed. Braunschweig: Vieweg. reprinted in Müller-Stach (2017, 53–109).

—. 1901. *Essays on the Theory of Numbers*. LaSalle, Illinois: Open Court Publishing Co. Translation by Wooster W. Beman; reissue: Dedekind (1963).

—. 1963. *Essays on the Theory of Numbers*. Mineola, New York: Dover Publications. Reissue of Dedekind (1901).

DiSalle, Robert. 2020. "Space and Time: Inertial Frames." in *The Stanford Encyclopedia of Philosophy*. Stanford, California: The Metaphysics Research Lab, Center for the Study of Language and Information, https://plato.stanford.edu/archives/sum2020/entries/spacetime-iframes/.

Dummit, David S. and Foote, Richard M. 2004. *Abstract Algebra*. 3rd ed. Hoboken, New Jersey: John Wiley and Sons, Inc.

Earman, John S. 1970. "Who's Afraid of Absolute Space?" *Australasian Journal of Philosophy* 48(3): 287–319, doi:10.1080/00048407012341291.

—. 1989. *World Enough and Space-Time: Absolute versus Relational Theories of Space and Time*. Cambridge, Massachusetts: The MIT Press.

Ehlers, Jürgen. 1973. "The Nature and Structure of Spacetime." in *The Physicist's Conception of Nature in the Twentieth Century*, edited by Jagdish Mehra, pp. 71–91. Dordrecht: Kluwer Academic Publishers, doi:10.1007/978-94-010-2602-4_6.

Einstein, Albert. 1905. "Zur Elektrodynamik bewegter Körper." *Annalen der Physik* 322(10): 891–921. Reprinted in Lorentz et al. (1923, 35–65), doi:10.1002/andp.19053221004.

Ewald, William Bragg, ed. 1996. *From Kant to Hilbert: A Source Book in the Foundations of Mathematics. Volume II*. Oxford: Oxford University Press, doi:10.1093/oso/9780198505365.001.0001.

Feynman, Richard P., Leighton, Robert B. and Sands, Matthew. 2005. *The Feynman Lectures on Physics: Definitive Edition*. Reading, Massachusetts: Addison-Wesley.

Field, Hartry. 1980. *Science Without Numbers: A Defense of Nominalism*. Princeton, New Jersey: Princeton University Press. Second edition: Field (2016).

—. 2016. *Science Without Numbers: A Defense of Nominalism*. 2nd ed. Oxford: Oxford University Press. First edition: Field (1980), doi:10.1093/acprof:oso/9780198777915.001.0001.

Friedman, Michael. 1983. *Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science*. Princeton, New Jersey: Princeton University Press, doi:10.1515/9781400855124.

Gallier, Jean H. 2001. *Geometric Methods and Applications: For Computer Science and Engineering*. 1st ed. Texts in Applied Mathematics n. 38. New York: Springer. Second edition: Gallier (2011), doi:10.1007/978-1-4613-0137-0.

—. 2011. *Geometric Methods and Applications: For Computer Science and Engineering*. 2nd ed. Texts in Applied Mathematics n. 38. New York: Springer. First edition: Gallier (2001), doi:10.1007/978-1-4419-9961-0.

Gödel, Kurt. 1931. "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter System I." *Monatshefte für Mathematik und Physik* 38(1): 173–198. English translation in van Heijenoort (1967, 596–616); reprinted in Gödel (1986, 145–195), doi:10.1007/BF01700692.

—. 1936. "Über die Länge von Beweisen." *Ergebnisse eines mathematischen Kolloquiums* 7: 23–24. Reprinted and translated in Gödel (1986, 396–398), reprinted, with the original page numbers, in Menger (1998, 341–342).

—. 1986. *Collected Works, Volume I: Publications 1929–1936*. Oxford: Oxford University Press. Edited by Solomon Feferman, John W. Dawson, Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay and Jean van Heijenoort, doi:10.1093/oso/9780195147209.001.0001.

Goldblatt, Robert L. 1987. *Orthogonality and Spacetime Geometry*. New York: Springer, doi:10.1007/978-1-4684-6345-3.

Hartshorne, Robin. 2000. *Geometry: Euclid and Beyond*. Undergraduate Texts in Mathematics. New York: Springer, doi:10.1007/978-0-387-22676-7.

Heath, Thomas L. 1908. *The Thirteen Books of Euclid's Elements, 3 Vol.* 1st ed. Cambridge: Cambridge University Press. Translated from the Text of Heiberg with Introduction and Commentary by Thomas L. Heath.

van Heijenoort, Jan, ed. 1967. *From Frege to Gödel: A Source Book in Mathematical Logic 1879–1931*. Cambridge, Massachusetts: Harvard University Press.

Hilbert, David. 1899. *Grundlagen der Geometrie*. Leipzig: B.G. Teubner. Originally published in "Festschrift zur Feier der Enthüllung des Gauss-Weber Denkmals in

doi: 10.48106/dial.v77.i2

Göttingen.” Later editions published in 1903, 1909, 1911, 1922 and 1930. Translation by Leo Unger of the 10th German Edition: Hilbert (1971).
—. 1971. *Foundations of Geometry*. LaSalle, Illinois: Open Court Publishing Co. Translation by Leo Unger of the 10th German Edition of Hilbert (1899).
Hintikka, Jaakko. 1969. *The Philosophy of Mathematics*. Oxford: Oxford University Press.
Hölder, Otto. 1901. “Die Axiome der Quantität und die Lehre vom Mass.” *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaften der Wissenschaften zu Leipzig, Mathematisch-physische Classe* 53: 1–64.
Huggett, Nick and Hoefer, Carl. 2015. “Absolute and Relational Theories of Space and Motion.” in *The Stanford Encyclopedia of Philosophy*. Stanford, California: The Metaphysics Research Lab, Center for the Study of Language and Information, https://plato.stanford.edu/archives/spr2015/entries/spacetime-theories/.
Huntington, Edward Vermilye. 1903. “Complete Sets of Postulates for the Theory of Real Quantities.” *Transactions of the American Mathematical Society* 4(3): 358–370, doi:10.1090/S0002-9947-1903-1500647-9.
Ketland, Jeffrey. 2021. “Foundations of Applied Mathematics I.” *Synthese* 199(1/2): 4151–4193, doi:10.1007/s11229-020-02973-w.
—. 2022. “A Formalization of Boolos’s ‘Curious Inference’ in Isabelle/HOL.” *Archive for Formal Proofs*, 2022, https://isa-afp.org/entries/Boolos_Curious_Inference.html.
Kopczyński, Wojciech and Trautman, Andrzej. 1992. *Spacetime and Gravitation*. Hoboken, New Jersey: John Wiley and Sons, Inc.
Kordos, Marek. 1969. “On the Syntactic Form of Dimension Axiom for Affine Geometry.” *Bulletin de l’Académie Polonaise des Sciences, Série des Sciences Mathématiques, Astronomiques et Physiques* 17: 833–837.
Longair, Malcolm S. 1984. *Theoretical Concepts in Physics: An Alternative View of Theoretical Reasoning in Physics for Final-Year Undergraduates*. New York: Cambridge University Press. Second edition: Longair (2003); third edition: Longair (2020).
—. 2003. *Theoretical Concepts in Physics: An Alternative View of Theoretical Reasoning in Physics*. 2nd ed. Cambridge: Cambridge University Press. First edition: Longair (1984), doi:10.1017/CBO9780511840173.
—. 2020. *Theoretical Concepts in Physics: An Alternative View of Theoretical Reasoning in Physics*. 3rd ed. Cambridge: Cambridge University Press. First edition: Longair (1984), doi:10.1017/9781108613927.
Lorentz, Hendrik A., Einstein, Albert, Minkowski, Hermann and Weyl, Hermann. 1923. *The Principle of Relativity: A Collection of Original Memoirs on the Special and General Theory of Relativity*. London: Methuen & Co. With notes by Arnold Sommerfeld, translated by W. Perrett and G.B. Jeffery; reprinted as Lorentz et al. (1952).

—, eds. 1952. *The Principle of Relativity: A Collection of Original Memoirs on the Special and General Theory of Relativity*. Mineola, New York: Dover Publications. Reprint of Lorentz et al. (1923).

Malament, David B. 2009. “Notes on Geometry and Spacetime.” Unpublished lecture notes, version 2.7., November 2009, https://philsci-archive.pitt.edu/id/eprint/16760.

—. 2012. *Topics in the Foundations of General Relativity and Newtonian Gravitation Theory*. Chicago Lectures in Physics. Chicago, Illinois: University of Chicago Press.

Menger, Karl, ed. 1998. *Ergebnisse eines mathematischen Kolloquiums*. Wien: Springer. Herausgegeben von Egbert Dierker, Karl Sigmund, mit Beiträgen von John W. Dawson Jr., Ryszard Engelking und Werner Hildenbrand, Geleitwort von Gérard Debreu, Nachwort von Franz Alt, doi:10.1007/978-3-7091-6470-9.

Minkowski, Hermann. 1909. “Raum und Zeit.” *Physikalische Zeitschrift* 10(3): 104–111.

Müller-Stach, Stefan, ed. 2017. *Richard Dedekind: Was sind und was sollen die Zahlen? Stetigkeit und irrationale Zahlen*. Berlin: Springer. second edition: Müller-Stach (2023), doi:10.1007/978-3-662-54339-9.

—, ed. 2023. *Richard Dedekind: Was sind und was sollen die Zahlen? Stetigkeit und irrationale Zahlen*. 2nd ed. Berlin: Springer. first edition: Müller-Stach (2017), doi:10.1007/978-3-662-68142-8.

Mundy, Brent. 1986. “Optical Axiomatization of Minkowski Space-Time Geometry.” *Philosophy of Science* 53(1): 1–30, doi:10.1086/289289.

Newton, Isaac. 1687. *Principia Mathematica Philosophiae Naturalis*. London: Jussu Societatis Regiae ac typis Josephi Streater. Second edition: Newton (1713); third edition: Newton (1726); modern translation: Newton (2016), doi:10.3931/e-rara-440.

—. 1713. *Principia Mathematica Philosophiae Naturalis*. 2nd ed. Cambridge: [publisher unknown]. First edition: Newton (1687), doi:10.3931/e-rara-1237.

—. 1726. *Principia Mathematica Philosophiae Naturalis*. 3rd ed. London: Innys, Regiae Societatis typographos. First edition: Newton (1687), doi:10.3931/e-rara-1235.

—. 2016. *The “Principia”: Mathematical Principles of Natural Philosophy; The Authoritative Translation and Guide*. 2nd ed. Berkeley, California: University of California Press. Translation by I. Bernard Cohen and Anne Whitman, assisted by Julia Budenz, doi:10.1525/9780520964815.

Pambuccian, Victor. 2011. “The Axiomatics of Ordered Geometry I: Ordered Incidence Spaces.” *Expositiones Mathematicae* 29(1): 24–66, doi:10.1016/j.exmath.2010.09.004.

Pasch, Moritz. 1882. *Vorlesungen über neuere Geometrie*. Leipzig: B.G. Teubner.

Penrose, Roger. 1968. “Structure of Space-Time.” in *Battelle Rencontres: 1967 Lectures in Mathematics and Physics*, edited by Cécile DeWitt-Morette and John

Archibald Wheeler, pp. 121–235. San Francisco, California: Benjamin Cummings.

Presburger, Mojżesz. 1930. "Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchen die Addition als einzige Operation hervortritt." in *Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchen die Addition als einzige Operation hervortritt*, pp. 92–101, 395. Warszawa: Państwow Wydawnictwo Naukowe (PWN). Translated in Stansifer (1984), translated and commented by Dale Jacquette in Presburger (1991).

—. 1991. "On the Completeness of a Certain System of Arithmetic of Whole Numbers in Which Addition Occurs as the Only Operation." *History and Philosophy of Logic* 12(2): 225–233. Translation and comment of Presburger (1930) by Dale Jacquette, doi:10.1080/014453409108837187.

Raatikainen, Panu. 2015. "Gödel's Incompleteness Theorems." in *The Stanford Encyclopedia of Philosophy*. Stanford, California: The Metaphysics Research Lab, Center for the Study of Language and Information, https://plato.stanford.edu/archives/spr2015/entries/goedel-incompleteness/.

Rindler, Wolfgang. 1969. *Essential Relativity: Special, General, and Cosmological*. Berlin: Springer.

—. 1977. *Essential Relativity: Special, General, and Cosmological*. 2nd ed. Theoretical and Mathematical Physics. Berlin: Springer. First edition: Rindler (1969), doi:10.1007/978-3-642-86650-0.

Robb, Alfred A. 1911. *Optical Geometry of Motion: A New View of the Theory of Relativity*. Cambridge: Heffer & Sons.

—. 1936. *Geometry of Time and Space*. Cambridge: Cambridge University Press.

Saunders, Simon W. 2013. "Rethinking Newton's *Principia*." *Philosophy of Science* 80(1): 22–48, doi:10.1086/668881.

Schutz, Bernard F. 1980. *Geometrical Methods of Mathematical Physics*. Newcastle upon Tye: Cambridge Scholars Publishing, doi:10.1017/CBO9781139171540.

Schutz, John W. 1997. *Independent Axioms for Minkowski Space-Time*. Reading, Massachusetts: Addison Wesley Longman.

Schwabhäuser, Wolfram, Szmielew, Wanda and Tarski, Alfred. 1983. *Metamathematische Methoden in der Geometrie*. Berlin: Springer, doi:10.1007/978-3-642-69418-9.

Sears, Francis W., Zemansky, Mark W. and Young, Hugh D. 1979. *University Physics*. 5th ed. Reading, Massachusetts: Addison-Wesley.

Shapiro, Stewart. 1983. "Conservativeness and Incompleteness." *The Journal of Philosophy* 80(9): 521–531, doi:10.2307/2026112.

Simpson, Stephen G. and Yokoyama, Keita. 2013. "Reverse Mathematics and Peano Categoricity." *Annals of Pure and Applied Logic* 164(3): 284–293, doi:10.1016/j.apal.2012.10.014.

Stansifer, Ryan. 1984. "Presburger's Article on Integer Arithmetic: Remarks and Translation." TR84-639. Ithaca, New York: Cornell University, Computer Science Department, https://hdl.handle.net/1813/6478.

Stein, Howard. 1967. "Newtonian Space-Time." *Texas Quarterly* 10(3): 174–200, http://strangebeautiful.com/other-texts/stein-newtonian-spacetime.pdf.

Szczerba, Lesław W. and Tarski, Alfred. 1965. "Metamathematical Properties of Some Affine Geometries." in *Logic, Methodology and Philosophy of Science II, Proceedings of the 1964 International Congress in Jerusalem*, edited by Yehoshua Bar-Hillel, pp. 166–178. Amsterdam: North-Holland Publishing Co.

—. 1979. "Metamathematical Discussion of Some Affine Geometries." *Fundamenta Mathematicae* 104(3): 155–192. Extension and elaboration of Szczerba and Tarski (1965), doi:10.4064/fm-104-3-155-192.

Tao, Terence. 2008. "What is a Gauge?" Blog post from September 27, 2008, https://terrytao.wordpress.com/2008/ 09/27/what-is-a-gauge/.

Tarski, Alfred. 1930. "O pojęciu prawdy w odniesieniu do sformalizowanych nauk dedukcyjnych." *Ruch Filozoficzny* 12: 210a–211b, https://diglib.eodopen.eu/view/uuid:8ab4ad95-f676-3642-a531-f9e636a7040b.

—. 1933. "Pojęcie prawdy w językach nauk dedukcyjnych." *Prace Towarzystwa Naukowego Warszawskiego, Wydział III, nauk matematyccznofizycznych (Travaux de la Société des Sciences et des Lettres de Varsovie, Classe III, Sciences Mathématiques et Physiques)* 26. Lectures given at the Logic Section of the Philosophical Society in Warsaw on October 8, 1930 and at the Polish Philosophical Society in Lwów on December 5, 1930; reprinted in Zygmunt (1995, 13–172); a summary of the chief results was published as Tarski (1930) and a revised German translation was published as Tarski (1935); English translation Tarski (1956, 152–278).

—. 1935. "Der Wahrheitsbegriff in den formalisierten Sprachen." *Studia Philosophica (Commentarii Societatis Philosophicae Polonorum)* 1: 261–405. Translation of Tarski (1933).

—. 1936. *O logice matematycznej i melodzie dedukcyjnej*. Lwów: Ksiaznica-Atlas. Translated into German: Tarski (1937).

—. 1937. *Einführung in die mathematische Logik und in die Methodologie der Mathematik*. Wien: Springer Verlag. Translation of Tarski (1936), doi:10.1007/978-3-7091-5928-6.

—. 1941. *Introduction to Logic and to the Methodology of Deductive Sciences*. New York: Oxford University Press. Revised and enlarged edition of Tarski (1936), translated by Olaf Helmer; fourth edition: Tarski (1994).

—. 1946. *Introduction to Logic and to the Methodology of Deductive Sciences*. 2nd ed. New York: Oxford University Press. First edition: Tarski (1941); fourth edition: Tarski (1994).

—. 1951. "A Decision Method for Elementary Algebra and Geometry." R-109. Santa Monica, California: RAND Corporation. Prepared for publication by J.C.C. McKinsey, dated August 1, 1948, revised May 1951, https://www.rand.org/pubs/reports/R109.html.
—. 1956. *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*. 1st ed. Oxford: Oxford University Press. Trans. J.H. Woodger, 2nd edition: Tarski (1983).
—. 1959. "What is Elementary Geometry?" in *The Axiomatic Method with Special Reference to Geometry and Physics*, edited by Leon HENKIN, Patrick SUPPES, and Alfred TARSKI, pp. 16–29. Amsterdam: North-Holland Publishing Co. Reprinted in Hintikka (1969), doi:10.1016/S0049-237X(09)70017-5.
—. 1983. *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*. 2nd ed. Indianapolis, Indiana: Hackett Publishing Co. Trans. J.H. Woodger, edited and introduced by John Corcoran, 1st edition: Tarski (1956).
—. 1994. *Introduction to Logic and to the Methodology of Deductive Sciences*. 4th ed. Oxford Logic Guides n. 24. Oxford: Oxford University Press. Edited by Jan Tarski, with a revised translation.
TARSKI, Alfred and GIVANT, Steven. 1999. "Tarski's System of Geometry." *The Bulletin of Symbolic Logic* 5(2): 175–214. A letter to Wolfram Schabhäuser, written originally ca. 1978, doi:10.2307/421089.
TRAUTMAN, Andrzej. 1966. "Comparison of Newtonian and Relativistic Theories of Space-Time." in *Perspectives in Geometry and Relativity: Essays in Honor of Václav Hlavatý*, edited by Banesh HOFFMANN, pp. 413–425. Bloomington, Indiana: Indiana University Press.
VEBLEN, Oswald. 1904. "A System of Axioms for Geometry." *Transactions of the American Mathematical Society* 5(3): 343–384, doi:10.1090/S0002-9947-1904-1500678-X.
WALD, Robert M. 1984. *General Relativity*. Chicago, Illinois: University of Chicago Press.
WEINBERG, Steven. 1972. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. Hoboken, New Jersey: John Wiley and Sons, Inc.
ZYGMUNT, Jan, ed. 1995. *Alfred Tarski, Pisma Logiczno-Filozoficzne, vol. 1*. Warszawa: Państwow Wydawnictwo Naukowe (PWN).

Weyl, Gödel, and the *Grundlagenstreit*

Patrizio Contu

The present paper provides a reconstruction of Weyl's and Gödel's interpretations of intuitionism, embedding the discussion in the context of the *Grundlagenstreit* and the origins of constructive logic. The two interpretations exhibit a striking affinity and deviate substantially from the mainstream view, usually referred to as the *Brouwer-Heyting-Kolmogorov* explanation of constructive proofs. Gödel's objections to intuitionism are fairly well-known, but the connection with Weyl appears to have received little attention from commentators. The crux of the matter is the concept and role of *ideal elements* in mathematics. The paper explains how different interpretations of intuitionism deal with this problem.

The mainstream view on constructive semantics, as codified by the Brouwer-Heyting-Kolmogorov (*BHK*) interpretation of logical constants, has not always gone unchallenged. Gödel—and Weyl before him—had quite different opinions on the way constructivism is to be understood. In particular, instead of laying the blame on the law of excluded middle, they focused on the logic of quantification and posed heavy restrictions on the interaction between quantifiers and propositional connectives, particularly negation. Both authors relied heavily on choice functions. In the following, we outline Weyl's and Gödel's main ideas, comparing them with each other and with mainstream constructive logics. We also sketch some important connections with Hilbert, as well as some puzzles around intuitionistic *reductio ad absurdum*.

1 Weyl on Constructivity

In Weyl (1918), Hermann Weyl had already published a book-length attempt of his own to build analysis on a predicative basis, but in Weyl (1921), he modified his former views and adhered to Brouwer's intuitionism, although with some substantial differences. What is of interest to us in this important paper is his theory of quantification, which diverges from Brouwer's views and from what has now become mainstream in constructive logic. We shall also refer to Weyl (2009) (whose original version was published in 1928) to reconstruct Weyl's approach. His theory could be broken down into the following principles:[1]

> Actuality. An existential statement can only be asserted when an instance has been found.
>
> Abstraction. Existential quantifiers build no real statements but only *statement abstracts* (*Urteilsabstrakte*), or *abstracts* for short. Universal quantifiers build *statement instructions* (*Urteilsanweisungen*).
>
> Constrained Inference. No inferences can be drawn from existential abstracts. No *logical* inferences can lead to a statement instruction.
>
> Negation. Negation cannot be applied to statement abstracts or instructions.
>
> Quantifier Nesting. An existential quantifier cannot occur in the scope of a universal quantifier.
>
> Propositional Innocence. Classical propositional logic is not to be blamed for non-constructivity, for the latter arises from the logic of quantification over infinite domains.

The basis of Weyl's claims is undoubtedly the principle of Actuality: constructively, so the principle goes, I cannot assert an existence statement based on the mere *possibility* of having a construction that I do not *actually* possess:[2]

> Nur die *gelungene* Konstruktion kann uns die Berechtigung dazu geben; von *Möglichkeit* ist nicht die Rede. (Weyl 1921, 55, original emphasis)

From this, Weyl concludes that existential statements are no real statements because no state of affairs (*Sachverhalt*), hence no independent meaning, is attached to them without the proof construction that has already taken place. Weyl formulates this in rather colourful tones:

1 We render Weyl's term "*Urteil*" (judgement) as "statement."

2 This may be a refusal of Husserl's identification of mathematical existence with possibility; cf., e.g., Husserl (1939, 450).

> Man muß solche Dinge nicht von außen erwägen, sondern sich innerlich ganz zusammenraffen und ringen um das "Gesicht", die Evidenz. Endlich fand ich für mich das erlösende Wort. *Ein Existentialsatz*—etwa "es gibt eine gerade Zahl"—*ist überhaupt kein Urteil im eigentlichen Sinne, das einen Sachverhalt behauptet*; Existential-Sachverhalte sind eine leere Erfindung der Logiker. "2 ist eine gerade Zahl": das ist ein wirkliches, einem Sachverhalt Ausdruck gebendes Urteil; "es gibt eine gerade Zahl" ist nur ein aus diesem Urteil gewonnenes *Urteilsabstrakt*. (Weyl 1921, 54, original emphasis)

Universally quantified statements, on the other hand, are general instructions on how to build real statements. Given the general statement $m+1 = 1+m$, it can be transformed by a uniform principle into a special case, e.g., $9+1 = 1+9$:

> Auch eine allgemeine Aussage weist nicht auf einen an sich bestehenden Sachverhalt hin, sie ist nicht gemeint als logisches Produkt unendlich vieler Einzelaussagen, sondern hypothetisch: angewandt auf eine einzelne bestimmte vorliegende Zahl liefert sie ein bestimmtes Urteil. (Weyl 2009, 72–73)

Therefore, one could think of the application of statement instructions as a sort of conversion in lambda calculus:

$$(\forall x A(x))(a) \triangleright A(a).$$

Since the instruction has to hold for all objects in the domain, for its application, it is not required to exhibit a previously constructed object a, but we can simply use any denoting name. In Gentzenian terms, this justifies the elimination rule

$$\forall x A(x) \Rightarrow A(a).$$

The restriction on existential abstracts, on the other hand, clearly justifies the introduction rule:

$$A(a) \Rightarrow \exists x A(x).$$

In fact, these are the only quantifier rules that Weyl gives in Weyl (2009, 32). Such is, then, the principle of CONSTRAINED INFERENCE. Hence, the questions arise:

1. How do we draw consequences from an existential abstract?

2. How do we establish a universal statement instruction?

Weyl's answer to the first question is simply that we do not. We cannot draw conclusions from a pseudo-statement. Whenever we want to infer a conclusion from $\exists x A(x)$, we have to resort to the statement $A(a)$ that we have established and draw our inferences from that statement (cf. Weyl 2009, 72). As to the second question, Weyl seems to be saying that general sentences are always arrived at through domain-specific axioms, never by logic alone. In the case of natural numbers, for example, the method of attaining universal sentences will be mathematical induction, together with the generality provided by definitions (Weyl 2009, 72). In spite of the duality between existential and universal quantifiers, which is reflected in the above rules, Weyl does not assign the same status to abstracts and instructions. Instructions implicitly contain infinitely many real statements; hence, they are different from abstracts, which are pseudo-statements (cf. Weyl 1921, 56).

Based on the previous principles, Weyl can now limit the applicability of negation to quantified sentences. An abstract cannot be negated because it is a pseudo-statement; just like one cannot draw inferences from it, one cannot negate it either. In particular, the forbidden $\neg\exists x A(x)$ can be written legitimately as $\forall x \neg A(x)$. For Weyl, the existential quantifier is mathematically and logically idle. The extension of the same principle to the universal quantifier is prima facie puzzling. Weyl claims:

> Die Negation einer allgemeinen Aussage über Zahlen wäre ein Existentialsatz; da dieser nichtssagend ist, sind die allgemeinen Urteile nicht negationsfähig. (Weyl 2009, 72)

Here, he seems to be appealing to the law

$$\neg\forall x A(x) \Rightarrow \exists x \neg A(x),$$

which is constructively invalid. It is probably the case that Weyl has classical negation in mind, and the argument is meant to show that *this* negation does not apply (see below). Weyl concludes that the law of excluded middle fails for quantified sentences since it cannot even be formulated (cf. Weyl 1921, 56).

The next principle that we need to examine is that of quantifier nesting (cf. Weyl 1921, 57). According to Weyl, if we have proved $\forall x A(x, a)$, then we can abstract legitimately and obtain $\exists y \forall x A(x, y)$. If, on the other hand, an instruction is to be a rule that can be applied to all objects of the domain to yield a real or proper statement:

$$(\forall x A(x))(a) \rhd A(a),$$

then it is clear that we cannot have the situation in which the result is a pseudo-statement:

$$(\forall x \exists y A(x, y))(a) \rhd \exists y A(a, y),$$

whence the scope restriction: an existential quantifier cannot occur within the scope of a universal quantifier. Weyl's way out consists in interpreting $\forall x \exists y A(x, y)$ as an abstract of an instruction rather than the opposite, i.e., by using what would later come to be known as Skolem functions f such that $\exists f \forall x A(x, f(x))$. This interpretation allows a legitimate conversion:

$$(\forall x A(x, f(x)))(a) \rhd A(a, f(a)).$$

The principle of nesting reflects again the asymmetry between ∀ and ∃ sentences.

Finally, it is implicit in Weyl's analysis that the failure of constructivity is due to quantification over infinite totalities rather than propositional logic. It is only when classical negation is applied to quantifiers that constructivity is violated. But classical negation itself, within the scope of propositional logic, is by itself harmless. We call this the principle of Propositional Innocence. That this is actually at work in Weyl's conception can be seen from the fact that Weyl (2009) defines propositional connectives by means of truth tables (cf. Weyl 2009, 30). It is important to stress, however, that while all other principles are clearly stated by Weyl, the principle of Propositional Innocence is my own extrapolation and remains, therefore, hypothetical.[3]

Summarizing, for Weyl, the existential quantifier is mathematically and logically idle, whereas the universal quantifier is mathematically not idle (since statement instructions are proved by means of mathematical definitions and axioms) and logically idle to a lesser degree (as statement instructions imply infinitely many proper statements). But in both cases, the crucial quantifier rules that are subject to parameter restrictions have no place in deduction.

3 Brouwer himself stressed that the excluded middle is not problematic over finite domains; see, e.g., Brouwer (2020, 21). I am indebted to an anonymous referee for pointing this out in this context.

doi: 10.48106/dial.v77.i2

In particular, quantified sentences cannot be meaningfully negated. Weyl is silent as to the question whether other logical connectives can be applied to quantified sentences, presumably because their impact is not as crucial.

2 Hilbert's Programme

Weyl's paper had a profound influence on Hilbert and the formulation of his programme. Hilbert held Weyl in high esteem and was deeply upset by his allegiance to intuitionism. Accordingly, he took Weyl's challenge seriously, as testified by Hilbert (1922), where his words echo Weyl's arguments closely:

> Bei unendlich vielen Dingen hat die Negation des allgemeinen Urteils $\forall x A(x)$ zunächst gar keinen präzisen Inhalt, ebensowenig wie die Negation des Existentialurteils $\exists x A(x)$. Allerdings können gelegentlich diese Negationen einen Sinn erhalten, nämlich, wenn die Behauptung $\forall x A(x)$ durch ein Gegenbeispiel widerlegt wird oder wenn aus der Annahme $\forall x A(x)$ bzw. $\exists x A(x)$ ein Widerspruch abgeleitet wird. Diese Fälle sind aber nicht kontradiktorisch entgegengesetzt; denn wenn $A(x)$ nicht für alle x gilt, wissen wir noch nicht, daß ein Gegenstand mit der Eigenschaft Nicht-A wirklich vorliegt; ebensowenig dürfen wir ohne weiteres sagen: entweder gilt $\forall x A(x)$ bzw. $\exists x A(x)$ oder diese Behauptungen weisen einen Widerspruch wirklich auf. Bei endlichen Gesamtheiten sind "*es gibt*" und "*es liegt vor*" einander gleichbedeutend; bei unendlichen Gesamtheiten ist nur der letztere Begriff ohne weiteres deutlich. (Hilbert 1922, 155–156, original emphasis)

This passage shows that Hilbert had taken up many of Weyl's views and that the negation that Hilbert had in mind is classical negation. The distinction between different kinds of negation is brought to clarity in the later treatise Hilbert and Bernays (1934, 33–34), where the example is given of an elementary arithmetic statement, say $f(m) = n$, whose contradictory negation is another statement to the effect that $f(m) = k$, with $n \neq k$. Here, we have two claims on the result of a given procedure. They contradict each other exactly in the sense that they only deviate in the claimed result, but they coincide in the basic procedure. Now consider an existential statement. If we say that there is no n such that $A(n)$, we cannot mean it in the mild sense (*in unscharfem Sinne*) that such n is not available, but rather in the sense that it *cannot* have the property A. Hilbert and Bernays call this a *sharpened* negation (*verschärfte Negation*). From a finitary standpoint, the mild or unsharp negation is the exact contradictory of an existential statement because it lies at the same epistemological level as the negated statement (available/not available; exists/does not exist), whereas the sharpened negation works at an entirely different level, that is, that of general laws:

> Die Existentialaussage und ihre verschärfte Negation sind nicht, wie eine elementare Aussage und ihre Negation, Aussagen über die beiden allein in Betracht kommenden Ergebnisse *einer und derselben Entscheidung*, sondern sie entsprechen zwei getrennten Erkenntnismöglichkeiten, nämlich einerseits der Auffindung einer Ziffer von einer gegebenen Eigenschaft, andererseits der Einsicht in ein allgemeines Gesetz über Ziffern. Daß eine von diesen beiden Möglichkeiten sich bieten muß, ist nicht logisch selbstverständlich. (Hilbert and Bernays 1934, 33, original emphasis)

Thus, the lack of a contradictory negation causes the law of excluded middle to fail. The same holds true of general statements: we cannot assume, from a finitistic point of view, that either $A(x)$ is true of all x or that an x can be found that is not A (cf. Hilbert and Bernays 1934, 34).

There is an underlying agreement with Weyl on the problem of applying negation to quantified statements, which ultimately does not translate, however, in Weyl's prohibition. In fact, the doctrine of quantifiers does not obey Weyl's principle of abstraction: existential sentences do not express pseudo-statements but only statements with partial information (*Partialurteile*). On the other hand, it is clear that Hilbert and Bernays fully subscribed to the principle of Propositional Innocence, for the source of infinitary reasoning was supposed to be quantification.

As in Weyl, Hilbert's method of dealing with quantifiers is based on choice functions. His initial approach made use of the τ term-forming operator, such that $\tau_x A(x)$ is to be interpreted as "the least likely to be A."[4] The corresponding axiom is

$$A(\tau_x A(x)) \Rightarrow A(x).$$

4 This wording derives from DeVidi (2004), but the same idea is clearly explained by Hilbert himself.

doi: 10.48106/dial.v77.i2

He was soon to change this by adopting a dual operator ε, with $\varepsilon_x A(x)$ to be read as "the most likely to be A," ruled by the axiom

$$A(x) \Rightarrow A(\varepsilon_x A(x)).$$

ε behaves as a choice function, as the following example illustrates:

$$A(y_1, \dots, y_n, x) \Rightarrow A(y_1, \dots, y_n, \varepsilon_x A(y_1, \dots, y_n, x)).$$

With the ε operator, defining the quantifiers becomes easy:

$$\exists x A(x) \Leftrightarrow A(\varepsilon_x A(x)),$$

$$\forall x A(x) \Leftrightarrow A(\varepsilon_x \neg A(x)).$$

The latter amounts to

$$\forall x A(x) \Leftrightarrow A(\tau_x A(x))$$

since

$$\varepsilon_x \neg A(x) = \tau_x A(x)$$

(the least likely to be A is the most likely to be $\neg A$), from which follows

$$A(\varepsilon_x \neg A(x)) \Leftrightarrow A(\tau_x A(x)).$$

From these definitions, the full rules of quantification can be deduced (cf. Hilbert and Bernays 1939).

In summary, Hilbert and Bernays concluded that quantification, when applied to infinite domains, is devoid of clear meaning, but instead of giving up classical laws, they set out to prove that infinitary (or ideal) methods can be justified indirectly by proving the consistency of the system. The sense in which a consistency proof solves the problem is explained as follows: Given the lack of semantic transparency of infinitary mathematics, it is theoretically possible that some infinitary results be shown to be invalid by finitary methods, in analogy to the discovery of the set-theoretic antinomies. But if a consistency proof has been established, such contradiction between different methods can never take place (cf. Hilbert and Bernays 1934, 42). Thus, Hilbert ultimately rejected all of Weyl's principles apart from the principle of PROPOSITIONAL INNOCENCE.

3 Negation and Quantification

We have seen that the net effect of Weyl's approach was to adopt classical propositional logic and curtail the logic of quantification. Such a diagnosis of constructivity was very much at variance with Brouwer's focus on negation in general, not limited to quantifiers.[5] In keeping with Brouwer's conception, the essence of Heyting's formulation of intuitionistic logic was the rejection of the principle of PROPOSITIONAL INNOCENCE, which in one stroke made it possible to lift all other prohibitions imposed by Weyl.[6] With the notion of falsity as *reductio ad absurdum*, negation could be applied to quantified sentences as well, and the excluded middle failed on all sentences. In formalistic terms, this amounts to identifying the source of non-constructive methods with propositional logic, whereas the logic of quantification is now thought to be innocent and can coincide with that of classical logic. As we shall see shortly, this is not, strictly speaking, the case, for even quantifiers are interpreted differently.

The problem now was the exact meaning of negation. Just by refusing classical negation and defining $\neg A \coloneqq A \Rightarrow \bot$, we still do not have a clear answer as to what to do with absurdity $\bot$. The principle on which everyone agreed was constructive *reductio ad absurdum*: $(A \Rightarrow B) \Rightarrow ((A \Rightarrow \neg B) \Rightarrow \neg A)$, but the main point of controversy lay in the *ex falso sequitur quodlibet* law: $\bot \Rightarrow B$. Constructively, it was implicit in Brouwer's conception that in order to justify a hypothetical judgement, one has to provide at least a method that transforms the antecedent into the consequent, and in the case of *ex falso*, it was not *prima facie* clear what such a method could be. Before Heyting, Kolmogorov (1925) had defined a version of Brouwer's logic without the *ex falso* law, since he thought that this rule has no "intuitive foundation" (Kolmogorov 1925, 419). Kolmogorov also provided the first double-negation translation with classical logic, but his work did not achieve wide circulation. A logical system with the positive fragment and constructive *reductio* that does not contain the *ex falso* is now called minimal logic, after Johansson (1937). Although the semantic justification of the *ex falso* was at first unclear, ultimately, Heyting's acceptance of this principle, apparently with the tacit agreement of Brouwer, became influential. Under Heyting's suggestion, Glivenko (1929) had included the *ex falso* in his axiom system, and the final detailed version of intuitionistic

5 However, Brouwer's own comments on Weyl (1921), reported in Mancosu (1998, 119–122), do not contain any remarks on the parts of the paper devoted to logic.

6 This is not to say that Heyting devised his formulation as an explicit rejoinder to Weyl.

doi: 10.48106/dial.v77.i2

logic, also featuring *ex falso*, was published as Heyting (1930). It is telling that all these works focused on propositional logic.

Kolmogorov (1932) provided a semantic justification of *ex falso* in terms of problems and their solutions. In general, for Kolmogorov, a proposition A represents a problem, and a proof of A represents its solution. What is, then, the status of a problem that we know is unsolvable? Kolmogorov considers the following problem: under the assumption that the number π is rational, prove that also the number e can be expressed as a rational number. He remarks that

> Die Voraussetzung [...] der [...] Aufgabe [ist] unmöglich, und folglich ist die Aufgabe selbst *inhaltslos*. Der Beweis, daß eine Aufgabe inhaltslos ist, wird weiter immer als ihre Lösung betrachtet werden. (Kolmogorov 1932, 59, original emphasis)

In the case of *ex falso*, represented as $\neg A \Rightarrow (A \Rightarrow B)$, we have as a simple consequence that if we have proved the premiss, then the resulting implication is devoid of content and therefore solved:

> Sobald $\neg A$ gelöst ist, [ist] die Lösung von A unmöglich und die Aufgabe $A \Rightarrow B$ inhaltslos. (Kolmogorov 1932, 62)

Thus, we do not need a *specific construction* to prove B from $\bot$ for the simple reason that the premiss can never be solved. This view is now the established interpretation (cf. Troelstra and van Dalen 1988, 10). Observe that *ex falso* is similar to Hilbert's τ operator: if the most unlikely to be true is actually true, then anything else is true.

The somewhat non-constructive flavour of *ex falso* can be gleaned from the intuitionistic law $(\neg A \vee B) \Rightarrow (A \Rightarrow B)$, whose proof is based on *ex falso* and *a fortiori*, as can be seen from the derivation in natural deduction:

$$\dfrac{\dfrac{[\neg A \vee B] \qquad \dfrac{\dfrac{\dfrac{[\neg A] \qquad [A]}{\bot}}{B}}{A \Rightarrow B} \qquad \dfrac{[B]}{A \Rightarrow B}}{A \Rightarrow B}}{(\neg A \vee B) \Rightarrow (A \Rightarrow B).}$$

The law has a non-constructive flavour because it makes the meaning of intuitionistic implication dangerously close to that of classical logic: a sufficient condition for an implication is that either the antecedent is false or the consequent is true. Heyting, Kolmogorov, and the standard theory after them all assume that ⊥ can never be proved, and, therefore, no specific construction is needed to obtain an arbitrary proposition from it. Johansson, on the other hand, points out that we may not know whether a proof of ⊥ could be obtained; hence, in general, we could obtain one if our axioms are inconsistent (cf. Johansson 1937, 128), and then we would have the burden of providing a specific construction to prove any B from it. Heyting was aware of the somewhat problematic status of *ex falso*. In later years, referring to this law, he stressed the open-ended character of intuitionistic mathematics:

> It must be remembered that no formal system can be proved to represent adequately an intuitionistic theory. There always remains a residue of ambiguity in the interpretation of the signs, and it can never be proved with mathematical rigour that the system of axioms really embraces every valid method of proof. (Heyting 1956, 102)

However, if the previous reasoning is correct, things are much worse: it would mean that as long as a contradiction cannot be proved, we are safe with the standard justification of *ex falso* (i.e., no specific construction is needed), whereas if there is a proof of a contradiction, then we are in a situation in which we have to exhibit a specific transformation that from a contradiction proves any proposition, which we may not be able to produce. It is no escape to say that we do have the inference rule because that is precisely what we have to justify semantically. Hence, it appears that if a contradiction is produced, *ex falso* may well cease to be valid. We might call this the *paradox of absurdity*. Having *ex falso* yields an elegant mathematical symmetry between truth and falsity since both $\bot \Rightarrow A$ and $A \Rightarrow \top$ are valid for any A. But there is a price to pay in terms of conceptual justification. This appears to be an important open question for constructive semantics, but we will not discuss it further in this paper.

The role of quantification in constructive logic remains to be considered. We have seen that the original formalization of intuitionistic logic turned upon propositional operators in order to work out the rules for negation. What is, then, the role of quantification? It turns out that the crucial issue is with the existential quantifier. We have seen how Weyl rejected an elimination rule for the existential quantifier because that would break the principle of actuality of proofs. In Hilbert's terms, it would introduce ideal elements. In

doi: 10.48106/dial.v77.i2

keeping with this idea, in a classical system, one can add a rule of existential instantiation based on Hilbert's ε operator:

$$\frac{\exists x A(x)}{A(\varepsilon_x A(x)).}$$

Gentzen (1935), on the other hand, formulated the elimination rule for ∃ as

$$\frac{\begin{array}{c}\Pi_1\\ \exists x A(x)\end{array} \qquad \begin{array}{c}[A(x/a)]\\ \Pi_2\\ C\end{array}}{C}$$

(where a does not occur in C). Gentzen's rule is intuitionistically valid, whereas existential instantiation is not, as the following example demonstrates:

$$\dfrac{\dfrac{\dfrac{\dfrac{[A] \qquad A \Rightarrow \exists x B(x)}{\exists x B(x)}}{B(\varepsilon_x B(x))}}{A \Rightarrow B(\varepsilon_x B(x))}}{\exists x (A \Rightarrow B(x)).}$$

Since the only laws that we are using are those for implication and the existential quantifier, if one accepts the constructive meaning of the implication rules, it follows that the problem is due to existential instantiation. This shows that Weyl's misgivings about drawing consequences from existentially quantified statements were not unfounded. It is now well-known that if one adds an extensionality condition for the ε operator:

$$\forall x (A(x) \Leftrightarrow B(x)) \Rightarrow \varepsilon_x A(x) = \varepsilon_x B(x),$$

even the Law of Excluded Middle can be derived (this was first proved by Diaconescu in the context of topos theory). Thus, Weyl's use of choice functions defined on objects in a suitable domain, as in $\exists f \forall x A(x, f(x))$, cannot be extended to choice functions defined on the property $A(x)$ itself, as in $\varepsilon_x A(x)$, without overstepping the bounds of constructivism, as Weyl rightly saw.

The above example also shows that the quantifier rules are crucial in characterising constructivism.[7]

7 The standard explanation of constructive proofs, known as the *Brouwer-Heyting-Kolmogorov* (*BHK*) interpretation, provides a non-classical account for all logical constants; hence, it does

4 Gödel's Functional Interpretation

There are striking similarities between Weyl's analysis of constructivity and Gödel's critique of intuitionistic logic in the 1930s. Whether Gödel was acquainted with Weyl's papers, however, is not clear, at least to this author. Whereas Weyl wrote on these topics before the formulation of the *BHK*, Gödel had that and the formalization of intuitionistic logic before his eyes. Here, we focus on Gödel's early discussions rather than the published version of his system (Gödel 1958). In Gödel (1933, 51–53) and Gödel (1938, 90), he states the following principles:

> Finite Generation. The universal quantifier can only be applied to totalities whose elements are finitely generated (e.g., natural or rational numbers).

> Existential Dispensability. Existential statements are mere abbreviations of statements, including a witness of the proved property; otherwise, they are dispensable. Therefore, the existential quantifier should not be a primitive symbol.

> Quantifier Scope. Negation must not be applied to universal statements because that would require a dependency on existential statements, which are dispensable. The only admissible meaning

---

not rely on propositional logic alone. While not a formal definition, this interpretation remains the conceptual reference for mainstream constructivism. We report its clauses for the reader's convenience:

> BHK (∧). π proves $A \wedge B$ iff $\pi = \langle \pi_1, \pi_2 \rangle$, where π_1 proves A and π_2 proves B.

> BHK (∨). π proves $A \vee B$ iff π proves A or π proves B.

> BHK (⇒). π proves $A \Rightarrow B$ iff π is an effective function (construction) $\lambda x.\phi(x)$ such that for each proof ρ of A, $\phi(\rho)$ proves B.

> BHK (∃). π proves $(\exists x \in D)A(x)$ iff $\pi = \langle a \in D, \rho \rangle$, where ρ proves $A(a)$.

> BHK (∀). π proves $(\forall x \in D)A(x)$ iff π is an effective function (construction) $\lambda x.\phi(x)$ such that for each $a \in D$, $\phi(a)$ proves $A(a)$.

> BHK (¬). π proves $\neg A$ iff π is an effective function (construction) $\lambda x.\phi(x)$ such that for each proof ρ of A, $\phi(\rho)$ proves $\bot$, where $\bot$ is a propositional constant of which nothing constitutes a proof.

doi: 10.48106/dial.v77.i2

of universal negation is the availability of a counterexample. Gödel (1938, 90) extends the negation restriction to all propositional connectives.

Constrained Inference. Existential statements are only governed by the introduction rule. Universal statements cannot be proved by logical means.[8]

Decidability. Only decidable relations and computable functions are allowed constructively.

The principles of Existential Dispensability, Quantifier Scope, and Constrained Inference are obviously very close to Weyl's analysis. There is no counterpart of the principle of Finite Generation in Weyl, whereas the principle of Decidability restricts the principle of Propositional Innocence: we can only apply classical inferences because we are using decidable predicates. From Gödel's principle of Quantifier Scope, it also follows that the definition of negation as *reductio ad absurdum* is not generally admissible because it allows us to deny a universal statement even in the absence of a counterexample. Furthermore, from the principle of Finite Generation, it follows that the *BHK* definition of implication, and hence also of negation, is not admissible because it quantifies over all proofs of the antecedent, and constructive proofs are not a well-defined domain of quantification in the sense of being finitely generated (cf. Gödel 1933, 52–53).

The remark on *reductio ad absurdum* is deeply ingrained in Gödel's analysis of intuitionism. Gödel (1932) extended a result of Glivenko, showing that classical arithmetic can be embedded into Heyting arithmetic through a suitable translation. This result demonstrates that intuitionistic arithmetic, contrary to expectations, is more general than classical arithmetic. Gödel explains this result thus:

> Der Grund dafür liegt darin, daß das intuitionistische Verbot, Allsätze zu negieren und reine Existentialsätze auszusprechen,

8 I quote:

> It follows that we are left with essentially only one method for proving general propositions, namely, complete induction applied to the generating process of our elements. (Gödel 1933, 51)

> in seiner Wirkung dadurch wieder aufgehoben wird, daß das Prädikat der Absurdität auf Allsätze angewendet werden kann, was zu formal den gleichen Sätzen führt, wie sie in der klassischen Mathematik behauptet werden. Wirkliche Einschränkungen scheint der Intuitionismus erst für die Analysis und Mengenlehre zu bringen, doch sind diese nicht durch Ablehnung des Tertium non datur, sondern der imprädikativen Begriffsbildungen bedingt. (Gödel 1932, 294)

The same argument is reiterated in Gödel (1941, 190). Gödel (1941) builds upon the above principles to provide a positive account of the strengthened constructivism that he envisioned and that was not satisfied by Heyting's theory. The official formulation of this account was to be Gödel (1958), where Gödel dropped the foundational discussion of intuitionism and focused on the proof of relative consistency.

Gödel's approach consists in defining a system **T** extending recursive number theory by admitting computable functionals of finite type, i.e., typed functionals such as

$$F^{\sigma\to\tau}(f^{\sigma}) = g^{\tau}.$$

On the logical side, the first consequence of the principle of indispensability and QUANTIFIER SCOPE is that sentences can only be in prenex form and with universal quantifiers only (cf. Gödel 1941, 192). Existential quantifiers are accepted as abbreviations, only governed by the introduction rule and therefore eliminable:

> An existential assertion can only appear as the last formula of a proof and the last but one formula of the proof must give the corresponding construction. (Gödel 1941, 193)

Gödel remarks that this is not an explicit definition of the existential quantifier but

> a definition of use, which states how propositions containing the new symbol are to be handled in proofs, i.e., from which premises they can be inferred, namely these [premises of the introduction rule], and what can be inferred from them, namely nothing. Now such an implicit definition must satisfy the requirement of eliminability. To be more exact: If a proposition not containing the new symbol can be proved with the help of the new symbol, it

> must be demonstrable without the help of the new symbol (otherwise we would not have to do with a definition but with a new axiom). But this requirement is trivially satisfied by this manner of introducing the existential quantifier. (Gödel 1941, 193)

Apart from the total overlap with Weyl's conception of the existential quantifier, we can observe that Gödel is stressing an important point here, that is, a definition of use based on the introduction rule alone makes the defined operator trivially eliminable.

From the previous discussion, it follows that the general form of a statement is $\exists \mathbf{x} \forall \mathbf{y} A(\mathbf{x}, \mathbf{y})$, with $A(\mathbf{x}, \mathbf{y})$ quantifier-free and where $\mathbf{x}$ and $\mathbf{y}$ are sequences of individual or functional variables. The situation is strongly reminiscent of Weyl's principle of quantifier nesting. As for Weyl, the strategy for obtaining sentences in the desired form consists in the use of choice functions. For example, $\forall z A(z)$, with $A(z) \equiv \exists x \forall y B(x, y, z)$, is interpreted in $\mathbf{T}$ as $\exists f \forall z \forall y B(f(z), y, z)$. The implication of $\mathbf{T}$-formulas $\exists x \forall y A(x, y)$ and $\exists u \forall v B(u, v)$ is not a formula of $\mathbf{T}$:

$$\exists x \forall y A(x, y) \Rightarrow \exists u \forall v B(u, v), \tag{1}$$

but, according to Gödel's analysis, it can be so transformed by first observing that given an x as in the antecedent, a u as in the consequent can be constructed, and such correlation should be given by a computable function f:

$$\exists f \forall x (\forall y A(x, y) \Rightarrow \forall v B(f(x), v)). \tag{2}$$

Furthermore, the implication within brackets can be interpreted as saying that a counterexample of the consequent implies a counterexample of the antecedent, which, by functional dependence, becomes:

$$\exists g \forall v (\neg B(f(x), v) \Rightarrow \neg A(x, g(v))). \tag{3}$$

Now, the internal implication is decidable because it is quantifier-free and all relations are decidable; hence, we can apply the classical contrapositive to obtain

$$\exists g \forall v (A(x, g(v)) \Rightarrow B(f(x), v)), \tag{4}$$

and by reintroducing the external quantifiers

$$\exists f \forall x \exists g \forall v (A(x, g(v)) \Rightarrow B(f(x), v)), \tag{5}$$

we only need to apply the axiom of choice again to obtain the final form:

$$\exists f \exists g \forall x \forall v (A(x, g(x, v)) \Rightarrow B(f(x), v)). \tag{6}$$

This rather laborious process of application of choice functions, when extended to all operators,[9] allows Gödel to prove his fundamental result: if a sentence is provable in Heyting arithmetic, then it is provable in **T**. With this result, Gödel is able to conclude that the sense in which intuitionistic logic, as applied to number theory, is constructive, consists in the fact that any provable existential statement of intuitionistic number theory is translatable into a provable existential statement of **T** for which, by construction, a witness t is readily available (cf. Gödel 1941, 199). Gödel was confident that his approach could be extended to other branches of constructive mathematics:

> If you apply intuitionistic logic in any branch of mathematics you can reduce it to a finitistic system of this kind under the sole hypothesis that the primitive functions and primitive relations of this branch of mathematics are calculable, respectively, decidable. [...] This finitistic system [...] is always obtained by introducing functions of higher types analogous to these, with the only difference that the individuals upon which the hierarchy of functions is built up are no longer the integers but the primitive objects of the branch of mathematics under consideration. (Gödel 1941, 195–196)

Summarizing, Gödel saw that Heyting's formalization of intuitionistic logic and mathematics contained some *prima facie* non-constructive methods of proof, not unlike those that we identified in the previous section, giving the possibility of proving an existential statement without a constructed witness, e.g., by applying the elimination rule for the existential quantifier, or the *reductio ad absurdum*. The significance of his result, as Gödel himself remarked, is that at least as far as number theory is concerned, intuitionistic logic is constructively sound because a witness can always be recovered. Gödel's system **T** could perhaps be viewed as a particular implementation of Weyl's analysis of constructivity, which is not to say, of course, that Weyl would have agreed with the details of Gödel's approach.

9 Gödel can define negation as *reductio ad absurdum*: $\neg A := A \Rightarrow \bot$ because now negation is never applied to quantifiers.

5 Conclusions

One crucial problem for constructivism consists in being able to provide a witness for the proof of an existential statement. Weyl and Gödel intended to address this problem by curtailing the deductive strength of the existential quantifier and, more specifically, by forsaking the elimination rule. This is because statements derivable by the elimination rule deviate from the requirement of *actuality* of constructions, being mere possibilities and thereby introducing a potentially non-constructive ("ideal") element in inference.

In our discussion, we have compared three main positions:

1. Intuitionism of Brouwer and Heyting: a constructive proof is not constrained by the availability of actual witnesses, for it suffices to be able, *in principle*, to compute them. The proof-theoretic systems introduced by Gentzen follow this paradigm, in which all of the deductive rules corresponding to Heyting's logic are admissible because witnesses can be obtained by cut-elimination or normalization.
2. Hilbert's finitist standpoint: ideal elements, including those posited by classical mathematics, are harmless as long as they can be justified by a finitary consistency proof. This can be understood at least partly as an attempt to meet the challenge of intuitionism without rejecting classical logic.
3. Strict constructivism, as defined by Weyl and Gödel: a truly constructive proof cannot include any ideal elements, and the notion of proof itself should follow the same standards. Prima facie, similar restrictions have a potential for reducing the deductive power of constructive theories, but, in fact, Gödel's approach is only partially revisionistic: his view is that while his variety of constructivism can be more restrictive in general, when confined to a theory built along the lines of **T** for arithmetic (i.e., based exclusively on computable functions and decidable predicates), the full power of Heyting's logic can be recovered (see Gödel 1941, 195–196).

There is now one looming question: When it comes to ideal elements, how strict can constructivism be? In particular, can Gödel claim to have succeeded in providing a firmer conceptual foundation for constructivism?

One crucial problem is whether Gödel's computable functionals can really be conceptualized without breaking the principle of Quantifier Scope, as formulated by Weyl and himself,[10] since functionals, like any function, require a $\forall\exists$ condition: f is a function such that for each argument, it computes a value, or $\forall x \exists y (f(x) = y)$. But if we transform that into its Skolem form, the result is not quite explanatory: $\exists g \forall x (f(x) = g(x))$. That is, f is a function that behaves exactly like some other function g. One can perhaps say that there is no need for such an explanation *within* **T**, but only for our understanding of **T** in the metatheory. This, however, would make Gödel's conceptual reform of constructivism much less convincing, for it would rely on an implicit grasp of **T** that does not follow **T**'s own principles.

More generally, are higher-order concepts such as functionals to be classified as ideal elements? A related objection has been levelled by Tait (2006). According to Tait, Gödel's replacement of proofs by computable functionals is unwarranted, on the grounds that determining that a functional is computable may involve resources of arbitrary complexity and, in general, all of Heyting arithmetic. The rationale behind Tait's main argument is that constructivity should be defined in terms of methods of reasoning rather than on the assumption of computability and decidability. Gödel's later view appears to have been that, ultimately, both the finitist standpoint and constructive logic are forced to include ideal elements and drop the assumption that proof constructions must be intuitively given spatiotemporal arrangements (see Gödel 1958, 244).

In sum, the question of the conceptual semantics of constructivism remains open, as the discussion of *ex falso* also illustrates. Standard constructivism attempts to strike a balance between the ideality of classical logic and the finitist quest for actuality, and while a fully satisfactory balance seems to be hard to attain, the foundational research conducted along the way has provided a rich account of the logical phenomena involved.*

Patrizio Contu
0000-0003-1332-228X
ADD INSTITUTE
patrizio.contu@mathesis.ch

10 I am indebted to an anonymous referee for raising this type of difficulty.

* My gratitude goes to an anonymous referee for useful comments and suggestions.

doi: 10.48106/dial.v77.i2

References

BROUWER, Luitzen Egbertus Jan. 2020. *Intuitionismus*. 2nd ed. Mathematik im Kontext. Berlin: Springer. Edited by Dirk van Dalen and David E. Rowe, doi:10.1007/978-3-662-61389-4.

DEVIDI, David. 2004. "Choice Principles and Constructive Logics." *Philosophia Mathematica* 12(3): 222–243, doi:10.1093/philmat/12.3.222.

GENTZEN, Gerhard. 1935. "Untersuchungen über das logische Schließen." *Mathematische Zeitschrift* 39: 176–210, 405–431. Republished as Gentzen (1969), doi:10.1007/BF01201353.

—. 1969. *Untersuchungen über das logische Schliessen*. Darmstadt: Wissenschaftliche Buchgesellschaft.

GLIVENKO, Valery Ivanovich. 1929. "Sur quelques points de la logique de M. Brouwer." *Académie Royale de Belgique, Bulletin de la Classe des sciences* 15(5): 183–188.

GÖDEL, Kurt. 1932. "Zur intuitionistischen Arithmetik und Zahlentheorie." *Ergebnisse eines mathematischen Kolloquiums* 4: 34–38. Reprinted and translated in Gödel (1986, 286–295), reprinted, with the original page numbers, in Menger (1998, 216–220).

—. 1933. "The Present Situation in the Foundations of Mathematics." in *Collected Works, Volume III: Unpublished Essays and Lectures*, pp. 45–53. Oxford: Oxford University Press. Edited by Solomon Feferman, John W. Dawson, Jr., Warren Goldfarb, Charles Parsons and Robert M. Solovay, doi:10.1093/oso/9780195072556.001.0001.

—. 1938. "Vortrag bei Zilsel." in *Collected Works, Volume III: Unpublished Essays and Lectures*, pp. 86–113. Oxford: Oxford University Press. Translation by Charles Parsons, doi:10.1093/oso/9780195072556.001.0001.

—. 1941. "In What Sense Is Intuitionistic Logic Constructive?" in *Collected Works, Volume III: Unpublished Essays and Lectures*, pp. 189–201. Oxford: Oxford University Press. Edited by Solomon Feferman, John W. Dawson, Jr., Warren Goldfarb, Charles Parsons and Robert M. Solovay, doi:10.1093/oso/9780195072556.001.0001.

—. 1958. "Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes." *Dialectica* 12(3/4): 280–287. English translation published as Gödel (1980); reprinted in Gödel (1990, 240–250), doi:10.1111/j.1746-8361.1958.tb01464.x.

—. 1980. "On a Hitherto Unexploited Extension of the Finitary Standpoint." *Journal of Philosophical Logic* 9(2): 133–142. Reprinted in Gödel (1990, 241–250), doi:10.1007/bf00247744.

—. 1986. *Collected Works, Volume I: Publications 1929–1936*. Oxford: Oxford University Press. Edited by Solomon Feferman, John W. Dawson, Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay and Jean van Heijenoort, doi:10.1093/oso/9780195147209.001.0001.

—. 1990. *Collected Works, Volume II: Publications 1938–1974*. Oxford: Oxford University Press. Edited by Solomon Feferman, John W. Dawson, Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay and Jean van Heijenoort, doi:10.1093/oso/9780195147216.001.0001.

van Heijenoort, Jan, ed. 1967. *From Frege to Gödel: A Source Book in Mathematical Logic 1879–1931*. Cambridge, Massachusetts: Harvard University Press.

Heyting, Arend. 1930. "Die formalen Regeln der intuitionistischen Logik I." *Sitzungsberichte der Preussischen Akademie der Wissenschaften* 2: 42–56. English translation in Mancosu (1998, 311–327).

—. 1956. *Intuitionism: An Introduction*. Studies in Logic and the Foundations of Mathematics n. 17. Amsterdam: North-Holland Publishing Co. Third edition: Heyting (1971).

—. 1971. *Intuitionism: An Introduction*. 3rd ed. Amsterdam: North-Holland Publishing Co. First edition: Heyting (1956).

Hilbert, David. 1922. "Die logischen Grundlagen der Mathematik." *Mathematische Annalen* 88(1/2): 151–165, doi:10.1007/BF01448445.

Hilbert, David and Bernays, Paul. 1934. *Die Grundlagen der Mathematik, Band I*. Berlin: Springer Verlag. Second edition: Hilbert and Bernays (1968).

—. 1939. *Die Grundlagen der Mathematik, Band II*. Berlin: Springer Verlag, doi:10.1007/978-3-642-86896-2.

—. 1968. *Die Grundlagen der Mathematik, Band I*. 2nd ed. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen n. 40. Berlin: Springer Verlag. First edition: Hilbert and Bernays (1934), doi:10.1007/978-3-642-86894-8.

Husserl, Edmund. 1939. *Erfahrung und Urteil, Untersuchungen zur Genealogie der Logik*. Praha: Academia.

Johansson, Ingebrigt. 1937. "Der Minimalkalkül, ein reduzierter intuitionistischer Formalismus." *Compositio Mathematica* 4: 119–136, http://eudml.org/doc/88648.

Kolmogorov, Andrey Nikolaevich. 1925. "O principe tertium non datur." *Matematiceskoj Sbornik* 32(4): 646–667. Translated by Jan van Heijenoort as "On the Principle of Excluded Middle" in van Heijenoort (1967, 416–437).

—. 1932. "Zur Deutung der intuitionistischen Logik." *Mathematische Zeitschrift* 35: 58–65, doi:10.1007/BF01186549.

Mancosu, Paolo, ed. 1998. *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s*. Oxford: Oxford University Press, ark:/13960/s217mj22r58.

Menger, Karl, ed. 1998. *Ergebnisse eines mathematischen Kolloquiums*. Wien: Springer. Herausgegeben von Egbert Dierker, Karl Sigmund, mit Beiträgen von John W. Dawson Jr., Ryszard Engelking und Werner Hildenbrand, Geleitwort von Gérard Debreu, Nachwort von Franz Alt, doi:10.1007/978-3-7091-6470-9.

Tait, William Walker. 2006. "Gödel's Interpretation of Intuitionism." *Philosophia Mathematica* 14(2): 208–228, doi:10.1093/philmat/nkj004.

TROELSTRA, Anne Sjerp and VAN DALEN, Dirk. 1988. *Constructivism in Mathematics: An Introduction. Volume I.* Studies in Logic and the Foundations of Mathematics n. 121. Amsterdam: North-Holland Publishing Co.

WEYL, Hermann. 1918. *Das Kontinuum: Kritische Untersuchungen über die Grundlagen der Analysis*. Leipzig: Veit, doi:10.1515/9783112451144.

—. 1921. "Über die neue Grundlagenkrise der Mathematik." *Mathematische Zeitschrift* 10(1/2): 39–79, doi:10.1007/BF02102305.

—. 2009. *The Philosophy of Mathematics and Natural Science*. Princeton, New Jersey: Princeton University Press. Translated by Olaf Helmer, with a new intrduction by Frank Wilczek, doi:10.1515/9781400833337.

Should We Hope Apparent Atrocities Are Illusory?

Exploring a Puzzle in Moral Axiology

Jimmy Alfonso Licon

Philosophers have recently turned to axiological questions related to God and, to a minor extent, related to morality. This paper contributes to the latter project. The world contains atrocities such as famine and war. Can we rationally hope that these atrocities are merely moral illusions? First, we have good reason to hope that moral atrocities are only apparent because our world would be morally worse if they were real. Some critics argue that they know atrocities are real. However, setting aside whether we have such moral knowledge, perhaps we shouldn't hope that atrocities are morally illusory because that outcome would undercut our moral reliability, imply that we have false and unjustified moral beliefs, result in moral opportunity costs, and potentially deny the dignity of victims of (even only apparent) atrocities.

> Hope? Let me tell you something, my friend. Hope is a dangerous thing. Hope can drive a man insane. It's got no use on the inside. You'd better get used to that idea.
>
> —Red, *The Shawshank Redemption* (1994)

Perhaps the strongest argument against the existence of the traditional conception of God is that the world contains ubiquitous, apparently gratuitous suffering, such as genocides, war, famine, and so forth. Many philosophers have questioned how a powerful, perfectly loving God could allow such suffering, especially when it affects seemingly innocent victims. However, the point here is not to emphasize the atheological implications of the terrible atrocities in the world but rather to draw attention to them in order to pose an axiological question: Should we, and can we rationally, hope that moral atrocities are merely apparent? This question is axiological in nature and includes issues like what would make the world objectively better and what we should and can rationally hope, morally speaking. That is the focus of this paper.

Before we start with that question, though, we should step back to look at the recent emergence of interest in axiological issues across philosophical subdomains. The most prominent example in the literature relates to the question of whether the world would be better objectively if God existed or not. To quickly cover some ground: some philosophers argue that the world would be objectively better for some persons if God exists (Penner and Lougheed 2015), others argue that the world would be objectively worse for some persons if God exists (Kahane 2011; Lougheed 2017), and still others argue the world would be better irrespective of persons if God exists, e.g., the world would be intrinsically better (Davison 2018; Plantinga 2004). Some philosophers have asked, not only whether the world would be objectively better if God existed, but whether we should *hope* He does (Licon 2021). And finally, Kahane (2012) offers broad factors to think better about the value of contrasting viewpoints in metaphysics.

Recent interest in axiology, though focused on God's existence, doesn't end there. Philosophers have, recently and to a lesser extent, explored axiological questions in the moral domain, too: one philosopher argues the world would be better if moral realism is true instead of rival views in metaethics (Blanchard 2020). And in applied ethics, Hendricks (2021) argues that it would be better if the pro-choice position on the ethics of abortion is correct, as it would mean that the world is morally a better place, *ceteris paribus*, than if the pro-lifers are right about the issue.

This paper adds to recent work by philosophers on questions in moral axiology by examining whether we should, and can rationally, hope that moral atrocities are merely apparent. And we should emphasize, though, an important caveat in this paper: *the author does not take a position as to what to think about the moral axiology of atrocities*. The paper's aim is to explore different reasons on opposing sides of the issue. Our thesis is to explore whether hope against the actuality of moral atrocities is rational by weighing the pros and cons.

The plan of the paper is simple. We begin by examining the nature of rational hope to see under what conditions we can rationally hope for something. From there, we investigate how we can rationally hope for something—that moral atrocities are merely apparent—must necessarily be either true or false. Then, we explore three *possible* scenarios in which moral atrocities would be merely apparent for different reasons. And we conclude by weighing reasons, pro and con, to think we should, and can rationally, hope that moral atrocities are merely apparent. As we shall see, there is a compelling moral reason to hope and several reasons to hope not.

1 Preliminary Issues

The puzzle examined in this paper is whether we should, and can rationally, hope that moral atrocities are merely apparent. This may strike many readers as an odd question, considering the seemingly strong evidence we have that moral atrocities are real. However, we should set aside the issue of whether we believe moral atrocities exist—just like we would do when discussing whether it would be better if God existed—but instead, explore whether the world would be objectively and morally better if moral atrocities were only apparent. Additionally, we consider whether we can and should rationally hope for this to be the case. We refer to the position that we can rationally hope for moral atrocities to be merely apparent as *aspirational illusionism*.

One compelling reason that strongly supports aspirational illusionism is that, if true, it would make the world objectively and morally superior compared to a world where apparent moral atrocities were real, all other things being equal. However, there are also significant reasons that oppose aspirational illusionism. Before delving into the arguments for and against aspirational illusionism, we must address an initial objection to the possibility of moral axiology.

1.1 *An Initial Objection*

Metaethicists widely maintain that certain moral truths are metaphysically and logically necessary. If this is the case, then it would appear that we cannot make axiological comparisons between possible worlds where there are only apparent moral atrocities that are not actual and possible worlds where apparent moral atrocities are indeed real. This limitation arises because some moral propositions are necessarily true or false. Consequently, there are no possible worlds available to serve as a basis for these comparisons. There is no nearest possible world in which the truth of the proposition "Genocide is a moral atrocity" differs from the actual world (Braddock 2017).

How do we address this challenge? One plausible suggestion is to treat aspirational illusionism and its negation as live epistemic possibilities rather

doi: 10.48106/dial.v77.i2

than metaphysical ones. This shift is acceptable since we only require epistemic possibility for rational hope. While this approach may be debatable, it is unclear why we cannot evaluate rational hopes by comparing live epistemic possibilities. After all, some philosophers argue that we can rationally hope for epistemically opaque matters in the past, even if they have been settled as factual (Martin 2014, 68). For instance, Sam can rationally hope that he aced the final exam, even though it has already been graded. Presumably, Sam can have this rational hope because it remains a genuine epistemic possibility for him, even if, as a matter of fact, he did not pass the exam. Therefore, if rational hopes about (even though settled but unknown) past matters of fact can be rational when grounded in live epistemic possibilities, the mere fact that some moral claims are metaphysically necessarily true or false is not sufficient reason to reject axiological evaluations of moral issues. Consequently, we only need to rely on live epistemic possibilities for rational hope, and we will explore that further in the following section.

1.2 *The Nature of Rational Hope*

There are a few aspects to assessing the rationality of the hope aspect of aspirational illusionism. For our purposes, an agent, S, has a rational hope that *p* if, where the evidence and knowledge are concerned, *p* is a real epistemic possibility, S lacks adequate justification to believe that *p* with epistemic certainty, and S desires that *p* (Martin 2014; Meirav 2009; Pojman 1986, 161–163). The epistemic domain of hope, though, doesn't include only the future since one can rationally entertain hopes about the past to the extent one's knowledge of the past is incomplete. Rational hope can involve past events, given that those events are epistemically opaque (Martin 2014, 68; Benton 2020, 2021). However, we cannot have rational hope for past events, where we have adequate knowledge of them, since the past is fixed (Smith 1997); e.g., Mary cannot rationally hope that John was faithful if she knows that he cheated.

Here, we face a preliminary worry: there cannot be a live epistemic possibility that something is false if we *know* that it is true, e.g., if Sammy *knows* that it is eight o'clock, then it cannot be a live epistemic possibility for Sammy that it is seven o'clock. We thus face an obstacle to the mere possibility of a moral axiology puzzle: it looks prima facie like it cannot be that moral atrocities are merely apparent *if we know* that there are actual moral atrocities. We will address this issue later on. First, though, we must explain how it could be possible that moral atrocities are merely apparent given the striking moral appearances we have to the contrary.

2 A Few Possible Scenarios

Suppose we can base axiological comparisons on live epistemic possibilities. Even in that case, we still require an explanation for how there could be apparent moral atrocities that are not actual. Given the presence of moral atrocities in the world, we need an account that elucidates the misleading appearances and provides an explanation for how, at least for some epistemic agents, there could exist a live epistemic possibility that these moral appearances are false. It is important to acknowledge that what constitutes a live epistemic possibility for one person may be considered a dead epistemic possibility for another. Evidence and perspectives differ among individuals, leading to varying epistemic possibilities. Let us assume that there are moral facts in the world and actual moral atrocities. While Beth believes in these claims, Sammy harbors doubts about the existence of metaphysical entities, such as moral facts. For Sammy, it remains a live epistemic possibility that moral atrocities are illusory.

How could moral atrocities be only apparent, even for some epistemic agents? Suppose there exists an omnipotent, omniscient, and perfectly benevolent creator of the universe. Many theists already hold that, while there are moral atrocities (in a sense), God allows them to happen as it either (*a*) prevents greater moral atrocities from occurring or (*b*) facilitates something morally good that couldn't be without the atrocity (Licon 2021, 292). By His nature, God wouldn't allow gratuitous suffering to happen since He "would prevent the occurrence of any intense suffering He could, unless He could not do so *without thereby leaving things worse off* than they otherwise would be" (Howard-Snyder and Howard-Snyder 1999, 117—emphasis mine).

There is a perspective within theism that suggests there are no moral atrocities in a strict sense—morally horrendous events that are morally gratuitous. This viewpoint does not deny that people suffer and die due to events such as famine and war but rather emphasizes that these events are morally balanced. According to this perspective, God permits these events because there are moral factors that morally counterbalance the suffering and the inherent badness of apparent moral atrocities. In contrast, in a world where apparent moral atrocities lack sufficient moral factors to offset them, we would find them morally atrocious in their gratuitousness rather than morally balanced as in the theistic framework.

doi: 10.48106/dial.v77.i2

So, there's one sense here where there aren't moral atrocities to the extent that moral atrocities are morally bad events that aren't morally offset by greater goods, i.e., God is morally justified in allowing them. On this scenario, we don't deny that the suffering and death associated with moral atrocities are real. But instead, we highlight the possibility that this suffering is morally offset by a greater good and, so, isn't an atrocity *overall*. We can imagine something similar, even if more radical, if a view like moral nihilism holds: it isn't that there aren't events that happen that we would call *atrocities*, like war and famine, but instead that there are no moral properties in the world that would make them morally wrong or bad.

Perhaps the reader isn't theologically inclined. There is another, distinct metaphysical scenario where, by some fluke, it just so happens that, contrary to our best moral evidence, e.g., robust moral intuitions, apparent atrocities aren't morally bad, unjust, or immoral. On this view, the mere fact that a recent war was punctuated by horrific events, like genocide, isn't morally good or bad but morally neutral instead. This metaphysical scenario lacks a good explanation, unlike with the theistic scenario, to explain why it is that apparent moral atrocities are merely illusory.

We could even imagine a further scenario that could motivate our moral axiological puzzle: the world would morally be a better place if we lived in a simulation, and those individuals who appear to suffer an atrocity are, in fact, simulants lacking moral standing than the world would be if they had moral standing (Bostrom 2003; Chalmers 2010; Crummett 2021). In this scenario, apparent atrocities wouldn't be morally bad or wrong since they only happened to individuals lacking moral standing (e.g., perhaps the early hosts in the fictional world *Westworld*).

Many readers will likely consider these scenarios highly unlikely. Despite this, there is a non-negligible possibility that one of these scenarios holds for some of us, given what we know and believe about moral matters. One of the major reasons to canvass these scenarios is to consider different ways apparent atrocities might be illusory: it could be that they aren't atrocities *overall* (theistic scenario); as a brute fact, there are only merely apparent atrocities (metaphysical fluke scenario); and it could be that apparent atrocities only (or mostly) happen to individuals who lack moral standing because they are primitive simulants, so there was no one actually harmed by them (simulation scenario).

3 A Major Reason to Hope

Suppose that despite compelling moral appearances, the world doesn't contain moral atrocities, or even that it contains many fewer moral atrocities than it appears. That is, despite how things appear, there aren't any, or at least far fewer, *actual* atrocities like famine, war, and slavery. And putting aside how clearly counterintuitive this claim is, it should be clear that the world would be a better place were the claim true. It would mean that, despite our moral appearances, there are far fewer morally horrendous events that have happened than would first appear, and thus that the world is a morally better place than it would appear. There would be less injustice, depravity, and so on than had moral appearances been veridical. This point assumes, obviously, that if the world is less unjust or morally bad, *ceteris paribus*, then the world is a morally better place than it would otherwise be had the moral atrocities been actual. And, for our purposes, that assumption looks entirely reasonable.

An illustration would be helpful. Start with theism: many theists hold that if God exists, then the suffering we observe isn't gratuitous—even if they may not agree on the reason why it isn't gratuitous, theists agree that, somehow, God allows the suffering to happen for a good reason, either because allowing it is necessary, with respect to God, to prevent greater suffering from happening or to yield a greater good. Suppose we think that a genocide is gratuitous suffering, but God exists, and He has allowed the genocide for moral reasons that are beyond our ability to understand. This situation would be morally better, *ceteris paribus*, than had the same genocide occurred without sufficient reasons to morally offset it.

This doesn't mean that genocide isn't bad—of course it is, hence the need for offsetting moral reasons—but that the world would be a better place than it would otherwise be if there were sufficient moral reasons to allow the genocide than if the genocide occurred in the absence of such reasons. Some philosophers have argued this is a good reason to hope theism is true (Licon 2021). A similar point holds of moral atrocities: the world would be morally and objectively better, *ceteris paribus*, if moral atrocities were merely apparent—the result would be less injustice and gratuitous suffering in the world than there would be otherwise.

We can reasonably assume that suffering and death from war, disease, and famine are morally bad to the extent that they aren't morally offset, i.e., they aren't necessary to produce a greater good or prevent a greater evil. If moral atrocities are morally illusory, the world would morally be a better place,

doi: 10.48106/dial.v77.i2

ceteris paribus, than if they were actual. So, we have strong reason to think that the world would be a better place if moral atrocities were illusory.

However, we may question why we should *hope* that moral atrocities are illusory, even granting that the world would be a better place if they were. The connection between the world being a better place if something is true and hoping that it is true is fuzzy, e.g., even if the world would be a better place if atrocities were morally illusory, it might be we still cannot rationally hope that they are since we know otherwise. Here, though, we do have a strong intuition that there is a defeasible connection between them. We can state that intuition as follows:

> ABP. If S has solid reason to believe that *q* would make the world morally better than not-*q*, and *q* is a live epistemic possibility for S, then S defeasibly[1] can and should hope that *q*.

How does (ABP) work? We know knowledge and maximal credence undercut hope: to know that *p* is to foreclose the rational hope that *not-p*. For example, we cannot rationally hope we went to the best high school if we know that we attended the worst. We cannot rationally hope that *p* without reasonable belief that the truth of *p* would make the world objectively better than if *p* were false. Broadly speaking, there are two aspects to aspirational questions:

(1) Would the world be better if X is true?

and

(2) Should we hope X is true?

While there is often a strong, defeasible connection between something making the world better and hoping it is true, there will be cases where the truth of something would make the world a morally better place than if that something was false, but where, for whatever reason, we cannot rationally hope that that something was true. We discuss reasons for that sort next.

1 There may be cases where it is epistemically and axiologically permissible to hope that *p*, but where there are other, stronger reasons to hope that not-*p*, e.g., it would be morally icky to hope that *p*. The nature of the defeasibility operating in this bridging is part of the (indirect) issues at play in this paper and in discussions of moral axiology more broadly.

4 Some Serious Reasons Not to Hope

The world would be a better place if moral atrocities were merely apparent. However, even while the world would be better, it is a separate question whether we should, and rationally can, hope that apparent moral atrocities aren't actual. We just examined the best reason in favor of adopting aspirational illusionism: the world would turn out to be a morally and objectively better place than it would seem based on our moral appearances. There are, however, several reasons not to hope that atrocities are only apparent. We begin with the fact that many people believe they *know* that apparent atrocities are actual.

4.1 *We Have (Salient) Moral Knowledge*

Some readers will no doubt be puzzled by this puzzle in moral axiology. "Surely," they will say, "we can't rationally hope that apparent moral atrocities really aren't since we *know that they are*!" This is a reasonable response to the question of whether we can and should rationally hope that atrocities are morally illusory. We may think that the world would be a better place if atrocities were morally illusory, but that we cannot rationally hope this as we know otherwise. And this reason, among others, is exactly why (ABP) has a "defeasibility" clause governing both whether we can and whether we should hope that something is the case.

Some issues, like the moral status of murder, will be less contentious than, say, the moral status of abortions, the ethics of markets in blood and organs, etc. It is likely that many readers think the issue of whether *atrocities are morally illusory* fits the bill: we know that what we think is a moral atrocity, even if not invariantly, is usually a moral atrocity. As the philosopher Michael Huemer argues in *The Problem of Political Authority*:

> It is false that in general we do not know what is substantively morally correct. *Sometimes* we do not know what is substantively just. But often we do know. I do not know, for example, whether a ban on abortion would be unjust. But I know that the Jim Crow laws were unjust. (Huemer 2013, 172—original emphasis)

And the philosopher Perry Hendricks, in evaluating whether we should hope that the pro-choice or pro-life position on abortion is correct, argues that because the abortion issue is highly contentious, we should hope that the

doi: 10.48106/dial.v77.i2

pro-choice position is correct since that would make the world morally better compared to a world where the pro-life position on abortion is right—if the pro-lifers are correct, then that would presumably mean millions of fetuses are murdered in the womb each and every year, and who would hope for that? However, he doesn't think we can extend this axiological thinking generally to morally repugnant practices that are more certain to actually be moral atrocities since

> it does not make sense to hope that slavery is just because we know that slavery is unjust. It does not make sense to hope that something you know is false turns out to be true; it makes no sense to hope, for example, that the Seahawks won the 2006 Superbowl. In other words, hope that *p* entails that we do not know that $\sim p$. But we (or, at least, most of us) know that slavery is not justified, and hence we should not hope slavery is justified *even though the world would be better if it is.* The same goes for Nazis and rapists: we know that the Nazis were wrong, and we know that rapists are wrong. So, though the world would be better if Nazis and rapists were right, it makes no sense to hope that they were. (Hendricks 2021, 785)

So, if that's right, then even if the world would be better, we cannot rationally hope that moral atrocities are merely apparent. One important fact overlooked by Huemer and Hendricks is that not everyone agrees that we know, for example, that Jim Crow laws are wrong—just as some may hold that the permissibility of abortion is obvious but that the moral status of unjustified killing remains up for grabs morally speaking.

Here, we are not talking about racists or other moral degenerates, but those who either doubt that we have moral knowledge of *any* kind (Mackie 1977; Joyce 2001) or folks who, although they fall short of endorsing views like moral skepticism and nihilism, recognize that they could be wrong about their moral views, even if they assign low credence to such a possibility. Or we could hold that it is *likely* the case there are objective moral facts but still accept that there is a non-negligible probability there are no objective moral facts. For many people, the possibility that there are no or fewer moral atrocities than our moral appearances bear witness to is a stable and live epistemic possibility as a consequence of their more mundane metaethical views.

Even many folks who take themselves to know that there are moral atrocities that aren't merely apparent may still believe that there are genuine moral atrocities, with robust justification, who accept that it is possible—perhaps with a probability slightly higher than zero—even if highly unlikely, that there aren't moral atrocities. With respect to those folks, we can ask the question whether they should, as it looks like they can rationally, hope that apparent moral atrocities are illusory. As it happens, there are reasons, both epistemic and moral, that cut against aspirational illusionism, even for folks for whom the position is a live epistemic possibility.

4.2 *Less Reliable Moral Cognition*

We should care, as both epistemic and moral agents, about the reliability of our moral cognition (Dogramaci 2017; Braddock 2016). If our moral cognition is unreliable or even less reliable than we believe, then the result will be that we have a greater number of epistemically unjustified beliefs than we realize. And, in turn, those beliefs will, in principle, influence which actions we think are morally permissible—to the extent that we act according to what we think morality requires. We take a moral risk when acting on moral cognition with diminished or low levels of reliability: we could sincerely believe, say, that the consequences of our actions aren't nearly as bad as we think because the reliability of our moral cognition is less than what we believe.

And to the extent we want to minimize moral risk (within reason), this cuts against aspirational illusionism. Suppose that the world we reside in is filled with only apparent moral atrocities: there is nothing unjust, immoral, or morally bad about atrocities, despite moral appearances to the contrary. A serious epistemic and moral consequence would, of course, be that our moral cognition is less reliable than it would be otherwise. After all, consider that nearly universally, at least in morally enlightened societies, we take it as a moral given that famine, war, genocide, and the like are moral atrocities that should be prevented or mitigated as best we can. However, if these atrocities are only apparent, then our moral cognition—e.g., our moral intuitions about what morality requires of us—are even more unreliable than we realize.

It would be a huge miss, by our moral cognition, to be so deeply wrong about the moral nature of atrocities in that they look like brutalities beyond imagination that require our attention and effort to prevent and mitigate. We would thus be hoping for a world where, apparently, the most pressing moral issues and concerns are mere illusions. It would be hard to see how our moral

doi: 10.48106/dial.v77.i2

cognition could be anything but unreliable if we are wrong about the big moral stuff.

To hope that our world is such that atrocities are morally illusory would be to hope for a world where we take serious moral risks, unwittingly, due to the unreliability of our moral cognition. And where our moral cognition is so unreliable that we must worry with every action whether it really is morally required of us or whether the action we forego really is morally prohibited.

Not only that, if we lived in a world where our moral cognition is unreliable to such a degree, we would see moral illusions almost everywhere and would be left wondering what moral actions we should take. However, those actions would be highly morally risky, too, since they are based on moral intuitions that are generated by our unreliable moral cognition. So, while a world with less injustice would be better than a world with more injustice, *ceteris paribus*, we should bear in mind that to hope we live in such a world is to hope that we have less reliable moral cognition and take greater moral risks than we realize based on our moral appearances.

4.3 *Many Ungrounded, False Moral Beliefs*

There are epistemic costs to aspirational illusionism, too. As epistemic agents, ideally, we want to avoid or discard false beliefs and acquire true ones. We should want to avoid false beliefs to avoid the bad consequences of those false beliefs. To have false beliefs, as least related to what matters to us like survival and navigation, without the negative consequences of those false beliefs, would likely require "all manner of compensating false beliefs to make" the original false beliefs "fit with what else we know" (Joyce 2001, 179). The hitch, among others, is that often our beliefs influence not only our actions but also other beliefs that we are likely to take on board. If we have a false belief that *tigers are harmless cats who love to play chase*, then in an environment with many tigers, this false belief may get us killed. That false may not get us killed, however, if we have a false but compensating belief that *tigers like it best if we avoid them entirely to make the chase more challenging* (Plantinga 1993, 225–226).

We never know if, when, or how a belief will be called into action—where we must rely on the belief to achieve an important and valuable goal—and, "given this, it is better that [the beliefs we form are] true than false" (Joyce 2001, 179). This is one of many reasons why it matters whether our beliefs are true. And yet, to hope that moral atrocities are apparent is, by implication, to hope that we have a large inventory of false beliefs, ungrounded salient moral facts. Even if we don't recognize it, we would then have many moral and non-moral beliefs about moral atrocities that would be false. To hope that apparent moral atrocities aren't actual is to hope, by implication, that many of our beliefs about history, public policy, and, of course, moral beliefs themselves are false. This isn't to claim that people who hope the world is morally better than it looks intend to hope for the epistemic costs of their hopes, but it would be one of the costs nonetheless of their view, even if they don't realize such would be an (unfortunate) implication.

4.4 *Opportunity Costs*

There would be many moral, practical, and cognitive opportunity costs if the world is such that apparent atrocities are merely illusory (Buchanan 1991). If the world is that way, it means that many moral problems that aren't atrocities are neglected, to varying degrees. This is because we spent substantial amounts of time, resources, and effort trying to prevent, mitigate, and address atrocities when it was morally unnecessary, given aspirational illusionism, and those resources would be wasted. Let's start with the moral opportunity costs.

First, the moral opportunity costs of trying to prevent and mitigate merely apparent atrocities would be very high. There are many events in the world that are morally wrong but fall short of moral atrocities, which have received less attention and resources because some of the attention is diverted to addressing merely apparent atrocities. So, to hope that the world is such that there are fewer or no moral atrocities is to hope that the world is such that we've wasted time and resources attempting to mitigate and prevent events that should have been applied elsewhere. For instance, there are no doubt many small evils in the world that aren't moral atrocities but that we could have mitigated had we focused more of our energies there instead of mitigating apparent atrocities.

To hope the world contains no actual atrocities, only apparent ones, would be to hope that we wasted many opportunities trying to prevent war and genocide rather than focusing on small but morally bad and evil events like bad headaches, heartbreak, discouraging bullying, and whatnot. For example, it looks like, in the aggregate, enough small evils and suffering would amount to a moral atrocity (as related to the problem of evil, see Case 2020). There are many people who suffer where it falls short of a moral atrocity, whose,

doi: 10.48106/dial.v77.i2

say, bellies hurt and teeth ache, and could be helped if we spent resources on them instead of mitigating merely apparent atrocities.

Next, consider that we are limited epistemic agents—we only have so much time and cognitive resources to shift through beliefs and memories to find what we need. The more that is stored in memory, *ceteris paribus*, the more records our cognitive systems must shift through to find the needed record. Though limited epistemic agents like us may have, practically, a nearly endless storage capacity, and assuming that

> there are obvious advantages of having virtually unlimited capacity in that domain, the limitations on retrieval access can be viewed as a necessary filter. In the interest of speed, accuracy, and avoiding confusion, *we do not want every item in our memories to be accessible*. (Bjork and Vanheule 1992, 157—my emphasis)

We do not want to recall every memory and belief because doing so would clutter our cognitive lives too much past the point where those records are useful. And if we stored numerous false moral beliefs, given aspirational illusionism. We should avoid storing false beliefs, not only to avoid wasting resources in retrieval but also because "retrieved records will often trigger additional thoughts [...] retrieving more records generally requires additional thinking" (Michaelian 2011, 411). So, if atrocities aren't actual, there are weighty cognitive opportunity costs that result from spending cognitive resources to solve moral problems that wouldn't be real.

We have reviewed some of the problems and costs of aspirational illusionism. There remains, though, something off-repellant about the hope that moral atrocities are merely apparent, but it is one that is difficult to flesh out. We attempt to unpack it in the penultimate section.

4.5 *Moral Repugnance*

There is an indirect, but still valuable, reason against aspirational illusionism. The world would be a better place if atrocities were merely apparent in the sense that the world would be morally less bad, *ceteris paribus*. Nonetheless, there is something morally repugnant about hoping that atrocities are merely apparent, even if the world would be better for it. This is deeply puzzling: there appears to be an obvious axiological bridging principle that we should—or at least we are permitted to—hope that something is the case if we have good reason to believe it would make the world a better place, and it is an epistemic possibility. On its face, it is puzzling why hoping so would be repugnant. Perhaps, though, there are a couple solutions to the puzzle.

The first solution is the most obvious: while the world would be a better place if atrocities were merely apparent, some of us cannot rationally hope that is the case since we believe we know that atrocities cannot be merely apparent. Even if the world would be a better place if atrocities were morally illusory, some individuals cannot rationally hold aspirational illusions, given their firm belief that they have moral knowledge to the contrary.

There are individuals who, even though they aren't moral skeptics or moral nihilists, don't take themselves to have moral knowledge; they believe, however, there are solid moral reasons and moral evidence (e.g., moral intuitions) to believe that apparent atrocities are actual. Should we conclude that such agents could rationally hope that atrocities are only apparent? My strong intuition here: there is something bizarre about someone with strong evidence that apparent moral atrocities are actual, hoping they are only apparent. This intuition, though, is puzzling: if an epistemic agent had good reason to believe that moral atrocities were actual and not merely apparent—they find arguments for moral skepticism slightly convincing—it looks like they're still in a position, given the world would be morally better if so, to hope that moral atrocities are merely apparent. So why the strong intuition otherwise?

Here's a tentative explanation: perhaps the reason the author has a strong contrary intuition is that humans are deeply moral creatures: most of us, for various reasons, have a strong sense of right, wrong, justice, and fairness, to name but a few. Our moral identity and how we morally evaluate our life events help to shape a fundamental and abiding aspect of our psychological identity: it matters not only how we treat others but how others treat us and how we see each other as moral agents (Hardy and Carlo 2011; Sauer 2019).

Whether this moral sense is merely the product of evolutionary and cultural processes or partly the result of something more metaphysical is beside the point: we clearly have a deep sense of justice, fairness, and right and wrong—one that cannot, for most of us, be easily ignored or forgotten. It would be hard for many of us to ignore the fact that we were mugged on the way home from a play by an assailant with a knife. It isn't simply that we were scared that it would happen again, but that the mugger profoundly wronged us with his actions—he didn't simply violate our sense of safety; though he did that, too, he violated our moral sense of agency.

doi: 10.48106/dial.v77.i2

Imagine you were told by someone you respected that the violent mugging you endured was merely illusory, and thus, despite how it appeared to you, it was a morally neutral event. (We will assume that a violent mugging is, at least, a minor atrocity—if you object to this, then pick your favorite example.) The mugger didn't actually harm you, despite your feelings of betrayal and the resulting trauma. To be told this by someone you love and respect, even if accurate, would be hard to square with the profound sense of injustice you felt as a result of the mugging. This isn't to claim everyone would feel this way about the issue, but it is likely many people would. There's an odd sense in which hoping that apparent atrocities are illusory undercuts an important and deep respect for people as moral agents.

5 Conclusion

This paper asked whether it would be rational to hope that atrocities like war and genocide are merely illusory. Even if basic moral truths hold necessarily, axiological judgments are based on live epistemic possibilities, not metaphysical ones. For agents who lack salient moral knowledge that apparent atrocities are actual, we can rationally ask whether they should hope they are. We explored a solid reason to hope so: if atrocities are only apparent, then the world is objectively and morally better than it would otherwise be if they were actual; but if atrocities aren't merely apparent, then the world is as morally bad as it appears, and perhaps worse.

In contrast, there are some reasons we either should not or cannot rationally hope that atrocities are merely apparent. The most obvious: some individuals *know they are real atrocities*. However, even for those who lack such moral knowledge that atrocities are actual, there are reasons that cut against the hope: our moral cognition would be less reliable than it would be otherwise, we would have many false, epistemically ungrounded moral beliefs, and we would have wasted resources trying to address merely apparent atrocities. Not to mention one final reason: there is something morally suspicious about hoping moral atrocities are only apparent that is deeply undignified with respect to victims of atrocities. So, while it isn't clear what we should hope for, what is clear is that moral axiology is worth further exploration.*

* Thanks to Liz Jackson for reading the manuscript and offering insightful and useful feedback that seriously improved the paper. The errors that remain are my own.

Jimmy Alfonso Licon
000-0001-9451-291X
School of Historical, Philosophical, and Religious Studies
Arizona State University
Jimmy.Licon@asu.edu

References

Benton, Matthew A. 2020. "Epistemological Aspects of Hope." in *The Moral Psychology of Hope*, edited by Claudia Blöser and Titus Stahl, pp. 135–151. Lanham, Maryland: Rowman & Littlefield, doi:10.5771/9781786609731.

—. 2021. "Knowledge, Hope, and Fallibilism." *Synthese* 198(suppl. 7): 1673–1689, doi:10.1007/s11229-018-1794-8.

Bjork, Robert A. and Vanheule, Marc. 1992. "Retrieval Inhibition and Related Adaptive Peculiarities of Human Memory." in *Advances in Consumer Research, volume 19*, edited by John F. Sherry and Brian Sternthal, pp. 155–160. Duluth, Minnesota: The Association for Consumer Research.

Blanchard, Joshua. 2020. "Moral Realism and Philosophical Angst." in *Oxford Studies in Metaethics*, volume XV, edited by Russ Shafer-Landau, pp. 118–139. Oxford: Oxford University Press, doi:10.1093/oso/9780198859512.003.0006.

Bostrom, Nick. 2003. "Are we Living in a Computer Simulation?" *The Philosophical Quarterly* 53(211): 243–255, doi:10.1111/1467-9213.00309.

Braddock, Matthew. 2016. "Evolutionary Debunking: Can Moral Realists Explain the Reliability of Our Moral Judgments?" *Philosophical Psychology* 29(6): 844–857, doi:10.1080/09515089.2016.1163667.

—. 2017. "Debunking Arguments from Insensitivity." *International Journal for the Study of Skepticism* 7(2): 91–113, doi:10.1163/22105700-20171195.

Buchanan, James M. 1991. "Opportunity Cost." in *The World of Economics*, edited by John Eatwell, Murray Milgate, and Peter Newman, pp. 520–525. The New Palgrave. London: MacMillan Publishing Co., doi:10.1007/978-1-349-21315-3_69.

Case, Spencer. 2020. "Small Evils and Live Options: A New Strategy Against the Argument from Evil." *Philosophia Christi* 22(2): 307–321, doi:10.5840/pc202022226.

Chalmers, David J. 2010. *The Character of Consciousness*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780195311105.001.0001.

Crummett, Dustin. 2021. "The Real Advantages of the Simulation Solution to the Problem of Natural Evil." *Religious Studies* 57(4): 618–633, doi:10.1017/S0034412519000726.

Davison, Scott A. 2018. "God and Intrinsic Value." in *Does God Matter? Essays on the Axiological Consequences of Theism*, edited by Klaas J. Kraay, pp. 39–45. Routledge

Studies in the Philosophy of Religion n. 18. London: Routledge, doi:10.4324/9781315210995-2.

Dogramaci, Sinan. 2017. "Explaining Our Moral Reliability." *Pacific Philosophical Quarterly* 98(s1): 71–86, doi:10.1111/papq.12153.

Hardy, Sam A. and Carlo, Gustavo. 2011. "Moral Identity: What Is It, How Does It Develop, and Is It Linked to Moral Action?" *Child Development Perspectives* 5(3): 212–218, doi:10.1111/j.1750-8606.2011.00189.x.

Hendricks, Perry. 2021. "The Axiology of Abortion: Should We Hope Pro-Choicers or Pro-Lifers Are Right?" *Ergo* 7(29): 774–788, doi:10.3998/ergo.1126.

Howard-Snyder, Daniel and Howard-Snyder, Frances. 1999. "Is Theism Compatible with Gratuitous Evil?" *American Philosophical Quarterly* 36(2): 115–130.

Huemer, Michael. 2013. *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey*. Basingstoke, Hampshire: Palgrave Macmillan, doi:10.1057/9781137281661.

Joyce, Richard. 2001. *The Myth of Morality*. Cambridge: Cambridge University Press, doi:10.1017/cbo9780511487101.

Kahane, Guy. 2011. "Should we Want God to Exist?" *Philosophy and Phenomenological Research* 82(3): 674–696, doi:10.1111/j.1933-1592.2010.00426.x.

—. 2012. "The Value Question in Metaphysics." *Philosophy and Phenomenological Research* 85(1): 27–55, doi:10.1111/j.1933-1592.2012.00589.x.

Licon, Jimmy Alfonso. 2021. "Aspirational Theism and Gratuitous Suffering." *Religious Studies* 57(2): 287–300, doi:10.1017/s0034412519000210.

Lougheed, Kirk. 2017. "Anti-Theism and the Objective Meaningful Life Argument." *Dialogue. Revue canadienne de philosophie / Canadian Philosophical Review* 56(2): 337–355, doi:10.1017/s0012217317000373.

Mackie, John Leslie. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin Books.

Martin, Adrienne M. 2014. *How We Hope: A Moral Psychology*. Princeton, New Jersey: Princeton University Press, doi:10.23943/princeton/9780691151526.001.0001.

Meirav, Ariel. 2009. "The Nature of Hope." *Ratio* 22(2): 216–233, doi:10.1111/j.1467-9329.2009.00427.x.

Michaelian, Kourken. 2011. "The Epistemology of Forgetting." *Erkenntnis* 74(3): 399–424, doi:10.1007/s10670-010-9232-4.

Penner, Myron Arthur and Lougheed, Kirk. 2015. "Pro-theism and the Added Value of Morally Good Agents." *Philosophia Christi* 17(1): 53–69, doi:10.5840/pc20151714.

Plantinga, Alvin. 1993. *Warrant and Proper Function*. Oxford: Oxford University Press, doi:10.1093/0195078640.001.0001.

—. 2004. "Supralapsarianism, or 'O Felix Culpa'." in *Christian Faith and the Problem of Evil*, edited by Peter van Inwagen, pp. 1–25. Grand Rapids, Michigan: William B. Eerdmans Publishing Co.

Pojman, Louis P. 1986. "Faith Without Belief?" *Faith and Philosophy* 3(2): 157–176, doi:10.5840/faithphil19863213.

Sauer, Hanno. 2019. "The Argument from Agreement: How Universal Values Undermine Moral Realism." *Ratio* 32(4): 339–352, doi:10.1111/rati.12233.

Smith, Nicholas J. J. 1997. "Bananas Enough for Time Travel?" *The British Journal for the Philosophy of Science* 48(3): 363–389, doi:10.1093/bjps/48.3.363.

doi: 10.48106/dial.v77.i2

On the Plurality of Parts of Classes

Daniel Patrick Nolan

abstract is missing

The ontological picture underpinning Lewis's *Parts of Classes* (1991) has some unusual features. It posits many, many simple, abstract objects that serve to be the subject matter of set theory. (We require so many, as Lewis points out, since standard set theory is committed to so many sets.) However, when we put the ontology posited by *Parts of Classes* together with the doctrines of Lewis's *On the Plurality of Worlds* (1986), two problems surface. The first, to do with the relationship between sets and possible worlds, is perhaps a drawback but is a result a Lewisian could comfortably accept. However, the second problem, concerning how to integrate this ontology with Lewis's understanding of possible worlds, may look more like an inconsistency, though I will argue that we can interpret Lewis consistently here. The second tension is a more serious problem in the combination of Lewis's views unless it is dealt with. There are two ways to resolve this second tension, each of which goes beyond what Lewis explicitly says in interesting ways. I think Lewis would have been best off extending his system in the second way I will suggest, and indeed, there is some textual evidence that he may have been tempted to extend it in this way as well. This gives Lewis an additional reason to embrace a proper class of worlds and *possibilia*, over and above others explored in the literature.

Lewis's central conjecture in *Parts of Classes* is that "the parts of classes are all and only their subclasses." By "class," Lewis meant things with members: the empty *set* was excluded from Lewis's use of "class," and while he counted all other sets as classes, he also defended the view that there are some classes that are not sets (so-called *proper classes*). From his central conjecture and the exclusion of the empty set, it follows that unit classes (i.e., classes with exactly one member) are atomic, lacking proper parts altogether. (All of the other classes are fusions of these unit classes.) How many unit classes are there, according to Lewis? As many as there are sets at all, since each set belongs to a unit class. (For Lewis, proper classes are distinguished by not belonging to classes.) There are thus "proper-class many" atoms in the ontology of Lewis (1991), since there are more sets than the cardinality of any set whatsoever.

1 Wholly Impossible Atoms

Put this together with the commitments of Lewis (1986), and the first problem for the view emerges. Lewis (1986) is committed to there only being a set of possible worlds and a set of possible objects (1986, 104), so almost all the atoms postulated to be the unit sets in Lewis (1991) must lie outside the possible worlds, in the sense of not being part of those worlds.[1] (A proper class of objects minus a set of objects leaves a proper class of objects.) Any objects that exist outside all of the possible worlds must be *impossible*: to be a possible object is to be part of a possible world, according to Lewis. Is this inconsistent, to be committed both to the claim that certain things exist and that it is impossible that such things exist? Not according to Lewis's system. Lewis already admits that there are entities that do not exist in any possible world: since he accepts unrestricted mereological composition, he accepts that objects in different worlds make up fusions that cannot be entirely found in any single world and that, in a sense, these trans-world fusions "cannot possibly exist" (Lewis 1986, 211). That is, no single possible world is a witness to their existence, and there is no world that they are a proper part of (as opposed to parts of them being parts of worlds). So it is not inconsistent for Lewis, given this sense of "possible," to say that there exist objects that do not possibly exist.

However, the atoms postulated for the purposes of mathematics by Lewis (1991) are arguably in a worse position than trans-world fusions. At least the fusions resolve into parts, each of which is part of a world, and the aggregate of all worlds has all of them as parts. The trans-world fusions are "in" the

1 Lewis distinguishes three ways of "being in" a possible world in Lewis (1983, 39–40). Lewis (1986, 96) adds a fourth way involving counterparts that need not concern us here. The first is to be part of the world in question; the second to be partially in a world (i.e., to share a part with a world); and the third is to exist "from the standpoint of a world": in effect, to be one of the things that an inhabitant of a world that shared our ordinary way of talking might correctly talk about as existing. This third way of "being in" was intended, in Lewis (1983), to include sets or properties that might not count as part of a given world, such as, e.g., the pure sets. Even if we are licensed as counting some entities as being "in" our world without overlapping it, my focus in this paper is on what objects are parts of worlds. For the sake of an idiomatic discussion, I will talk as if all and only parts of worlds are "in" those worlds, unless indicated otherwise, though my point concerns Lewis's commitments about what are parts of the possible worlds.

world*s*, at least collectively. However, these atoms postulated by Lewis (1991) must be "completely impossible," as I have put it elsewhere (Nolan 2002, 156, n.9). They are not parts of any world, and no part of them is part of any world either. (It may well be that there is a different "advanced modalising" sense of possible, in which every existence claim that is true is possibly true—see Lewis 1986, 6; Divers 1999, 229–230—but I do not want to wade into any debates about how this might be best understood here.) No doubt Lewis could stipulate that these atoms are possible in some sense and perhaps intends to with his talk of sets existing "from the standpoint of a world"; see Lewis (1983, 39–40) and footnote 1 above. That would not stop them from failing to be possible in the way that possible objects are typically possible in Lewis's system and would not stop them from sharing the "impossibility" that Lewis admits trans-world fusions have. It strikes me that it would be better to have a non-disjunctive account of possibility in terms of possible worlds if such a thing could be had.[2]

Lewis does not tell us much about these atoms. We do not have answers about what intrinsic nature they have, if any, or what relations they may stand in, if any (Lewis 1991, 31–35, 142–143): only that they are the singleton sets of other entities (either individuals or other classes).[3] If we move to the structuralist understanding of set theory set out in the appendix of Lewis

2 One referee has suggested that Lewis better preserves the mathematical platonist intuition that numbers, sets, and other such mathematical objects are not parts of concrete reality and are not found in space and time by holding that they are all disjoint with his possible worlds. That will be an advantage for some of Lewis's way of going. But for other mathematical platonists, it may seem like an undesirable upshot of his attempt to account for possibility with alternative *concrete* cosmoi when reality contains a non-concrete aspect as well.

3 Lewis also says that we do not know if they have locations and indeed "haven't a clue" whether they do (Lewis 1991, 33). This marks a departure from the view he expresses in Lewis (1986, 94–96) that sets are located where their members are.

If classes do have spatiotemporal locations, that would make them worldmates with individuals, at least in worlds like ours, and so parts of possible worlds: so given his *Priority Thesis*, that no class is part of any individual, some or all of the possible worlds would fail to be individuals (having parts that are classes). He is also committed to all possible worlds being individuals (Lewis 1986, 83), which leaves his views in conflict. (At least unless he concedes he does not have a clue whether his own theory is correct.) His own views, by the time of Lewis (1991), committed him to denying sets of spatiotemporal locations. His implicit commitment to nearly all the singletons being outside all the worlds also requires that most of them lack spatiotemporal locations. I think the Lewis of 1991 would be well-advised to renounce his scepticism about the location of singletons and instead admit that they all *lack* spatiotemporal location. A more contemporary Lewisian tempted by the more radical revisions suggested towards the end of this paper may wish to revise that commitment again, however.

doi: 10.48106/dial.v77.i2

(1991) or in Lewis (1993), then we do not even require that our atoms stand in a distinctive singleton relation. The demand that there be proper-class many of them while there are only set-many objects in the possible worlds will remain, however, so this aspect of his view will require that nearly all the atoms postulated will be "completely impossible" in the sense above, even when we move to Lewis's structuralist framework.

Another feature of Lewis's system ensures that *all* the atoms needed to be classes are disjoint from the possible worlds, whether or not we move to the structuralism mentioned above. Possible worlds and their contents are treated as *individuals* in Lewis's system; that is, they are the ur-elements and do not themselves have members (Lewis 1986, 83). Lewis furthermore insists on a *Priority Thesis* (Lewis 1991, 7): that no class is part of any individual. So, in particular, no class can be a part of any possible world. So we are left with the result that there must be proper-class many atoms outside all of the possible worlds, serving as the ontology of class theory, even if we go structuralist about the relationship between those atoms and the entities that they are the singleton sets of.

2 Are the Mathematical Atoms Worlds After All?

The second problem to be addressed in this paper emerges when we come to consider which things count as worlds. Given the letter of Lewis (1986), it might seem that these atoms must be parts of worlds after all. Lewis defines a *worldmate* relation: his first pass is to say "things are worldmates iff they are spatiotemporally related" (1986, 71), and then extends this to include as worldmates entities that are "*analogically* spatiotemporal" (1986, 75–76), to handle alien possibilities where the connections between entities are not the actual, familiar, spatiotemporal relations. Lewis also says that a world "is a maximal sum: anything that is a worldmate of any part of it is itself a part" (1986, 69). Furthermore, it is clear from the context that these are the only parts of worlds, and nothing further is required to be a world than to be such a maximal sum since he has taken himself to have given "the unity relation for possible worlds" (1986, 70).[4]

Lewis also accepts unrestricted composition: for any entities, there is a sum of those entities (Lewis 1986, 201; 1991, 74). Now, consider two cases for each

4 This account of worldmates would have to be modified were Lewis to accept the existence of immanent universals, as he points out in Lewis (1986, 69) and especially (1986, 208–209).

of these allegedly beyond-wordly mathematical atoms. Either it has some worldmates, or it has no worldmates. In the first case, there will be a sum of it and its worldmates (and their worldmates, etc.), and so it is part of a possible world. In the second case, it has no worldmates; therefore, the sum of it alone satisfies the condition "anything that is a worldmate of any part of it is itself a part." It is degenerately maximal under the worldmate relation in this way. So it is a possible world all by itself and, so, part of a world (an improper part of itself). But, as pointed out above, these atoms must not be parts of any possible world. We have reached a contradiction.

Let us deal with this apparent inconsistency first. One potential repair is obvious: instead of understanding maximal interrelation in the manner presupposed by the previous paragraph, Lewis could insist that possible worlds are *non-degenerately* maximally interrelated by spatiotemporal, or analogously spatiotemporal, external relations. For example, he could say that every world w is a sum with at least one part, and w includes all the worldmates of that part *and that its parts all have worldmates*. A single atom not standing in worldmate relations to anything would not count as a world on this revised definition. This is the natural way to understand the spirit of specifying things as "worldmates": if something is not even its own worldmate, plausibly, it is not in any world. I expect Lewis intended that everything that was part of a possible world would stand in spatiotemporal relations, or at least analogously spatiotemporal ones, and that this is how we should read his definition of a possible world.

This definition of worlds need not even rule out worlds of a single mereological atom, since it may be that atoms stand in spatiotemporal relations or analogously spatiotemporal relations *to themselves*. On this proposal, it is not trivial that everything is its own worldmate, but nevertheless, things that stand in the right kinds of relations to themselves can be their own worldmates. We would need to draw a distinction between being zero distance from oneself and not being in any spatiotemporal relationship to oneself at all if we wanted some atoms to be their own worldmates but some (indeed, most) to not be, but we should probably want to draw this distinction in any case if we are to allow it to be coherent for something to not be in space and time, since such a thing is not located at all and, so, not co-located with itself.

Avoiding the contradiction in this fashion, however, does have an unwelcome consequence for Lewis's system. It will rule out as a possibility that there could be an individual object that did not stand in spatiotemporal relationships (or analogously spatiotemporal relationships) to anything. On

doi: 10.48106/dial.v77.i2

the face of it, there does not seem to be anything metaphysically necessary about there being spatiotemporality. Why couldn't there just be an electron on its own, with charge and spin but no spatiotemporal features? You might reply that electrons are essentially spatiotemporal, so it would have to have location and perhaps duration. But what about some radically different kind of individual, existing by itself, not in space and time? It does not seem to be essential to being a non-class that something is in space or time (or is in relations analogous to spatio-temporal ones). There is nothing, on the face of it, incoherent about such a scenario, yet if it does not occur in any possible world, by Lewis's standards, it is not possible, at least in the "ordinary" sense.

Counting something that is apparently possible, in the standard sense of metaphysically possible, as being impossible is a mark against this version of Lewis's theory. However, this problem is similar to other kinds of marks against Lewis's theory: Lewis also cannot allow that there could be nothing concrete and cannot allow, in the ordinary sense of possible, that it is possible for there to be co-existing objects that are not spatiotemporally related to each other (and not analogously-spatiotemporally related to each other). In each of these other cases, Lewis bites the bullet, allowing that these apparent possibilities are not indeed possible, and he considers these as costs worth paying for the attractions of his theory (Lewis 1986, 71–74). So a Lewisian who refused to countenance the possibility of an entity not standing in spatiotemporal or analogously spatiotemporal relations, even to itself or its own parts, would probably bite the bullet on this in a similar way.

One option for Lewis here, suggested by a referee, would be to allow that atoms standing in no relations could be their own worldmates, but to put a constraint on the worldmate relation so that *non-individual* atoms (i.e., singletons) never counted as their own worldmates. A featureless individual could then be possible, without all the singletons being their own worldmates and thus their own possible worlds. I would be uncomfortable with solving the problem through redefinition like this without an explanation of *why* it makes a difference to the metaphysics of possibility whether a featureless atom is a class or not, though other's tastes may differ. At any rate, I think this sort of solution will be difficult to plausibly implement were we to move to the structuralist approach preferred by Lewis in Lewis (1993), where there would be no intrinsic difference (or difference in natural external relations) between the featureless atoms that played the structural role of singletons and those that would not. While the referee is right that there is an option here, I will turn to revisions I find more satisfying.

3 A Natural Resolution: Worlds Form a Proper Class

Ruling out the mathematical simples as counting as worlds, or indeed being in worlds at all, also retains the strikingly implausible feature of the *Parts of Classes* system mentioned above. Since each of these atoms is not a part of any possible world, it remains completely impossible. The other way of responding to the question of whether these atoms postulated to be the singletons are worlds would be to embrace the claim that each *is* a possible world after all and that when a thing is not a worldmate of anything else, it is a possible world all by itself. To modify his views in this way, Lewis would need to drop the claim that the possible worlds form a set and that the possible objects form a set. Once that is done, we can allow that there are proper-class many atomic *possible* worlds alongside all of the worlds embraced in Lewis (1986). These possible worlds can then serve as the ontology for mathematics. There is now no need to say that those objects are absolutely impossible, since they are just additional possible worlds. As a bonus, we can now recognise as a genuine possibility that something exists without being spatiotemporally related to anything (nor standing in a relation analogous to spatiotemporal ones). Lewis would need to answer "yes" to the question of whether there are indistinguishable possible worlds if nearly all of them are featureless atoms, and this was a question he wished to stay neutral on, but giving up neutrality for a good theoretical reason does not seem like a cost.[5]

We face some choices about whether to treat every possible world as an individual. (That is, in this context, a member of a class that is not itself a class.) On the current proposal, some are, and some are not. If we did want all possible worlds to be individuals while insisting that all the atoms serving as singletons were in worlds, we could instead adopt a position where some or all possible worlds had individual parts and singleton parts. (This would require that some "mixed fusions" of classes and individuals were themselves individuals, contrary to the letter of Lewis 1991, 7–8 and Lewis's *Priority Thesis*, but the modification makes little difference to the overall system.) We

5 Divers (1994) argues that a Lewisian should reject indistinguishable possible worlds, largely on the grounds of quantitative parsimony. Parsimony arguments are at their strongest when theories are equal, or nearly equal, in other respects. But if a Lewisian theory with many duplicate featureless worlds provides an ontology for mathematics without "completely impossible" objects, while its rival requires nearly all the entities committed to lie entirely outside the possible worlds, then the former theory plausibly has a theoretical advantage that outweighs any cost in parsimony, especially if the latter theory is arguably just as unparsimonious, only about the number of entities outside possible worlds.

doi: 10.48106/dial.v77.i2

would also want to tweak Lewis's definition of the null set (Lewis 1991, 10–15) to continue to ensure that it had no classes as parts, perhaps by making it the fusion of all atomic individuals. Further choices may have to be made: Does every possible world contain classes? Does each contain all of them (perhaps through trans-world identity), or is the mathematical universe spread out amongst them? These are theoretical choices we can leave to partisans of this kind of view, should there ever be any.

A more radical option also becomes available once we no longer need proper-class many atoms outside the possible worlds. Instead of accepting the existence of proper-class many additional atoms, whether within or outside of worlds, we could instead allow the more usual inhabitants of possible worlds to provide the material for mathematics, provided only that there are enough of them. If there are proper-class many possible electrons, for example, a variation on the structuralism of the appendix of Lewis (1991) or of Lewis (1993) can be employed to let them be the ontology of mathematics while also preserving their role as individuals (i.e., ur-elements of sets). I have explored one way of developing a view like this, with different motivations: see Nolan (2002, chap. 7 and appendix) and Nolan (2019), Schwarz (2005) and Cowling (2017, chap. 7) offer introductions and some philosophical motivations for the system. This way of developing a megethological system requires minor modifications to two of Lewis's principles used to develop his *Parts of Classes* framework: both the *Division Thesis* and the *Fusion Thesis* must be tweaked. (Nolan 2002, 162–163, 195–200 on the *Division Thesis*; and 2002, 165–169 on the *Fusion Thesis*). The *Fusion Thesis*, that every fusion of individuals is itself an individual, needs to be given up in any case as soon as we have a proper class of individual atoms, unrestricted composition, and global choice (Nolan 2002, 169), so it would be very natural to restrict the *Fusion Thesis* in a setting like this in any case. Since my revisions require the use only of ontology found in possible worlds, the question of what to do with the proper-class many mysterious atoms lying outside all the possible worlds evaporates since the system no longer needs them.

Note that Lewis himself may have had some sympathies for this revision to his system. In Lewis (2002), Lewis says the case for postulating proper-class many possibilia such as electrons is "fairly persuasive" (2002, 8). If he endorsed that change, he would be able to accommodate all of his mathematical ontology within possible worlds after all. And given that he ended up endorsing a structuralist conception of the relationship between individuals and sets (Lewis 1993), he would have been able to have an ontology and ideology of mathematics that required no more than commitments he had already incurred in his theory of possible worlds.

Moving to a proper class of possible objects, and perhaps with it a proper class of possible worlds, would have some disadvantages as well, as Nolan (1996, 249–251) points out. Proper classes are not members of sets, so one has to be careful employing set-theoretic constructions out of possible objects or possible worlds for other purposes. Natural language semantics in the possible worlds tradition helps itself to functions from all sorts of classes that may well turn out to be proper classes on this proposal (see Partee 1989 for a classic introduction), and pressing classes of possible individuals into service in metaphysics (in the style, e.g., of Montague 1969) will also face problems. Lewis (2002, 8–10) discusses some of the moves that might need to be made in the face of this challenge.

There are many options available to those tempted to operate with a proper class of possible objects. Some are canvassed by Nolan (1996, 249–253). Another option is to reconceive the task of possible worlds semantics as not providing the once-and-for-all semantic values of expressions but just to be providing models of semantic values that have some perspicuous connections to the meanings of expressions. We can offer set-sized models with a set of "worlds" and a set of "possible objects" that can display, e.g., systematic connections between the semantic values of simple and progressive tenses, even if, in reality, there are more than set-many possible completed bakings of cakes and more than set-many possible bakings of cakes in progress. Operating as if semantic values can be modelled straightforwardly in set theory can be productive, even if there are foundational issues lurking about what these set-sized models have to do with modal space and the "real" semantic values of expressions, whatever those might be. The project of possible worlds semantics, as traditionally conceived, does not need to grind to a halt, even if the models semanticists are working with are more limited than they might have realised.

Bringing out the tension in the ontologies of Lewis (1986) and Lewis (1991) is no mere pedantry. Resolving the tension between the two works provides us with another motivation to endorse a proper class of possible worlds and possible individuals, besides those suggested by Nolan (2002). (Nolan 2002 argues that moving to a proper class of worlds and individuals gives the modal realist a more satisfactory principle of recombination and an appealing alternative to the *Parts of Classes* machinery for class theory.) A modal realist who wishes to resist this resolution owes us an account of why it is an acceptable

doi: 10.48106/dial.v77.i2

cost of her theory to deny that atomic possibilities of the sort described above are genuine possibilities and why it is worth postulating "entirely impossible" ontology, i.e., objects that not only do not exist in worlds but which do not divide into parts that exist in worlds. Without motivating these bullet-bitings, a modal realist who resists a proper class of possible individuals would seem to be settling for second-best modal realism.*

Daniel Patrick Nolan
0000-0003-0055-0611
Department of Philosophy
University of Notre Dame
dnolan2@nd.edu

References

COWLING, Sam. 2017. *Abstract Entities*. London: Routledge, doi:10.4324/9781315266619.

DIVERS, John. 1994. "On the Prohibitive Cost of Indiscernible Concrete Possible Worlds." *Australasian Journal of Philosophy* 72(3): 384–389, doi:10.1080/00048409412346181.

—. 1999. "A Genuine Realist Theory of Advanced Modalizing." *Mind* 108(430): 217–239, doi:10.1093/mind/108.430.217.

LEWIS, David. 1983. *Philosophical Papers, Volume 1*. Oxford: Oxford University Press, doi:10.1093/0195032047.001.0001.

—. 1986. *On the Plurality of Worlds*. Oxford: Blackwell Publishers.

—. 1991. *Parts of Classes*. Oxford: Blackwell Publishers.

—. 1993. "Mathematics is Megethology." *Philosophia Mathematica* 1(1): 3–23. Reprinted in Lewis (1998, 203–229), doi:10.1093/philmat/1.1.3.

—. 1998. *Papers in Philosophical Logic*. Cambridge: Cambridge University Press, doi:10.1017/cbo9780511625237.

—. 2002. "Tensing the Copula." *Mind* 111(441): 1–13, doi:10.1093/mind/111.441.1.

MONTAGUE, Richard. 1969. "On the Nature of Certain Philosophical Entities." *The Monist* 53(2): 159–194. Reprinted in Montague (1974, 148–187), doi:10.5840/monist19695327.

—. 1974. *Formal Philosophy: Selected Papers*. New Haven, Connecticut: Yale University Press. Edited and with an introduction by Richmond H. Thomason.

NOLAN, Daniel Patrick. 1996. "Recombination Unbound." *Philosophical Studies* 84(2/3): 239–262, doi:10.1007/bf00354489.

* Thanks to Sara Bernstein, Joshua Kelleher, Alexander Sandgren, and several anonymous referees for helpful feedback.

—. 2002. *Topics in the Philosophy of Possible Worlds*. New York: Routledge, doi:10.4324/9781315054872.

—. 2019. "Individuals Enough for Classes." On PhilPapers' Archive, said to be from 2004, https://philpapers.org/archive/NOLIEF.pdf.

PARTEE, Barbara Hall. 1989. "Possible Worlds in Model-Theoretic Semantics: A Linguistic Perspective." in *Possible Worlds in Humanities, Arts and Sciences*, edited by Sture ALLÉN, pp. 93–123. Research in Text Theory n. 14. Berlin: de Gruyter. Proceedings of Nobel Symposium 65, doi:10.1515/9783110866858.93.

SCHWARZ, Wolfgang. 2005. "Emperors, Dragons, and Other Mathematicalia." Unpublished manuscript, https://www.umsu.de/papers/emperors.pdf.

doi: 10.48106/dial.v77.i2

Published by *Philosophie.ch*

Verein philosophie.ch
Fabrikgässli 1
2502 Biel/Bienne
Switzerland
dialectica@philosophie.ch

https://dialectica.philosophie.ch/

ISSN 0012-2017

Dialectica is supported by the Swiss Academy of Humanities and Social Sciences.

Abstracting and Indexing Services

The journal is indexed by the Arts and Humanities Citation Index, Current Contents, Current Mathematical Publications, Dietrich's Index Philosophicus, IBZ — Internationale Bibliographie der Geistes- und Sozialwissenschaftlichen Zeitschriftenliteratur, Internationale Bibliographie der Rezensionen Geistes- und Sozialwissenschaftlicher Literatur, Linguistics and Language Behavior Abstracts, Mathematical Reviews, MathSciNet, Periodicals Contents Index, Philosopher's Index, Repertoire Bibliographique de la Philosophie, Russian Academy of Sciences Bibliographies.

Contents