A formalization of Kant's transcendental logic

Achourioti, T.; van Lambalgen, M.

Published in:
Review of Symbolic Logic

DOI:
10.1017/S1755020310000341

Link to publication

Citation for published version (APA):
Achourioti, T., & van Lambalgen, M. (2011). A formalization of Kant's transcendental logic. *Review of Symbolic Logic*, *4*(2), 254-289. https://doi.org/10.1017/S1755020310000341

The Review of Symbolic Logic
Volume 4, Number 2, June 2011

A FORMALIZATION OF KANT'S TRANSCENDENTAL LOGIC

T. ACHOURIOTI AND M. VAN LAMBALGEN
ILLC/Department of Philosophy, University of Amsterdam

Abstract. Although Kant (1998) envisaged a prominent role for logic in the argumentative structure of his *Critique of Pure Reason*, logicians and philosophers have generally judged Kant's logic negatively. What Kant called 'general' or 'formal' logic has been dismissed as a fairly arbitrary subsystem of first-order logic, and what he called 'transcendental logic' is considered to be not a logic at all: no syntax, no semantics, no definition of validity. Against this, we argue that Kant's 'transcendental logic' is a logic in the strict formal sense, albeit with a semantics and a definition of validity that are vastly more complex than that of first-order logic. The main technical application of the formalism developed here is a formal proof that Kant's Table of Judgements in Section 9 of the *Critique of Pure Reason*, is indeed, as Kant claimed, complete for the kind of semantics he had in mind. This result implies that Kant's 'general' logic is after all a distinguished subsystem of first-order logic, namely what is known as geometric logic.

§1. Introduction. The purpose of this article is to examine from the point of view of mathematical logic, Kant's formal logic[1] and its relation to what Kant called 'transcendental logic'. This may seem to be a hopeless enterprise. A typical modern dismissal of Kant's *formal* logic is: 'terrifyingly narrow-minded and mathematically trivial'[2] and of course every student of logic is aware of the more politely expressed verdicts in Frege's *Begriffsschrift* and Strawson's *The Bounds of Sense*. Worse, Kant's *transcendental* logic does not seem to be a logic in the modern sense at all: no syntax, no semantics, no inferences.

As against the received view, we will argue here that (i) Kant's formal logic is overly restrictive only if it is assumed that its underlying semantics is that of classical first-order logic, whereas Kant's implied semantics is very different, (ii) the implied semantics, centered around Kant's three different notions of object, can be given a precise mathematical expression, thus leading to a formalized transcendental logic, and (iii) on the proposed semantics, a logical system very much like Kant's formal logic is a distinguished fragment of first-order logic, namely, geometric logic.

What could be the purpose of such a, partly very technical, exercise? To say that the *Critique of Pure Reason*[3] is notoriously difficult is an understatement. In his attempt to characterize our cognition of objects, Kant employs a vast array of technical terms that are insufficiently defined (at least for readers with an analytical mind-set), such as 'judgement',

Received: November 5, 2009

[1] As expounded briefly in the *Critique of Pure Reason* (Kant, 1998) and in more detail in the various series of course notes produced by Kant's students, and the *Jäsche Logik*, commissioned by Kant and compiled from his notes by J.B. Jäsche; for all of which see Kant (1992). We cite the *Critique of Pure Reason* in the standard manner, by referring to the page numbers of the first (A) and second (B) edition, so for example, 'A57/B81').

[2] Hazen, A. (1999). Logic and analyticity, In Varzi, A. C., editor. *The Nature of Logic*, Stanford, CA: CSLI, pp. 79–110.

[3] Henceforth abbreviated to *CPR*.

doi:10.1017/S1755020310000341

'transcendental unity of apperception', 'transcendental object', 'synthesis', and 'schema'. Perhaps a mathematical formalization, however incomplete, can shed some light on these concepts and their relations. We are very far from claiming that the formalization presented exhausts Kant's concepts; there is hardly a better inducement to modesty than trying to come to grips with the complexities of *CPR*, not to mention the secondary literature. But the formalization may provide a starting point.

One important aim is to attempt to restore logic to its rightful place in the argumentative structure of *CPR*. Kant himself was clear about the importance of logic; the so-called Metaphysical Deduction (Section 10 of *CPR*) uses the presumed completeness of the Table of Judgements to argue for the completeness of the Table of Categories, the a priori concepts that are supposed to underlie our cognition of objects and their relations. The *communis opinio* is that the Table of Judgements rests on a thoroughly discredited view of logic, and this has led to many reductive readings of *CPR* in which logic hardly figures, if at all. One notable exception to this trend is Longuenesse's (1998) *Kant and the Capacity to Judge*, whose focal point is the connection between the logical forms of judgements and acts of the understanding. In fact, reading Longuenesse's commentary made us aware that there is much more to Kant's logic than meets the eye, and the present article can be seen as an attempt to translate some of her insights into logical terms.[4]

To our knowledge, ours is the first attempt to shed light on Kant's logic using the tools of modern mathematical logic. There exist overviews, such as Stuhlmann-Laeisz' *Kants Logik* (Kitcher, 1976), but these are set firmly in the context of traditional logic. There are also studies of specific philosophical aspects, such as MacFarlane's (2000) dissertation on Kant's use of the term 'formal'; or Posy (2003) on the question whether Kant is guilty of psychologism. A recent article on Kant's transcendental logic, Rosenkoetter's (2009) 'Truth criteria and the very project of a transcendental logic' interestingly suggests that transcendental logic is bound up with the search for a truth criterion that traditional logic cannot deliver (A57–9/B82–4), but there is no hint that transcendental logic can be conceived as a formal logic in its own right. This is precisely what we shall attempt. Before we delve into the intricacies of Kant scholarship and the sometimes considerable complexity of the formalization of transcendental logic, we want to summarize why we think such an exercise can contribute to a better understanding of the first *Critique*.

(i) Kant describes the role of transcendental logic as

> The part of transcendental logic that expounds the elements of the pure cognition of the understanding and the principles without which no object can be thought at all, is the transcendental analytic, and at the same time a logic of truth. For no cognition can contradict it without at the same time losing all content, i.e. all relation to any object, hence all truth. (A62–3/B87)

The word 'logic' need not be understood metaphorically in this passage, neither did Kant intend it as such: his discussion of logic in A50–7/B74–82 amply shows this. From a modern perspective, what is stated here is a somewhat outlandish completeness theorem, as for instance that for Friedman's quantifier 'almost all'

[4] This is not to say that Longuenesse (1998) is the only work taking the connection between the logical forms of judgements and acts of the understanding seriously. Two other authors that come to mind are Reich (1932) and Wolff (1995). Here we shall not discuss the differences in interpretation between these three authors.

(cf. Steinhorn, 1985): a sentence is consistent with the Friedman axioms if and only if it has a Borel model, that is, a model where the objects are reals,[5] and the relations are interpreted as (uncountable) Borel subsets of (finite products of) the reals. The situation is completely analogous here: in both cases one needs consistency with a given set of axioms to force the language to have the interpretation one is interested in. Since transcendental logic plays such a huge role in the Transcendental Analytic, it is worth the effort to see whether a formalization can clarify parts of the argument; much of the effort will go into the formalization of Kant's three notions of object.

(ii) A second potential benefit is conceptual clarification. As an example we may take the notion of judgement, which receives a different characterization in each of *CPR*, *Prolegomena*, and *Jäsche Logik*. The formalization will illuminate each of these characterizations and show that in the end they are identical.

(iii) There is however a potentially damning objection to this enterprise. As will be seen, the mathematical techniques used in proving the completeness theorem for transcendental logic and the completeness of the Table of Judgements are thoroughly modern, without any relation to what Kant could have known. Doesn't this invalidate the whole enterprise? What sense is there in attributing a logic to Kant whose formal properties are so far beyond what was available to him? We believe a defence can be mounted along the following lines. As an example, consider Kant and his treatment of the continuum, which is striking because he does not consider it to be composed of points, as a set theoretic treatment would have it. Consistent with a constructive treatment, points in the Kantian continuum are only boundaries or limitations. Kant reasons quite consistently with this conception, even though its formal treatment, for example, in terms of locales, was of course not available to him. Similarly, we believe that Kant's intuition about objects come very close to a mathematical structure (inverse systems), even though he could not articulate it in this way. These structures will provide a semantics for the language of transcendental logic, and as we will see this part is neither very sophisticated nor does it depart much from what Kant himself writes. Where we really go way beyond Kant is in using this semantics to prove limitative results, such as that the Table of Judgement is complete, as Kant conjectured (or rather, asserted). Still, one can view the formalization as providing a model of Kant's thinking, thus showing its consistency and coherence. To us, that seems more profitable than chopping off a large part of Kant's thought because it is presumed to be hopelessly outdated.

§2. Kant's aims and the role of logic. Kant's overall aim in *CPR* was to delimit the scope for metaphysical reflection by arguing that (i) all knowledge starts with perceptual experience, (ii) our cognition contributes principles structuring perceptual experience, principles which are themselves not derived from sensory experience (space, time, and the *Categories*), (iii) all our suprasensory knowledge is confined to these principles, which Kant called 'synthetic a priori', meaning that they are informative but that their truth can be ascertained without recourse to experience, (iv) it is possible to give a complete

[5] That is, these objects are much more special as compared to what one gets from the completeness theorem for first-order logic.

enumeration of the Categories and the associated synthetic a priori principles, and this yields a precise delimitation of metaphysics.

At this point, it may be instructive to compare Kant to modern cognitive science, in which the idea that sensory experience needs structuring principles not themselves directly derived from experience is fairly prominent. The classic statement of this view comes from Bartlett (1968, p. 1):

> Whenever anybody interprets evidence from any source, and his interpretation contains characteristics that cannot be referred wholly to direct sensory observation or perception, this person thinks. The bother is that nobody has ever been able to find any case of the human use of evidence which does not include characters that run beyond what is directly observed by the senses. So, according to this, people think whenever they do anything at all with evidence. If we adopt that view we very soon find ourselves looking out upon a boundless and turbulent ocean of problems.

In his famous paper 'Beyond the information given', Bruner (1973, p. 224) took this idea as a starting point, and argued for the importance of *coding systems*:

> We propose that when one goes beyond the information given, one does so by being able to place the present given in a more generic coding system and that one essentially reads off from the coding system additional information [. . .].

Examples of such coding systems abound in vision, mostly of a geometrical kind, for example, it is assumed that real-world surfaces can be described as smooth manifolds in the mathematical sense (the coding system). This brief excursion should suggest that the basic outlines of Kant's approach to cognition have contemporary relevance. For a much more extensive treatment, the reader should consult Kitcher (Stuhlmann-Laeisz, 1990).

We now come to the role that Kant envisaged for logic in this whole scheme. For Kant, all cognitive activity is directed toward *judgement*. This crucial term is easily misunderstood, especially when one first approaches it through the 'Table of Judgements', which seems to list possible syntactic forms of sentences (universal, particular, singular, hypothetical, disjunctive, etc.). But for Kant, judgement is first and foremost a cognitive act, the result of which has one of the listed syntactic forms.[6] This act-like character of judgement is clear if one looks at Kant's definitions of judgement:

> A judgement is the representation of the unity of the consciousness of various representations ... (*Jäsche Logik*, Section 17)
> A judgement is nothing but the manner in which given cognitions are brought to the objective unity of apperception (B141).

Common in these definitions is that a judgement is the act of binding together mental representations; this is what the term 'unity' refers to. But where the second definition differs from the first is that also the aim of judgement is indicated by means of the word 'objective', which is Kant's terminology for 'having relation to an object'.[7] We will return below to the multiple meanings of the term 'object', but for now it suffices to think of objects of experience. For Kant, objects are not found in experience, but they are constructed

[6] Thus a Kantian judgement is very much unlike a proposition in the modern sense.
[7] Kant also expresses this by saying that a judgement must be *objectively valid*.

('synthesized') from sensory materials under the guidance of the Categories, which are defined as 'concepts of an object in general, by means of which the intuition of an object is regarded as determined in respect of one of the logical functions of judgement' (B128). It is here that judgement plays an all-important role, since Kant's idea is that objects are synthesized through the act of making judgements about them. An example will help here. Consider the universal judgement 'All A are B', which Kant glosses with the help of a quantifier as 'To everything x to which A belongs, also B belongs'. Kant inquires what the presuppositions of such judgements are, and finds that these include the concepts of *unity* (to provide units that can act as substitution instances for the variable x), and *totality* (to ensure that x can run over the *whole* extension of A).[8] Thus, the conditions that have to be satisfied to make a universal judgement lead one to consider the Categories of Quantity: unity, plurality, totality. More generally, Longuenesse (1998, p. 78) writes:

> Kant asked himself which logical forms of judgement should be considered primitive *if* the *original* function of judgement is to "bring given cognitions to the objective unity of apperception", that is, to relate our representations to objects. Having discovered that the objectifying function of these forms provided him with the solution he was seeking to the problem of the categories,[9] Kant retained as primitive only those forms which he thought indispensable for generating the relation of our representations to an object ....

Thus Kant's argument is the following. Cognitive activity consists in binding together representations, and always issues in judgements. The mode of binding does not derive from the senses, but is contributed a priori by cognition (compare Bruner's 'coding systems'). Each primitive judgement form represents a particular a priori manner of binding together, and conversely, so that we have a one–to–one correspondence between judgement forms and Categories. Assuming that the list of judgement forms is complete, the corresponding list of Categories is complete as well. This is the content of the 'Metaphysical Deduction', which, if successful, would provide an exhaustive list of those concepts that legitimately go beyond sensory experience.

Starting with Section 4 in Frege's *Begriffsschrift*, Kant's logic came under heavy criticism. Viewed from the perspective of the newly developed predicate logic, Kant's logic seemed a tiny fragment, devoid of special significance; and of course, Frege's antipsychologism was inimical to associating cognitive operations to judgement forms. Another discussion of 'the trouble with Kant' can be found in Strawson's *The Bounds of Sense*, which focusses on Kant's aim to identify primitive logical forms, and to associate these with primitive functions of the understanding. At least in classical logic, it is impossible to talk about a distinguished set of primitive logical forms, since judgement forms are interdefinable (e.g., the hypothetical in terms of the disjunctive and negation).[10] Strawson

[8] While running over the extension of A, x describes a *plurality*, which may or may not be totalisable.

[9] The problem being how to provide a complete list of the Categories.

[10] At this point one can begin to see what could be wrong with this line of criticism: it is very much tied to classical logic, and would fail for example, intuitionistic logic. It is quite doubtful, to say the least, whether Kant adopts a material implication reading of the hypothetical. It is even unclear whether his disjunctive judgement, with its intended semantics in terms of parts and wholes, fully conforms to the classical model; see Footnote 41 for more on this.

observes that if one abandons the idea of primitive logical forms, one could say that each of indefinitely many logical forms gives rise to a Category, which seems absurd; or one could say that it is logical *ideas* and not logical forms that are fundamental, in particular

1. the formally atomic proposition (which for Strawson is in subject–predicate form)
2. truth-functional composition[11]
3. quantification.

This then greatly reduces the number of Categories; what one is left with is that the notions of particular object and universal kind or character as categories must have application in a world in which judgements can be made. More importantly, these two categories can be directly deduced from transcendental considerations, without the detour via the Table of Judgements.

All in all it seems then as if Kant's attempt to involve logic in determining 'the bounds of sense' is an utter failure; so why try to resuscitate it? One reason for doing so is that the various criticisms have completely bypassed issues of semantics. It has been taken for granted that Kant's judgement forms can be translated in classical predicate logic, and thus inherit its semantics. But a moment's reflection shows that this may well be inappropriate. The semantics for classical predicate logic is given by models consisting of a domain of objects over which relations are interpreted; thus here the objects are assumed to be *given*. We have just seen however that for Kant, judgements somehow play a role in constituting objects from sensory material, so that it seems wrong to take objects as given from the outset. Could it be that by constructing a semantics that is closer to Kant's intentions, the judgement forms that he considers somehow do acquire a special status?

§3. Our aims. Kant determined the judgement forms in the Table of Judgements not by looking at what the logic textbooks of the day had to offer, but

> Only insofar as [judgement forms] were selected and organized under the guidance of the transcendental question (what are the a priori conditions for the representations of objects in general?) could logical functions of judgement serve as a guiding thread to the table of categories (Longuenesse, 1998, p. 397).

Kant's procedure can be thought of along the lines of the 'method of analysis and synthesis' in geometry: in the analytic phase one assumes that one knows precisely how objects are constituted and one determines from this the logical forms of judgements that can have relation to an object, in so doing providing a transcendental justification of the Table of Judgements; in the synthetic phase (which takes up most of the Transcendental Deduction) one reverses the argument and shows how logical forms of judgement contribute to the constitution of objects.

Our argument is focussed on the first, 'analytic', part. We consider a subset of 'the a priori conditions for the representations of objects in general' (as discussed by Kant) that is sufficient to generate an abstract mathematical representation of Kant's view of objects—as will be seen below for this purpose we use (abstract descriptions of) the three syntheses of the A Deduction (A99–104) and the 'unity of self-consciousness' (A109). The mathematical representation of Kant's view of objects then provides a semantics for

[11] But see Footnote 10.

judgements that allows one to ask the question: '*which* logical forms of judgements are capable of having 'relation to an object'?'. Our main technical result is that this class of judgement forms consists of all and only those formulas that can be written as conjunctions of geometric implications, that is, of formulas of the form $\forall \bar{x}(\theta(\bar{x}) \rightarrow \psi(\bar{x}))$, where θ is a conjunction of atomic formulas, and ψ is constructed from atomic formulas using only $\vee, \wedge, \exists, \bot$.[12] No judgement form outside this class can claim objective validity.[13] In this sense there is a completeness result like the one Kant claimed for his Table of Judgements; and it is similarly based on transcendental considerations. At first sight it may seem that the class of judgement forms we have isolated *extends* the Table of Judgements: after all, it allows (some) quantifier combinations of the form $\forall\exists$, whereas Kant's Table apparently doesn't. We shall argue, however, that this difference is spurious, because Kant's examples, especially in his treatment of causality, show that what he calls a 'hypothetical' judgement can be as complex as geometric implications. And clearly the paradigm example of a synthetic a priori judgement

> Everything that happens (begins to be) presupposes something which it follows in accordance with a rule. (A189)

is a geometric implication.

Our main mathematical result does not answer the 'synthetic' question: '*how* do judgement forms enable us to constitute empirical objects as substitution instances of the "x" in judgements ?'.[14] This involves looking at judgement forms, to assess their contribution to the constitution of objects. But before the *how*-question can be answered, first those judgement forms need to be identified for which the *how*-question can be asked at all.

§4. Kant's logic: general and transcendental.

The key to such an enterprise is that judgements in Kant's sense participate in two kinds of logics: *general* logic and *transcendental* logic.

4.1. General logic.

> [G]eneral logic abstracts from all the contents of the cognition of the understanding and of the difference of its objects, and has to do with nothing but the mere form of thinking. (A54/B78)

Kant emphasizes time and again the *formality* of general logic, which contains 'absolutely necessary rules of thinking regardless of difference of objects'. In particular, it abstracts from empirical content and makes no distinction between imaginary and real or actual objects. Kant's general logic is treated in the notes taken by his students and the *Jäsche Logik* and treats such topics as:[15]

1. forms of judgements (categorical, hypothetical, disjunctive, with various subdivisions)

---

12 The relevant definitions will be given in Section §9.

13 For the meaning of this term, compare Footnote 7.

14 Compare p. 258 for Kant's introduction of bound variables.

15 We do not discuss Kant's treatment of concepts here.

2. the inferences in which these judgements participate (immediate inferences such as conversion of subject and predicate; and inferences of reason such as syllogisms, *modus ponens*/*modus tollens*, disjunctive syllogism)
3. the principle on which inferences of reason are based: the major premise must be a 'universal rule' and the minor premise must express subsumption under the condition of the rule.

Categorical judgements are judgements in subject–predicate form, combined with quantifiers and optional negation: (*affirmative*) all A are B, some A are B, (*negative*) no A are B, some A are not B, (*infinite*) all A are non-B, some A are non-B, *etcetera*.[16] Categorical judgements may occur as premises and as conclusion in syllogisms.

Disjunctive judgements are not what one would think, judgements of the form $p \vee q$. The *Jäsche Logik* provides the following definition:

> A judgement is disjunctive if the parts of the sphere of a given concept determine one another in the whole or toward a whole as complements... [A]ll disjunctive judgements represent various judgements as in the community of a sphere... (*Jäsche Logik*, Section 27, Section 28) (Kant, 1992, pp. 602–603)

As example Kant provides

> Every triangle is either right-angled or not right-angled.

Thus the logical form is something like $\forall x(C(x) \to A(x) \vee B(x))$, where C represents the whole, A, B its parts; here it is not immediately clear whether the parts can be taken to exist outside the context of the whole. In our formalization, we will act as if the $\vee$ is classical disjunction; but see Footnote 41.

It would similarly be a mistake to identify Kant's hypothetical judgements with a conditional $p \to q$, let alone material implication as defined by its truth table. A few examples from Kant will make this clear. In the context of a discussion of the possible temporal relations between cause and effect Kant writes in *CPR*:

> If I consider a ball that lies on a stuffed pillow and makes a dent in it as a cause, it is simultaneous with its effect. (A203/B246)

The hypothetical that can be distilled from this passage is

> If a ball lies on a stuffed pillow, it makes a dent in that pillow.

From this we see that Kant needs relations, not just unary predicates, to formulate his causal hypotheticals, and also that the logical form of this particular causal hypothetical is what we called above (p. 260) geometric implications.

We now give an extended quote from the *Prolegomena* Section29 which provides another example of a hypothetical judgement that really has geometric implication form, and which also shows that causality is related to the rule-like character of the hypothetical, which, as we shall see in Section 9.1, means that the existential quantifier(s) in the geometric implication can be replaced by a concrete function.

[16] Kant thus makes a distinction between sentence negation (negative judgements) and predicate negation (infinite judgements).

> It is, however, possible that in perception a rule of relation will be found, which says this: that a certain appearance is constantly followed by another (though not the reverse); and this is a case for me to use a hypothetical judgement and, e.g., to say: If a body is illuminated by the sun for long enough, it becomes warm. Here there is of course not yet the necessity of connection, hence not yet the concept of cause. But I continue on, and say: if the above proposition, which is merely a subjective connection of perceptions, is to be a proposition of experience, then it must be regarded as necessarily and universally valid. But a proposition of this sort would be: The sun through its light is the cause of the warmth. The foregoing empirical rule is now regarded as a law, and indeed as valid not merely of appearances, but of them on behalf of a possible experience, which requires universally and therefore necessarily valid rules [...] the concept of a cause indicates a condition that in no way attaches to things, but only to experience, namely that experience can be an objectively valid cognition of appearances and their sequence in time only insofar as the antecedent appearance can be connected with the subsequent one according to the rule of hypothetical judgements (Kant, 2002, p. 105).

The logical form of the first hypothetical is something like

> If x is illuminated by y between time t and time s and $s - t > d$ and the temperature of x at t is v, then there exists a w such that the temperature of x at s is $v + w$ and $v + w > c$,

where d is the criterion value for 'long enough' and c a criterion value for 'warm'. Again we see the importance of relations in the logical form of hypotheticals, which is that of geometric implications.[17] One last remark before we move on to inferences. Kant makes a distinction between categorical judgements 'To all x to which A belongs, also B belongs' and hypothetical judgements 'if x is A then x is B'. The difference is that in the latter judgement type it is not asserted that x belongs to A, whereas that is asserted in the former type. This difference can be accounted for by working in a many-sorted logic where the subject term of a categorical judgement refers to a separate sort, always taken to be nonempty. This does not affect the technical results given below.

Moving now to inferences, we see that general logic contains a specific set of inference rules, not a general semantic consequence relation. Simplifying a bit, one can say that the inference rules are like elimination rules in natural deduction. Corresponding introduction rules are lacking; for example, there are no formal introduction rules for the hypothetical or the disjunctive.[18] The conditions for asserting a judgement are not treated in general logic, but in transcendental logic, about which more will be said below. Before we do so

[17] A full treatment of causality would involve structures containing events and times (as well as axioms governing these), whereas below we consider only the case where quantifiers range over objects. Nevertheless we hope the message is clear: there is considerable logical complexity (including quantifiers) hidden in Kant's notion of a hypothetical judgement.

[18] In fact the usual introduction rule for $\vee$ would be incorrect on Kant's interpretation of the disjunction, since it is tantamount to the possibility to create wholes by arbitrarily putting together parts.

however, we add a few remarks on Kant's formal theory of truth, which also belongs to general logic.
Kant lists three 'formal criteria of truth' (*Jäsche Logik*, Kant, 1992, p. 560):

(1) the principle of [non-]contradiction
(2) the principle of sufficient reason 'on which rests the (logical) *actuality* of a cognition, the fact that it is grounded, as material for *assertoric* judgements'
(3) the principle of excluded middle.

The second principle seems out of place in logic, especially if one reads what Kant writes on the previous page

> *If all the consequences of a cognition are true, then the cognition is true too*. For if there were something false in the cognition, then there would have to be a false consequence too.
> From the consequence, then, we may infer to a ground, but without being able to determine this ground. Only from the complex of all consequences can one infer *to a determinate ground*, infer that it is the true ground.

This seems a blatant fallacy, but the principle will make a surprising appearance in our formal treatment of Kant's transcendental logic.

4.2. Transcendental logic. We began this section by saying that judgements participate in two logics simultaneously: general and transcendental. About the latter, Kant says the following:

> a science of pure understanding and of the pure cognition of reason, by means of which we think objects completely a priori. Such a science, which would determine the origin, the domain, and the objective validity of such cognitions, would have to be called transcendental logic since it has to do merely with the laws of the understanding and reason, but solely insofar as they are related to objects a priori and not, as in the case of general logic, to empirical as well as pure cognitions of reason without distinction. (A57/B81–2)

Roughly speaking, whereas general logic is concerned merely with inference rules, what transcendental logic adds to this is a semantics in terms of *objects*. At first sight this may seem rather trivial: after all, first-order predicate logic has a semantics in terms of domains of objects and relations defined on these domains, and thus there seems to be no ground for our earlier claim, when discussing Frege's and Strawson's criticism, that Kant's logic is different from first-order predicate logic. But there is a vast difference between the notion of object as it occurs in modern semantics and in Kant's logic. In the former, objects are mathematical entities supplied by the metatheory, usually some version of set theory. These objects have no internal structure, at least not for the purposes of the model theory. Kant's notions of object, as they occur in the semantics furnished by transcendental logic, are very different. For instance, there are 'objects of experience', somehow constructed out of sensory material by 'the laws of the understanding and reason'. Transcendental logic deals with a priori and completely general principles which govern the construction of objects, and relate judgements to objects so that we may come to speak of *true* judgements. Since transcendental logic deals with construction of objects of experience, for the purposes of *this* logic objects must have internal structure; and a large part of this article will be taken up by a discussion of what that structure is.

§5. 'The rules for the pure thinking of an object'. There appears to be an internal inconsistency in Kant's notion of an object, which follows from his treatment of causality. Kant was convinced by Hume's analysis of causality, showing that it cannot be inherent in the phenomena. Kant's proposal was to make causality a category, a 'pure concept of the understanding', which is instrumental in constituting phenomena but which cannot be derived from them. Now this seems to conflict with the natural tendency to say that representations within us are *caused* by objects outside us, that 'objects affect sensibility'. For in this case one extends the category of causality beyond its legitimate domain of application. We will adopt Longuenesse's analysis here. She proposes, following A92/B125, that Kant replaces the necessity inherent in causality by another form of necessity: instead of 'the object causes representations within us', one now has 'the object alone makes the representation possible', or 'without the representation the object is impossible'. Moreover, and most importantly, this form of necessity is internalized, what Longuenesse calls 'internalisation within representation', so that 'the object' as it occurs in the statement 'without the representation the object is impossible' is itself a kind of representation. Here is a quote from *CPR* making this point:

> It is necessary to make understood what is meant by the expression an object of representations. [A]ppearances themselves are nothing but sensible representations, which must not be regarded in themselves ... as objects (outside the power of representation). What does one mean, then, if one speaks of an object corresponding to and therefore also distinct from the cognition? It is easy to see that this object must be thought of only as something general = X, since outside of our cognition we have nothing that we could set over against this cognition as corresponding to it. (A104)

We thus have the following slightly paradoxical situation: we have to relate appearances to objects represented internally as existing outside us. How Kant conceived of this remarkable feat will be our next topic.

§6. Objects of appearance, objects of experience, transcendental objects. Longuenesse's interpretation hinges on an analysis of what Kant might mean when he glosses a categorical judgement such as 'all A are B' as 'To all x to which A belongs, B also belongs'. It is not open to Kant to let the variable x run over a particular set theoretically defined collection of objects; x must run over objects as he conceives them, that is, as particular internal representations. But now the problem arises that Kant has several notions of object: object of appearance, object of experience, and transcendental object, and in some way the variable x is concerned with all three. In a footnote defending her interpretation against Allison's, Longuenesse (1998, p. 111) puts this as follows:

> [R]eference to an object represented by the term 'x' in the logical form of judgement does have as one of its components reference to an independent object leading us to seek coherence among our representations of it and allowing us to think of the connection of representations as in some sense necessary, "even if the judgement is empirical and therefore

> contingent", which is precisely the role Kant assigned to reference to the "transcendental object = x" in the A Deduction.

This seems to be saying that we should think of a variable x as referring to different types of objects simultaneously. This is a very unusual understanding of variables, and it will be one of our main tasks to make this idea precise. In order to do so, we will briefly introduce the three kinds of objects.

In Section 2 we remarked that Kant's theory of cognition can be captured by Bruner's phrase 'going beyond the information given'. Kant attempts to show in the A Deduction that the psychological data philosophers of empiricist persuasion typically assume are already the result of mental operations. Receptivity by itself is not enough, but what it yields must be *processed* in order to be used by the understanding, which is concerned with relating cognitions to objects. The general term for these processes is *synthesis*:

> By synthesis in the most general sense ... I understand the action of putting different representations together with each other and comprehending their manifoldness in one cognition. (A77/B103)

6.1. Synthesis of apprehension in intuition. The A Deduction treats three such syntheses, which to a first approximation can be thought of as operating one after the other. The purpose of the first synthesis, the 'synthesis of apprehension in intuition', is to prepare the manifold received in intuition for further processing by representing it internally as a manifold of spatial parts and sensory qualities, together with the manner in which these are bound together:

> Every intuition contains a manifold in itself, which however would not be presented as such if the mind did not distinguish the time in the succession of impressions on one another; for as contained in one moment no representation can ever be other than absolute unity. Now in order for unity of intuition to come from this manifold ... it is necessary first to run through and then take together this manifoldness, which action I shall call the synthesis of apprehension ...(A99)

A concrete example of this process of 'run through and take together' is given later, in B162, where Kant envisages 'mak[ing] the empirical intuition of a house into perception through apprehension of its manifold' by, as it were, drawing its shape, for example, by a temporal process.

Although this first synthesis has made much material available for further processing (spatial parts, sensory qualities), by itself it does not produce objects (always thought of as internal representations); it does not distinguish between true cognition of objects and hallucinations. For this, additional mental processes are necessary, which are concerned with various forms of stability that objects must have.

6.2. Synthesis of reproduction in imagination. To begin with, experience presupposes the reproducibility of appearances, as can be seen from the following example, especially when considered in conjunction with the example of the house (B162) just given:

> Now it is obvious that if I draw a line in thought, or think of the time from one noon to the next, or even want to represent a certain number to myself, I must necessarily first grasp one of these manifold representations after another in my thoughts. But if I were always to lose the preceding

> representations (the first parts of the line ...) from my thoughts and not reproduce them when I proceed to the following ones, then no whole representation ... could ever arise. (A102)

Thus if one is engaged in mentally tracing the outline of the back of a house but in the meantime the representation resulting from tracing the front has dropped out of memory, no experience of the house will ever arise. In other words, to experience an object *qua* object, at any particular moment sensory experience has to go hand in hand with the reproduction of those aspects of the object that cannot be perceived at that very moment. Longuenesse (1998, p. 42) puts this as follows: '[T]he *aim to represent a whole* guides every associative reproduction of the imagination.' It is only through the synthesis effected by reproduction in imagination that we can strive toward complete knowledge, because at any particular moment sensory knowledge is limited. At the same time the *aim to represent a whole* is precisely that, an aim, because complete knowledge is beyond our reach.

The preceding two syntheses (which are 'inseparably combined' (A102)) yield objects of appearance. In the formal model to be presented in Section 7 objects of appearance will constitute the domains of models. Predicates are introduced through the following considerations.

6.3. Synthesis of recognition in a concept.

> Without consciousness that which we think is the very same as what we thought a moment before, all reproduction in the series of representations would be in vain. For it would be a new representation in our current state. (A103)

It is the concept's job to generate the *consciousness of the generic identity between the reproduced representation and all those preceding it* (Longuenesse, 1998, p. 46), that is, to recognize a representation as being of a certain kind. Kant has an interesting dual view of concepts: they are on the one hand intensional rules (or algorithms) to be applied to objects of appearance, on the other hand the extensions generated by the applications of the rules. If objects of appearance can be stably classified as being of a certain kind, they are called 'determinate objects of experience', where 'determinate' here means 'determined by concepts'.

6.4. From transcendental unity of apperception to transcendental object. We are not yet done, however; there is another kind of stability we have to consider. Even though all our representations are internal, there is sometimes a stable coherence among them, and this leads us to think that there is an object 'behind' the representations that is responsible for the coherence. But we can know nothing about objects 'behind' representations, and the question arises how such an object can be internalized as a particular kind of representation. The answer is given in the following long sequence of quotes:

> We find however, that our thought of the relation of all cognition to its object carries something of necessity to it [...] since insofar as they are to relate to an object our cognitions must also necessarily agree with each other in relation to it, i.e., they must have that unity that constitutes the concept of an object. It is clear however, that since we have to do only with the manifold of our representations, and that X which corresponds with them (the object) [...] is nothing for us, the unity that the object makes necessary can be nothing other than the

> formal unity of the consciousness in the synthesis of the manifold of representations. (A104–5)

> Now no cognitions can occur in us, no connection and unity among them, without that unity of consciousness that precedes all data of the intuitions, and in relation to which all representation of objects is alone possible. This pure, original, unchanging consciousness I will now name transcendental apprehension. (A107)

> [A]ppearances are not things in themselves, but themselves only representations, which in turn have their object, which therefore cannot be further intuited by us, and that may therefore be named the non-empirical, i.e., the transcendental object = X. The pure concept of this transcendental object (which in reality throughout all our cognition is always one and the same = X), is what can alone confer upon all our empirical concepts in general relation to an object, i.e., objective validity. [T]his concept cannot contain any determinate intuition at all, and therefore concerns nothing but that unity which must be encountered in a manifold of cognition insofar as it stands in relation to an object. This relation, however, is nothing other than the necessary unity of self-consciousness, thus also of the synthesis of the manifold, through a common function of the mind for combining it in one representation. (A109)

Hence, on the one hand the transcendental object is responsible for grasping objects of experience as the same object, on the other hand the transcendental object is an internal representation of the 'necessary unity of self-consciousness'. To come back to the starting point of this section, the variable x must refer to objects of experience as well as to the transcendental object, and the latter must somehow be constructed from the 'unity of self-consciousness'.

We seem to have strayed very far from logical considerations. It will turn out however, that it is possible to give a formalization of transcendental logic that captures the formal relationships that are hinted at by Kant in the preceding quotes. The first step toward formalization will be to propose semantic structures that can replace first-order models with their inappropriate notion of object; these will be inverse systems of first-order models and their limits. We will then show how to interpret formulas on such structures, in so doing making sense of the multiple interpretations of the same variable that Longuenesse believes is essential. Once we have done so, we can set up a proof system for transcendental logic and prove its soundness and completeness. A by-product of the completeness proof is that there is a sense in which Kant's Table of Judgements is complete.

§7. Structures for objects. The purpose of the structure to be presented is to give an account of Kant's notion of object, in terms of formal counterparts of the three syntheses, the transcendental object, and unity of self-consciousness. This leaves out important topics of Kant's transcendental logic, notably his theory of causality. The advantage of this restriction is that in our chosen domain, the structures necessary for formalization already exist, namely inverse systems.

To make the complexity of the construction somewhat more manageable, we split it in two parts. We first deal with inverse systems of sets and the relation of these inverse systems to the first two syntheses. We then consider inverse systems of models, that is, of sets with relations defined on them, and we relate these to the third synthesis. We emphasize that we introduce this division only for expository purposes; in reality the three syntheses are inseparably linked.

7.1. Inverse systems of sets and synthesis. An inverse system is a collection of sets which are connected by mappings. We first give the pertinent definitions, then relate these to abstract properties of synthesis.

DEFINITION 7.1. *A directed set is a set T together with an ordering relation $\leq$ such that*
(1) $\leq$ is a partial order, that is, transitive, reflexive, antisymmetric
(2) $\leq$ is directed, that is, for any $s, t \in T$ there is $r \in T$ with $s, t \leq r$.

DEFINITION 7.2. *An inverse system indexed by T is a set $\mathcal{D} = \{D_s | s \in T\}$ together with a family of mappings $\mathcal{F} = \{h_{st} | s \geq t, h_{st} : D_s \longrightarrow D_t\}$. The mappings in $\mathcal{F}$ must satisfy the coherence requirement that if $s \geq t \geq r$, $h_{tr} \circ h_{st} = h_{sr}$.*

7.1.1. Interpretation of the index set In our case, the index set represents some abstract properties of synthesis.[19] Recall that the 'synthesis of apprehension in intuition' proceeds by a 'running through and holding together of the manifold' and is thus a process that takes place in time. We may now think of an index $s \in T$ as an interval of time available for the process of 'running through and holding together'. More formally, s can be taken to be a set of instants or events, ordered by a 'precedes' relation; the relation $t \leq s$ then stands for: t is a substructure of s. It is immediate that on this interpretation $\leq$ is a partial order. The directedness is related to what Kant called 'the formal unity of the consciousness in the synthesis of the manifold of representations' (A105) or 'the necessary unity of self-consciousness, thus also of the synthesis of the manifold, through a common function of the mind for combining it in one representation' (A109)—the requirement that 'for any $s, t \in T$ there is $r \in T$ with $s, t \leq r$' creates the formal conditions for combining the syntheses executed during s and t in one representation, coded by r.[20]

7.1.2. Interpretation of the D_s and the mappings $h_{st} : D_s \longrightarrow D_t$ An object in D_s can thought of as a possible 'indeterminate object of empirical intuition' synthesized in the interval s. If $s \geq t$, the mapping $h_{st} : D_s \longrightarrow D_t$ expresses a consistency requirement: if $d \in D_s$ represents an indeterminate object of empirical intuition synthesized in interval s, so that a particular manifold of features can be 'run through and held together' during s, some indeterminate object of empirical intuition must already be synthesizable by 'running through and holding together' in interval t, for example, by combining a subset of the features characterizing d. This interpretation justifies the coherence condition $s \geq t \geq r$, $h_{tr} \circ h_{st} = h_{sr}$: the synthesis obtained from first restricting the interval available for 'running through and holding together' to interval t, and then to interval r should not differ from the synthesis obtained by restricting to r directly.

We do not put any further requirements on the mappings $h_{st} : D_s \longrightarrow D_t$, such as surjectivity or injectivity. Some indeterminate object of experience in D_t may have

19 At this stage the reader should think of syntheses executed by a single consciousness. This will be generalized later to several cognitive agents when we come to consider Kant's notion of 'objective validity'.

20 Of necessity, this discussion is very brief, although it marks the point where the Transcendental Aesthetic comes into contact with the Transcendental Analytic. We will offer a much more detailed explanation elsewhere.

disappeared in D_s: more time for 'running through and holding together' may actually yield fewer features that can be combined.[21] Thus we do not require the mappings to be surjective. It may also happen that an indeterminate object of experience in D_t corresponds to two or more of such objects in D_s, as when a building viewed from afar upon closer inspection turns out to be composed of two spatially separated buildings; thus the mappings need not be injective.

The interaction of the directedness of the index set and the mappings h_{st} is of some interest. If $r \geq s, t$ there are mappings $h_{rs} : D_r \longrightarrow D_s$ and $h_{rt} : D_s \longrightarrow D_t$. Each 'indeterminate object of empirical intuition' in $d \in D_r$ can be seen as a synthesis of such objects $h_{rs}(d) \in D_s$ and $h_{rt}(d) \in D_t$. For example, the 'manifold of a house' (B162) can be viewed as synthesized from a 'manifold of the front' and a 'manifold of the back'. (Note that we do not require that every pair of objects in D_t and D_s is unifiable in this sense.) The operation just described has some of the characteristics of the *synthesis of reproduction in imagination*: the fact that the front of the house can be unified with the back to produce a coherent object presupposes that the front can be reproduced as it is while we are staring at the back. The mappings $h_{rs} : D_r \longrightarrow D_s$ and $h_{rt} : D_s \longrightarrow D_t$ capture the idea that $d \in D_r$ arises from reproductions of $h_{rs}(d)$ and $h_{rt}(d)$ in r.

7.2. Inverse limit of an inverse system of sets. In the quote from A109, repeated below for convenience, Kant talks about the essential connection between the 'necessary unity of self-consciousness' and the transcendental object:

> Appearances are not things in themselves, but themselves only representations, which in turn have their object, which therefore cannot be further intuited by us, and that may therefore be named the non-empirical, i.e., the transcendental object = X. The pure concept of this transcendental object (which in reality throughout all our cognition is always one and the same = X), is what can alone confer upon all our empirical concepts in general relation to an object, i.e., objective validity. This concept cannot contain any determinate intuition at all, and therefore concerns nothing but that unity which must be encountered in a manifold of cognition insofar as it stands in relation to an object. This relation, however, is nothing other than the necessary unity of self-consciousness, thus also of the synthesis of the manifold, through a common function of the mind for combining it in one representation. (A109)

As we remarked above, this connection seems mysterious: 'unity of self-consciousness' and 'transcendental object' seem to belong to completely different conceptual domains. But if one agrees to represent unity of self-consciousness via a directed set of indices, there *is* a connection with the transcendental object, which is given through the following definition and Theorem 7.5 below.

DEFINITION 7.3. *Let* $(T, \{D_s \mid s \in T\}, \mathcal{F})$ *be an inverse system. Let* $\mathcal{D} \subseteq \Pi_{s \in T} D_s$ *the set of all* ξ *such that for* $s \geq t$, $h_{st}(\xi(s)) = \xi(t)$. *Such* ξ *will be called* threads. *Then* $\mathcal{D}$ *is called the* inverse limit *of the given inverse system.*

[21] For an empirical illustration of this phenomenon, one can think of Treisman's 'illusory conjunctions'(Treisman & Gelade, 1980).

We propose to identify the transcendental object of A109 with the inverse limit as just defined. Before we state the properties of inverse limits that we shall use in the sequel, we give an excerpt from a letter from Béatrice Longuenesse (quoted with permission) that sheds considerable light on the notion of transcendental object:

> I would say ”the transcendental object” should be taken as a mass noun referring to whatever it is that our intuition gives us access to without giving us any knowledge of it except as an appearance. The singular term ”the transcendental object” is not of the same nature as the singular term as we would use it to refer to an empirical object, e.g.: ”the dictionary sitting on my desk right now.” [One must not] think of ”the transcendental object” as a count noun, or of transcendental object(s) as falling under categories of unity and plurality just like empirical objects do. The situation in fact is: thinking of our judgments as ranging over empirical objects is also thinking of our judgments as being about *some object =X*, whatever it is, however it is to be sliced up, that our empirical objects represent for us. To say that that object is ”the same through out our cognition” to say that there is one and the same stuff out there that our cognitions try to capture and do capture by way of giving values to the variable ’x’ in our judgments.[22]

This makes clear that the transcendental object must be represented as some thing, otherwise it couldn’t be ‘sliced up’; hence considered as a classical set it must be nonempty. Now, without additional conditions, it is possible that the inverse limit is nonempty, so that in this formal model there wouldn’t be any ‘stuff’ for our cognition to capture. The comparison of the transcendental object with a mass noun suggests however that our formalization of the transcendental object as a set is not completely adequate; it would be more appropriate to choose a whole/part template. Technically, this would require a category theoretic formalization, but this would compound the expository problems. We therefore continue to work in a set theoretic framework, in which we have one transcendental object, the nonempty inverse limit, together with elements of the inverse limit, to be called *transcendental elements*, introducing a term not to be found in *CPR*.

If the transcendental object viewed as inverse limit is nonempty, it provides ‘coherence among our representations’ in the following sense

[22] The continuation of the letter shows that there are constructive aspects to the notion of the transcendental object: “[Thus], it is true that if given the choice, the transcendental object should be thought about on the model of part/whole *rather than* set/elements and in that sense, is analogous to the Kantian *totum continuum* of space. In fact this comparison comes up when Kant discusses the Transcendental Ideal, which is thought of as a whole of reality where whole precedes parts just as in space, as an ‘infinite given magnitude,’ the whole precedes parts. But, it is probably even truer that strictly speaking, the transcendental object—the object = X that ‘corresponds to’ our intuitions—transcends the distinction between the two ways of slicing up reality: whole/part *or* set/elements. For the fact that there is an *alternative* between these two ways of slicing up reality (roughly: whole/part being characteristic of intuition, set/element of understanding) is characteristic of our own, finite understanding, which does not intuit: for which intuition and concepts are two irretrievably separate modes of representation, which complement each other in our cognition.”

LEMMA 7.4. *Let* $(T, \{D_s \mid s \in T\}, \mathcal{F})$ *be an inverse system with nonempty inverse limit* $\mathcal{D}$. *Then for each* $s, t \in T$ *with* $s \geq t$ *there exists a mapping* $\pi_s : \mathcal{D} \longrightarrow D_s$ *satisfying* $h_{st} \circ \pi_s = \pi_t$.

Proof. Given a thread $\xi \in \mathcal{D}$, put $\pi_s(\xi) = \xi_s$. □

That is, there is a single transcendental object (the inverse limit) to which all intuitions (the D_s) relate. At the same time, 'reference to an object represented by the term 'x' in the logical form of judgement does have as one of its components reference to an independent object', in the sense that for some d, assignment of d to x can be thought of as deriving from an assignment to x of an 'independent object' that projects to x.

Although, as we remarked above, the elements of the transcendental object have no Kantian meaning per se, it is useful to think of them in the following manner. Recall that the D_s represent the possible objects of empirical intuition. Intuitively speaking, one can have many different intuitions of the same 'object'; think of all the different possible views of a house. There is no single intuition that contains within itself all the different intuitions one can have of the house; nevertheless these intuitions somehow form a unity. A transcendental element in the sense just defined is not an object of empirical intuition, since it is not an element of a D_s, but in virtue of the previous lemma it binds together objects of empirical intuition into a unity.

Our next concern is the nonemptiness of the transcendental object. In the general case that we study here, in which the mappings h_{st} need not be surjective, it is not guaranteed that the inverse limit is nonempty. If we require that the indexed sets D_s are finite, nonemptiness can be proved, as we shall see in a moment. There is a Kantian justification for this assumption, which has to with the connection between the Categories of Quantity and finite numbers, provided by the so-called Schemata. In a nutshell, the idea is this. The universal judgement 'all A are B' corresponds to the Category of Totality: to make such a judgement one has to be certain that one has considered everything that falls under the extension of A. Kant conceives of this 'considering everything that falls under the extension of A' as a serial process, taking place in time, that is, one must *enumerate* the objects falling under the extension of A. For Kant, the enumeration yields a totality only if the extension enumerated is finite.[23]

The existence proof that we are about to give will make clear why there is an essential connection between the nonemptiness of the transcendental object and the directedness of the index set (the 'unity of self-consciousness').

THEOREM 7.5. *Let* $(T, \{D_s \mid s \in T\}, \mathcal{F})$ *be an inverse system where each* D_s *is finite. Then the inverse limit is nonempty.*

Proof. For each $s \in T$ provide D_s with the discrete topology. Form the Cartesian product $\Pi_{s\in T} D_s$; this product is nonempty by Tychonov's theorem. We have to show that the subset of $\Pi_{s\in T} D_s$ consisting of the threads is nonempty as well.
Define for each $s \in T$,

$$\begin{aligned} A_s &= \{\xi \in \Pi_{s\in T} D_s \mid \forall t \leq s \; h_{st}(\xi(s)) = \xi(t)\} \\ &= \bigcap_{t \leq s} \{\xi \in \Pi_{s\in T} D_s \mid h_{st}(\xi(s)) = \xi(t)\}. \end{aligned}$$

[23] For the full argument, we refer to Longuenesse (1998, chap. 9).

The last representation shows, using the definition of inverse system, that each A_s is a closed nonempty set. We show that the collection of sets A_s has the finite intersection property. Consider A_s and A_t. We have to show that $A_s \cap A_t$ is nonempty. By directedness of T, choose $r \geq s, t$ then $A_r \subseteq A_s \cap A_t$, which shows that $A_s \cap A_t$ is nonempty. Analogously for all finite cardinalities.
Hence by compactness, the collection of threads, $\bigcap_{s \in T} A_s$, is nonempty. □

7.3. Inverse systems of models and the third synthesis. To formalize the 'synthesis of recognition in a concept', we represent concepts as relations on the indexed sets of an inverse system. This additional structure entails additional requirements on the mappings h_{st}.

DEFINITION 7.6. *Let T be a directed set. An inverse system of models indexed by T is a family of first-order models $\{\mathcal{M}_s | s \in T\}$ together with a family of homomorphisms $\mathcal{F} = \{h_{st} | s \geq t, h_{st} : \mathcal{M}_s \longrightarrow \mathcal{M}_t\}$. The mappings in $\mathcal{F}$ must satisfy the coherence requirement that if $s \geq t \geq r$, $h_{tr} \circ h_{st} = h_{sr}$. That $h_{st} : \mathcal{M}_s \longrightarrow \mathcal{M}_t$ is a* homomorphism *means that $\mathcal{M}_s \models R(a_1, a_2, \ldots)$ implies $\mathcal{M}_t \models R(h_{st}(a_1), h_{st}(a_2), \ldots)$.*[24]

Bearing in mind that $s > t$ means that in index s there is more time for synthesis than in index t, the homomorphism condition means that if an object of appearance a in $\mathcal{M}_s$ satisfies the concept C, its image $h_{st}(a)$ in $\mathcal{M}_t$ will also do so; but if $h_{st}(a)$ satisfies C in $\mathcal{M}_t$, the extra time available in s may lead one to reject $C(a)$ in $\mathcal{M}_s$.

For expository reasons we introduced inverse systems of sets before defining inverse systems of models. This does not mean that first the mappings h_{st} are fixed and that concepts have to conform to these via the homomorphism condition; on the contrary, there exists a mutual dependence between concepts and mappings, reflecting Kant's view that concepts play a constitutive role with regard to objects.

7.4. Inverse systems of models and their inverse limits. We are now ready to introduce the main definition, that will ultimately allow us to set up a sound and complete proof system for transcendental logic.

DEFINITION 7.7. *Let $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ be an inverse system. Let D_s be the domain of $\mathcal{M}_s$. Let $\mathcal{D} \subseteq \Pi_{s \in T} D_s$ the set of all ξ such that for $s \geq t$, $h_{st}(\xi(s)) = \xi(t)$. Define a model $\mathcal{M}$ with domain $\mathcal{D}$ by putting $\mathcal{M} \models R(\xi^1, \xi^2, \ldots)$ if for all $s \in T$, $\mathcal{M}_s \models R(\xi^1_s, \xi^2_s, \ldots)$. $\mathcal{M}$ is called the* inverse limit *of the given inverse system.*

THEOREM 7.8. *Let $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ be an inverse system of models where each domain D_s is finite. Then the inverse limit is nonempty.*

LEMMA 7.9. *Let $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ be an inverse system of models with nonempty inverse limit $\mathcal{M}$. Then the projection π_s defined by $\xi \in \mathcal{M} \mapsto \pi_s(\xi) := \xi(s)$ is a homomorphism that in addition satisfies for $s \geq t$: $h_{st}(\pi_s(\xi)) = \pi_t(\xi)$.*

§8. A language for objects: objectively valid judgements. Kant provides several definitions of judgement, ranging from the enumeration of syntactic forms in the Table of Judgements[25] to a definition of what the cognitive role of judgement is in Section 19 of

[24] We do not require surjectivity of h_{st}.

[25] This applies to the headings Quantity, Quality, and Relation; Modality refers rather to ways of using judgements.

CPR—where the latter pays scant attention to the possible syntactic form of judgements. Nevertheless, it is this definition that provides the clearest clue as to Kant's intended semantics for judgements, so we will give the relevant passages from Section 19 and then turn this into a formal definition. This will allow us to relate the cognitive role of judgements to the syntactic forms in the Table of Judgements, and also to provide a formal underpinning to Longuenesse's ideas about the various roles the variable x plays in Kant's representation of categorical judgements (see the beginning of Section 6).

> If, however, I investigate more closely the relation of given cognitions in every judgement [...] then I find that a judgement is nothing other than the way to bring given cognitions to the **objective** unity of apperception. That is the function of the copula **is** in them: to distinguish the objective unity of given representations from the subjective. (B141–2)
> Only in this way does there arise from this relation [between given cognitions] a **judgement**, i.e. a relation that is **objectively valid**, and that is sufficiently distinguished from the relation between these same representations in which there would only be subjective validity, e.g. in accordance with the law of association. In accordance with the latter I could only say "If I carry a body, I feel a pressure of weight," but not "It, the body, is heavy," which would be to say that these two representations are combined in the object [...] (B142)

We thus see again the connection between 'unity of apperception' and 'relation to the object' that we also encountered in A109, quoted in Section 6.4.[26] Following Longuenesse's (1998, p. 111) interpretation, we combine A109 with B141–2, and then (quotation repeated for convenience)

> [W]e cannot but come to the conclusion that reference to an object represented by the term 'x' in the logical form of judgement does have as one of its components reference to an independent object leading us to seek coherence among our representations of it and allowing us to think of the connection of representations as in some sense necessary [...] which is precisely the role Kant assigned to reference to the "transcendental object = x" in the A Deduction.

These considerations suggest that judgements must be interpreted both on objects of experience and on the transcendental object and that there should exist a relation between these interpretations. We illustrate this by means of the particular affirmative judgement 'Some A are B', which Kant glosses as 'To some x to which A belongs, also B belongs'. All we can ever do is find witnesses for x among objects of experience; but we hope that this is sufficient to have a witness for x that refers to an object internally represented as existing outside us. In terms of the formal structures introduced in Section 7 this means: if we have an inverse system $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ whose limit $\mathcal{M}$ exists, and if for all indices s, $\mathcal{M}_s \models \exists x(Ax \wedge Bx)$, then $\mathcal{M} \models \exists x(Ax \wedge Bx)$. More generally,

[26] Section 18 makes the point as well:

> The **transcendental unity** of apperception is that unity through which all of the manifold given in intuition is united in a concept of the object. It is called **objective** on that account [...]. (B139)

DEFINITION 8.1. *We may provisionally define a first-order sentence* φ *to be objectively valid if for any inverse system* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ *with inverse limit* $\mathcal{M}$*, if for all* s*,* $\mathcal{M}_s \models \varphi$ *then* $\mathcal{M} \models \varphi$*.*

So far we have talked about sentences only. If we also want to interpret formulas, we need to say something about assignments; this will give us one way of understanding how a variable x can refer both to objects of experience and to the transcendental object.

DEFINITION 8.2. *An assignment* F *on an inverse system* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ *is a function that assigns to each index* s *and each variable* x *an object in* D_s *(the domain of* $\mathcal{M}_s$*) such that for* $s \geq t$*,* $h_{st}(F(s,x)) = F(t,x)$*. We write* F_s *for the assignment on* $\mathcal{M}_s$ *that results if we hold the argument* s *fixed in* F*.*

DEFINITION 8.3. *Let* F *be an assignment on the inverse system* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$*. We write* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F}) \Vdash \varphi[F]$ *if for all* $s \in T$*,* $\mathcal{M}_s \models \varphi[F_s]$*. We think of* $\Vdash$ *as representing truth on the world of experience* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$*. If there is no danger of confusion, satisfaction on an inverse system will be abbreviated by* $\{\mathcal{M}_s\} \Vdash \varphi[F]$*.*

LEMMA 8.4. *If the inverse limit of* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ *exists,* $F(x)$*, considered as a function of the index* s *is a thread in the inverse limit.*

DEFINITION 8.5. *If* F *is an assignment on the inverse system* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$*, we write* $\widehat{F}$ *for the function that assigns to each variable* x *the thread* $F(x)$*.*

If we now look back at Kant's characterization of transcendental logic as given in Section 4.2, we see that transcendental logic is concerned with 'objective validity', and hence what truth of a judgement in the world of experience implies for truth about 'independent' objects. This can be captured formally in terms of a definition of validity as follows:

DEFINITION 8.6. *Let* Γ *be a set of formulas and* φ *a formula. We write* $\Gamma \Vdash_{ov} \varphi$ *if for any world of experience* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$*, any assignment* F*, and the associated transcendental object* $\mathcal{M}$*,* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F}) \Vdash \Gamma[F]$ *implies* $\mathcal{M} \models \varphi[\widehat{F}]$*. If* $\Gamma \Vdash_{ov} \varphi$*, we say that* Γ *forces* φ *to be true of independent objects.*

Objective validity of a formula φ in the sense of Definition 8.1 now corresponds to the condition $\varphi \Vdash_{ov} \varphi$. In Section 9.3 we shall see that this characterization captures a good part of Kant's intentions. Furthermore, if all formulas in Γ are objectively valid, $\Gamma \models \varphi$ entails $\Gamma \Vdash_{ov} \varphi$. In general the entailment relations $\models$ and $\Vdash_{ov}$ are very different however, and their interplay will shed light on Kant's transcendental logic.

§9. A class of objectively valid judgements: geometric implications. We will prove below that the class of objectively valid formulas coincides with a class of formulas that is of independent interest, the geometric implications.

DEFINITION 9.1. *Let* $\mathcal{L}$ *be a first-order language.*[27] *A formula is* geometric *in* $\mathcal{L}$ *if it is constructed from atomic formulas in* $\mathcal{L}$ *using* $\wedge$*,* $\vee$*,* $\perp$ *and* $\exists$*.*

DEFINITION 9.2. *A formula in* $\mathcal{L}$ *is a* geometric implication *if it is of the form* $\forall \bar{x}(\varphi(\bar{x}) \rightarrow \psi(\bar{x}))$*, where* φ*,* ψ *are positive primitive.*

[27] We assume $\mathcal{L}$ contains =.

We can view a geometric formula $\psi(\bar{x})$ as a special case of a geometric implication by writing it as $\forall \bar{x}(\varphi(\bar{x}) \rightarrow \psi(\bar{x}))$ where φ is a conjunction of identity statements for the variables occurring in $\bar{x}$.

LEMMA 9.3. *A geometric implication is equivalent to a conjunction of formulas* $\forall \bar{x}$ $(\theta(\bar{x}) \rightarrow \psi(\bar{x}))$, *where* θ *is a conjunction of atomic formulas, and* ψ *geometric.*

If we now compare the judgement forms just defined with Kant's Table of Judgements (restricted to Quantity, Quality, and Relation, since Modality is of a different character altogether), we see that the hypothetical and disjunctive judgements are included, as well as the affirmative categorical judgements; no distinction is made between affirmative and infinite judgements (this is because we consider for the moment first-order formulas only; see Footnote 33 for the required extension). Negative particular judgements are excluded, however, and this for good Kantian reasons, as we shall see in Section 9.3. Of course the class of judgements defined in Definition 9.1 goes somewhat beyond the letter of the Table of Judgements because of the $\forall\exists$ combination, although as we have argued, not beyond its spirit: Kant's hypothetical judgement form needs the full complexity of geometric implications, and some of the synthetic a priori judgements have this structure as well.

9.1. Constructive content. From a philosophical point of view it is interesting that the logic of geometric implications[28] is intuitionistic: let $\Gamma \cup \{\varphi\}$ be a finite set of geometric implications such that $\Gamma \vdash \varphi$ (classically) then there exists an intuitionistic proof of φ from Γ.[29] This shows that logical relations between geometric implications are compatible with a constructive understanding of the existential quantifier, as providing a witness given by a construction that takes instantiations of the universal quantifiers as input. This brings us close to the definition of judgement that Kant gives in the *Prolegomena* Section 23:

> Judgements, when considered merely as the condition of the unification of representations in a consciousness, are rules.

There is thus a logical relation between the definitions of judgement as given in *CPR* (B141–2) and as given in the *Prolegomena* Section 23. Coquand has in fact devised a proof system for geometric implications (called 'dynamic proofs') which brings out their rule-like character particularly well.

Geometric implications have strong connections to Euclidean geometry as well. For example, the axiom saying that given a center and a radius, one can construct a circle with that center and radius, can be represented by a geometric implication. Avigad *et al.* (2009) have recently shown in this journal that Euclid's *Elements* can be axiomatized using geometric implications only, by formalizing reasoning with diagrams. That is, proofs via diagrams correspond at the syntactic level to proofs involving geometric implications. Kant of course believed that geometric proofs proceed in intuition via construction, and do not correspond to symbolic proofs. At the same time, Transcendental Deduction B Section 24 introduces the intriguing notion of *synthesis speciosa*, the 'effect of the understanding on sensibility', which ensures the correspondence of the syntheses of the understanding according to the possible logical forms (*synthesis intellectualis*) to the syntheses of our

[28] The name derives from topos theory (see for example, Goldblatt, 2006, sec. 16.4).

[29] See Palmgren (2002) for a proof theoretic argument.

sensibility.[30] Seen in this light, it is surely of interest that the same judgement forms play a role in the domains of the understanding and of sensibility.

9.2. Geometric implications are objectively valid on inverse systems of finite models.

THEOREM 9.4. *Suppose given an inverse system $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$, where T is countable and the domain of each $\mathcal{M}_s$ finite. In this case the inverse limit $\mathcal{M}$ exists (Theorem 7.5). Let $\varphi(\bar{x})$ be a geometric formula and F an assignment such that $\{\mathcal{M}_s\} \Vdash \varphi[F]$. Then $\mathcal{M} \models \varphi[\widehat{F}]$.*

Proof. By induction on the complexity of φ. The statement of the theorem is true by definition if φ is atomic. If the statement holds for $\varphi_1, \ldots, \varphi_n$ it holds as well for $\varphi_1 \wedge \ldots \wedge \varphi_n$.
To proceed, we need a definition and two lemmata.

DEFINITION 9.5. *Let T be the index set of the inverse system. A subset $S \subseteq T$ is* cofinal *if for every $t \in T$ there exists $s \in S$ with $s \geq t$.*

LEMMA 9.6. *If $S \subseteq T$ is cofinal, the inverse limit of $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ is isomorphic to that of $(S, \{\mathcal{M}_s \mid s \in S\}, \mathcal{F})$.*

LEMMA 9.7. *(Lyndon) Let $h : \mathcal{A} \longrightarrow \mathcal{B}$ be a homomorphism from a model $\mathcal{A}$ to a model $\mathcal{B}$. Let $\psi(\bar{x})$ be geometric, and $\bar{a}$ a tuple of elements from $\mathcal{A}$, then $\mathcal{A} \models \psi[\bar{a}]$ entails $\mathcal{B} \models \psi[h(\bar{a})]$.*

Now let $\varphi = \varphi_1 \vee \varphi_2$, where the φ_i are geometric, such that for all s, $\mathcal{M}_s \models \varphi[F_s]$. We show that at least one disjunct φ_1, φ_2 holds on a cofinal set of indices (or rather on the models indexed by these indices). Suppose that for φ_1 there exists $t \in T$ such that for all $s \geq t$, $\mathcal{M}_s \not\models \varphi[F_s]$. It follows that for all $s \geq t$, $\mathcal{M}_s \models \varphi_2[F_s]$. The set $S = \{s \mid s \geq t\}$ is cofinal in T, so the inverse limit determined by T is isomorphic to that determined by S. By the inductive hypothesis, $\mathcal{M} \models \varphi_2[\widehat{F}]$, whence $\mathcal{M} \models \varphi[\widehat{F}]$.

Lastly, suppose $\varphi = \exists x \theta(x)$, such that for all s, $\mathcal{M}_s \models \exists x \theta(x)[F_s]$. Since the index set is countable and directed, we can construct a *linear* cofinal subset of T, as follows. Enumerate T as $t_0, t_1, t_2, \ldots$ and construct a linear cofinal subset S by putting $s_0 := t_0$; s_1 is an index larger than both s_0 and t_1; s_2 is an element larger than both s_1 and t_2, etcetera; put $S : \{s_0, s_1, s_2, \ldots\}$. Since we assumed the indexed models are finite, the collection of domains $\{D_s \mid s \in S\}$ can be given the structure of a finitely branching tree, using the homomorphisms h_{st} to define the branches. Pick $s \in S$, then by hypothesis, for some assignment f on $\mathcal{M}_s$ such that for all variables z different from x, $f(z) = F_s(z)$: $\mathcal{M}_s \models \theta[f]$. If $r \in S$ is such that $r \leq s$, then $h_{sr}(f)$ defines an assignment on $\mathcal{M}_r$ and $\mathcal{M}_r \models \theta[h_{sr}(f)]$. Since θ is geometric, it follows that the objects witnessing the existential quantifier in $\exists x \theta(x)$ determine a subtree of the finitely branching tree $\{D_s \mid s \in S\}$. Because $\mathcal{M}_s \models \exists x \theta(x)[F_s]$ for all s, this subtree is infinite. By König's Lemma it must therefore have an infinite branch. This branch is a thread ξ in the limit of the inverse system $(S, \{\mathcal{M}_s \mid s \in S\}, \mathcal{F})$ which can be extended to a thread in the inverse limit of $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ by cofinality of S. Define an assignment G by $G(s, z) := F(s, z), G(s, x); = \xi_s$. Then $\{\mathcal{M}_s\} \Vdash \theta[G]$, whence by the inductive hypothesis $\mathcal{M} \models \theta[\widehat{G}]$, and hence $\mathcal{M} \models \exists x \theta[\widehat{F}]$. □

[30] Longuenesse (1998, chap. 8) provides an elaborate discussion of this topic.

COROLLARY 9.8. *Suppose given an inverse system* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$*, where* T *is countable and the domain of each* $\mathcal{M}_s$ *finite. Let* F *be an assignment on the inverse system. Let* $\forall \bar{x}(\varphi(\bar{x}) \rightarrow \psi(\bar{x}))$ *(where* φ, ψ *are geometric) be a geometric implication such that* $\{\mathcal{M}_s\} \Vdash \forall \bar{x}(\varphi(\bar{x}) \rightarrow \psi(\bar{x}))[F]$*, then* $\mathcal{M} \models \forall \bar{x}(\varphi(\bar{x}) \rightarrow \psi(\bar{x}))[\widehat{F}]$.

Proof. To show that $\forall \bar{x}(\varphi(\bar{x}) \rightarrow \psi(\bar{x}))$ holds on the inverse limit $\mathcal{M}$ under $\widehat{F}$, assume that $\mathcal{M} \models \varphi[G]$, where G is an assignment that differs at most on $\bar{x}$ from F, so that for all $s \in T$, $\mathcal{M}_s \models \varphi[G_s]$ by Lemma 9.7. The hypothesis of the theorem gives that for all $s \in T$, $\mathcal{M}_s \models \psi[G_s]$, whence $\mathcal{M} \models \psi[\widehat{G}]$ by Theorem 9.4. □

9.3. A closer look at objective validity. With this theorem and especially its proof at our disposal, we can give a more detailed justification of the definition of objective validity. Consider a sentence $\exists x A(x)$. If this sentence is true on some $\mathcal{M}_s$, that is, if the existential quantifier is witnessed by an object of experience in $\mathcal{M}_s$, then $\exists x A(x)$ need not yet be objectively valid, that is, witnessed by an object represented as being independent of our cognition; for the relevant object of experience in $\mathcal{M}_s$ may disappear at stage $t > s$. Moreover, even if $\exists x A(x)$ is witnessed in several $\mathcal{M}_s$, and even though we intend $\exists x A(x)$ to be true of the same object in these $\mathcal{M}_s$, there is no guarantee this will be so. But if $\exists x A(x)$ is witnessed by an object of experience in all $\mathcal{M}_s$, one would like this to be so because an object independent of our cognition unifies these witnesses and provides them with objective reality; and the above proof shows that there is a transcendental element (a thread in the inverse limit) which does this.[31] But whether witnesses among objects of experience can be so related to the transcendental object depends on the type of judgement of which they are witnesses. For consider now $\exists x \neg A(x)$. It is easy to construct (see Figure 1) an inverse system whose limit contains a single thread satisfying A, and where $\exists x \neg A(x)$ is true at each indexed model, but where the witnesses for $\exists x \neg A(x)$ are totally unrelated.

In this context it is interesting what Kant has to say about the role of negation:

> To be sure, logically, one can express negatively any proposition that one wants, but in regard to the content of our cognition in general, that is, whether it is expanded or limited by a judgement, negative judgements have the special job solely of preventing error (A709/B737).

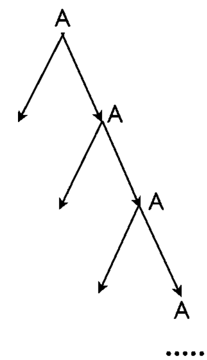
This seems to suggest that negative judgements need not have objective validity.[32] In this context, Wolff (1995, pp. 290–291) quotes C.S. Peirce, who wrote that in traditional logic, existential import of a judgement depends on its *Quality* not *Quantity*: negative judgements do not come with existential import. Peirce writes:

> It is probable that Kant also understood the affirmative proposition to assert the existence of its subject, while the negative did not do so; so that 'Some phoenixes do not rise from their ashes' would be true, and 'All phoenixes do rise from their ashes' would be false.

[31] As a consequence, if a geometric implication $\forall x(\varphi(x) \rightarrow \psi(x))$ has existential import on all models $\mathcal{M}_s$, that is, if in each $\mathcal{M}_s$ there is an object of experience satisfying $\varphi(x)$, then $\forall x(\varphi(x) \rightarrow \psi(x))$ has existential import on $\mathcal{M}$, and we can think of these objects of experience as unified by the transcendental object.

[32] Note though that it is a consequence of the formal definition of objective validity that negative universal judgements are objectively valid.

Fig. 1



The same interpretation is given in Thompson (1953), where it is also argued that the infinite judgement corresponding to the negative particular judgement—'Some A is non-B'—does carry existential import.[33]

9.3.1. Objective validity and truth It is important to distinguish objective validity from truth. Objective validity of a judgement does *not* mean that the judgement is true of transcendental objects, that is, of objects represented as existing independently of us. Rather, it is a conditional notion: if the judgements holds on the world of experience (represented by an inverse system of models), then it holds for transcendental objects as well. Of course we can never be fully certain that the given judgement holds of the world of experience, so truth on transcendental objects cannot be established definitively, and functions rather as a regulative norm governing the activity of judgement. But if a judgement is objectively valid, we know at least that it is not pointless to verify it on the world of experience, since in doing so we may 'relate our representations to an object' (Longuenesse , 1998, p. 82).

9.3.2. Truly objective validity Despite the relation established in A109 between 'unity of self-consciousness' and 'objective validity', Kant does not intend that objective validity relates to each subject taken individually only:

> Persuasion is a mere semblance, since the ground of the judgement, which lies solely in the subject, is held to be objective. Hence such a judgement also has only private validity, and this taking something to be true cannot be communicated. Truth, however, rests upon agreement with the object, with regard to which, consequently, the judgements of every understanding must agree [...] The touchstone of whether taking something to be true is conviction or mere persuasion is therefore, externally, the possibility of communicating it and finding it to be valid for the reason of every human being to take it to be true; for in that case there

[33] The infinite judgement 'some A are non-B' can be formalized using the infinite disjunction $\bigvee$ as

$$\bigvee_{C \cap B = \emptyset} (\exists x (A(x) \wedge B(x)).$$

interestingly, this is still a geometric formula, and Theorem 9.4 can be extended to show that this infinitary formula is objectively valid.

> is at least a presumption that the ground of the agreement of all judgements, regardless of the difference among the subjects, rests on the common ground, namely the object, with which they therefore all agree and through which the truth of the judgement is proved. (A820–1/B848–9)

This distinction, between 'private validity' and 'validity for every human being' plays a prominent (and controversial) role in Section 18–20 of the *Prolegomena*; see Longuenesse (1998, chap. 7) for discussion. Here we bring up this topic only to note that it calls for an extension of the interpretation of the index set T, which now contains the mental states in which syntheses are executed of all cognitive agents of interest, together with the requirement that the the mental states of each agent are cofinal in the set of mental states of every other agent (cf. Kant's remark on the 'possibility of communicating').

§10. The transcendental object revisited: generalized inverse limits. So far we have worked with a concrete representation of the transcendental object as an inverse limit of a countable inverse system of finite models. However, if we look carefully at the justification for this identification, we see that what is really important are the properties of the maps that exist between inverse limit and indexed models, and between the indexed models themselves, that is, the commutativity of the various diagrams of maps, since these express the unifying function of the transcendental object. Also, the proof that geometric formulas are objectively valid hinges on the compactness properties that hold in this concrete setting. In order to get rid of this special feature we define a model-theoretic notion of inverse limit which has preservation of geometric formulas built in.

DEFINITION 10.1. *Let an inverse system* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ *be given. The structure* $\mathcal{M}$ *is a* generalized inverse limit *of* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ *if it satisfies the properties*

(1) *there exists an injective homomorphism from* $\mathcal{M}$ *into the inverse limit of* $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ *(hence the inverse limit is nonempty).*
(2) *for each* s *there must exist a homomorphism* $\pi_s : \mathcal{M} \longrightarrow \mathcal{M}_s$ *satisfying the coherence condition: for* $b \in \mathcal{M}$, $s \geq t$: $h_{st}(\pi_s(b)) = \pi_t(b)$
(3) *for geometric formulas* φ*, and parameters* $\overline{a}$ *from* $\mathcal{M}$*, if for all* $s \in T$, $\mathcal{M}_s \models \varphi[\pi_s(\overline{a})]$*, then* $\mathcal{M} \models \varphi[\overline{a}]$
(4) $\mathcal{M}$ *is maximal in the following sense: any model* $\mathcal{N}$ *that satisfies the previous conditions (with projections* $\pi_n^{'}$*) and whose cardinality is less than that of* $\mathcal{M}$ *can be embedded into* $\mathcal{M}$ *via a homomorphism* g *satisfying* $\pi_n \circ g = \pi_n^{'}$*).*

A few comments on this definition will be helpful. The elements of $\mathcal{M}$ can be thought of as threads, via the mapping $a \in \mathcal{M} \mapsto \{\pi_s(a) \mid s \in T\}$, where the r.h.s. is a thread because of coherence Condition (2). This map will in fact yield the injective homomorphism required by (1); this follows from (3) applied to the geometric formula $x = y$. One way to think of Condition (3) is to assume that the language of $\mathcal{M}$ is expanded with a unique individual constant for each object of $\mathcal{M}$; we denote the resulting expansion of of $\mathcal{M}$ by $(\mathcal{M}, \overrightarrow{a})$. Let c be a name for the object a, then since $\mathcal{M} \models x = c[a]$ and $x = c$ is positive primitive, $\mathcal{M}_s \models x = c[\pi_s(a)]$. Thus the constants of the language of $(\mathcal{M}, \overrightarrow{a})$ have a unique interpretation in $\mathcal{M}_s$. In this setting Condition (3) can be formulated as

> for positive primitive sentences φ in the language of $(\mathcal{M}, \overrightarrow{a})$, if for all $s \in T$, $\mathcal{M}_s \models \varphi$, then $\mathcal{M} \models \varphi$.

LEMMA 10.2. *Let $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ be an inverse system with generalized inverse limit $\mathcal{M}$, and let φ be a geometric implication (in the language of $(\mathcal{M}, \overrightarrow{d})$) true on all $\mathcal{M}_s$. Then φ is true on $(\mathcal{M}, \overrightarrow{d})$).*

Proof. Immediate from Definition 10.1. □

We take a generalized inverse limit $\mathcal{M}$ to be a representation of the transcendental object associated to the given inverse system. While the mapping posited in (1) need not be surjective, in a sense the threads in the inverse limit are indistinguishable from the elements of $\mathcal{M}$: if a thread ξ in the inverse limit witnesses a geometric formula $\varphi(x)$, there will also be an element from $\mathcal{M}$ that does so, by (3).[34]

10.1. Objective validity redefined. Since we will now work with generalized inverse limits, the definition of then entailment relation $\Vdash_{ov}$ needs to be adapted.

DEFINITION 10.3. *Let $(T, \{\mathcal{M}_s \mid s \in T\}, \mathcal{F})$ be an inverse system with generalized inverse limit $\mathcal{M}$. Then $\Gamma(\overline{x}) \ \Vdash_{ov} \varphi(\overline{x})$ means that for all $\overline{a} \in \mathcal{M}$, all $s \in T$, $\mathcal{M}_s \models \Gamma[\overline{\pi_s(a)}]$ implies $\mathcal{M} \models \varphi[\overline{a}]$. We call $\varphi(\overline{x})$* objectively valid *if $\varphi(\overline{x}) \Vdash_{ov} \varphi(\overline{x})$.*

We see that the variables on the l.h.s. of $\Vdash_{ov}$ receive their interpretation from the transcendental elements that are assigned to these variables on the r.h.s.; this seems to be as close as one can get to Longuenesse's view, cited on page 21, on the meaning of the variable x in Kant's logic lectures and in *CPR*.

§11. A characterization of objectively valid formulas.

THEOREM 11.1. *Let φ be objectively valid. Then there are finitely many geometric implications $\gamma_1, \ldots, \gamma_n$ such that*

(1) $\varphi \models \gamma_i$*, for* $1 \leq i \leq n$
(2) $\gamma_1 \wedge \ldots \wedge \gamma_n \models \varphi$

Proof. Define $\Sigma = \{\gamma \mid \gamma \text{ is a geometric implication and } \varphi \models \gamma\}$. By the compactness theorem it suffices to prove that $\Sigma \models \varphi$, for this will give us the finitely many geometric implications $\gamma_1, \ldots, \gamma_n$.[35] The idea of the proof is that we pick a model $\mathcal{M}$ of Σ and show that it can be written as generalized inverse limit of a countable inverse system $\{\mathcal{M}_n \mid n \in \mathbb{N}\}$ such that for all n, $\mathcal{M}_n \models \varphi$. Since φ is assumed to be objectively valid, it follows that $\mathcal{M} \models \varphi$.

---

[34] Note that Condition (4) implies that each set of ξ of cardinality less than $|\mathcal{M}|$can be mapped homomorphically to $\mathcal{M}$; but there doesn't have to be an isomorphism.

[35] At this point there is a curious analogy with one of Kant's 'formal criteria of truth', discussed in Section 4.1:

> *If all the consequences of a cognition are true, then the cognition is true too.* For if there were something false in the cognition, then there would have to be a false consequence too.
> From the consequence, then, we may infer to a ground, but without being able to determine this ground. Only from the complex of all consequences can one infer *to a determinate ground*, infer that it is the true ground.

Σ stands for the 'complex of all consequences' of φ that are possible grounds; we try to prove that $\Sigma \models \varphi$, that is, 'if all the consequences of a cognition are true, then the cognition is true too', and in so doing infer the 'determinate ground', the $\gamma_1, \ldots, \gamma_n$.

This proof sketch can be turned into a full proof most easily if we work with saturated models. We shall set aside worries about the additional hypotheses needed to guarantee existence of saturated models, and in particular will not present the proof in terms of special models, so as not to obscure the main ideas. (For definitions of 'saturated' and 'special' see Chang & Keisler , 1990, chap. 5) or Hodges 1993, chap. 10).)

An essential ingredient in the proof is the following reformulation of Lemma 5.2.9 in Chang & Keisler (1990):

THEOREM 11.2. *Let $\mathcal{B}$ be a saturated model of cardinal κ, and let $\mathcal{A}$ be a model of cardinality $\leq \kappa$. Suppose every geometric sentence true on $\mathcal{A}$ is true on $\mathcal{B}$. Let $(\mathcal{A}, \overrightarrow{d})$ the expansion of $\mathcal{A}$ with a 1–1 enumeration of its objects. Then there exists a homomorphism from $\mathcal{A}$ to $\mathcal{B}$ with respect to geometric formulas in the language of the expansion $(\mathcal{A}, \overrightarrow{d})$.*

Choose a saturated model $\mathcal{M}$ of cardinality κ with $\mathcal{M} \models \Sigma$; we have to show $\mathcal{M}$ can be written as the generalized inverse limit of an inverse system $\{\mathcal{M}_n \mid n \in \mathbb{N}\}$ such that φ is true on each $\mathcal{M}_n$. We proceed by induction, and suppose $\mathcal{M}_n$ and a projection $\pi_n : \mathcal{M} \longrightarrow \mathcal{M}_n$ have been constructed. We assume $\mathcal{M}_n$ is saturated of cardinality $\lambda > \kappa$. We construct $\mathcal{M}_{n+1}$ as a saturated model of a certain set of sentences Ξ_{n+1} that we now proceed to specify.

(1) $\varphi \in \Xi_{n+1}$

(2) There must be a homomorphism $h_{n+1} : \mathcal{M}_{n+1} \longrightarrow \mathcal{M}_n$. By Theorem 11.2 this means that every geometric formula true on $\mathcal{M}_{n+1}$ is true on $\mathcal{M}_n$. Since we do not have $\mathcal{M}_{n+1}$ yet, this translates into: every geometric sentence false on $\mathcal{M}_n$ must be false on $\mathcal{M}_{n+1}$. Define $\Delta_{n+1} = \{\neg\psi \mid \psi \text{geometric}, \mathcal{M}_n \not\models \psi\}$ and put Δ_{n+1} in Ξ_{n+1}.

(3) There must exist a homomorphism $\pi_{n+1} : \mathcal{M} \longrightarrow \mathcal{M}_{n+1}$. Again by Theorem 11.2 this means that every geometric sentence true on $\mathcal{M}$ is true on $\mathcal{M}_{n+1}$. Define $\Gamma_{n+1} = \{\tau \mid \tau \text{ geometric}, \mathcal{M} \models \tau\}$ and put Γ_{n+1} in Ξ_{n+1}.

(4) To satisfy Condition (3) of Definition 10.1, consider the expansion $(\mathcal{M}, \overrightarrow{d})$. Every geometric sentence false on $(\mathcal{M}, \overrightarrow{d})$ must be false on some $\mathcal{M}_n$. We therefore need an enumeration $\theta_1, \theta_2, \ldots$ of geometric sentences θ_i and define $\Theta_{n+1} = \{\neg\theta_{n+1}(\overline{c}) \mid (\mathcal{M}, \overrightarrow{d}) \not\models \theta_{n+1}(\overline{c})\}$, and put Θ_{n+1} in Ξ_{n+1}.

(5) This concludes the specification of Ξ_{n+1}.

LEMMA 11.3. *Ξ_{n+1} is consistent.*

Proof. Suppose Ξ_{n+1} is inconsistent, then there finite sets $\neg\psi_1, \ldots, \neg\psi_k \in \Delta_{n+1}, \tau_1, \ldots, \tau_l \in \Gamma_{n+1}$, and $\theta_{n+1}(\overline{c}_1), \ldots, \theta_{n+1}(\overline{c}_k) \in \Theta_{n+1}$ such that

$$\varphi \models \neg(\neg\psi_1 \wedge \ldots \wedge \neg\psi_k \wedge \tau_1 \wedge \ldots \wedge \tau_l \wedge \neg\theta_{n+1}(\overline{c}_1) \wedge \ldots \wedge \theta_{n+1}(\overline{c}_k)),$$

which is equivalent to

$$\varphi \models (\tau_1 \wedge \ldots \wedge \tau_l \rightarrow \psi_1 \vee \ldots \vee \psi_k \vee \theta_{n+1}(\overline{c}_1) \vee \ldots \vee \theta_{n+1}(\overline{c}_k)).$$

Since the constants in the occurrences of θ_{n+1} do not occur in φ, we have

$$\varphi \models \forall\overline{x}_1 \ldots \overline{x}_k(\tau_1 \wedge \ldots \wedge \tau_l \rightarrow \psi_1 \vee \ldots \vee \psi_k \vee \theta_{n+1}(\overline{x}_1) \vee \ldots \vee \theta_{n+1}(\overline{x}_k)).$$

The formula on the r.h.s. of $\models$ is a geometric implication, hence in Σ. Since $\mathcal{M} \models \Sigma$ and all τ are true on $\mathcal{M}$, we have

$$\mathcal{M} \models \forall\overline{x}_1 \ldots \overline{x}_k(\psi_1 \vee \ldots \vee \psi_k \vee \theta_{n+1}(\overline{x}_1) \vee \ldots \vee \theta_{n+1}(\overline{x}_k)).$$

Now by the induction hypothesis a projection $\mathcal{M} \longrightarrow \mathcal{M}_n$ has been constructed so that the ψ_i are false on $\mathcal{M}$, whence we get

$$\mathcal{M} \models \forall \overline{x}_1 \ldots \overline{x}_k(\theta_{n+1}(\overline{x}_1) \vee \ldots \vee \theta_{n+1}(\overline{x}_k)),$$

which is a contradiction. □

Due to the contribution of Θ_{n+1}, the set Ξ_{n+1} can have cardinality κ. This was the reason we assumed $\mathcal{M}_n$ to be saturated of cardinality $\lambda > \kappa$. We now choose $\mathcal{M}_{n+1}$ as a saturated model of Ξ_{n+1} of cardinality λ, then all required homomorphisms exist.
We now have to verify that $\mathcal{M}$ satisfies the conditions of Definition 10.1 for being a generalized inverse limit. The essential observation here is that for a constant c in the language of the expansion $(\mathcal{M}, \overrightarrow{a})$, $(\mathcal{M}, \overrightarrow{a}) \models x = c[a]$ implies $(\mathcal{M}_{n+1}, \overrightarrow{\pi_{n+1}(a)}) \models x = c[\pi_{n+1}(a)]$, which in turn implies $(\mathcal{M}_n, \overrightarrow{h_{n+1}(\pi_{n+1}(a))}) \models x = c[h_{n+1}(\pi_{n+1}(a))]$. But we also have directly $(\mathcal{M}_n, \overrightarrow{\pi_n(a)}) \models x = c[\pi_n(a))]$, whence $h_{n+1}(\pi_{n+1}(a)) = \pi_n(a)$. This gives Conditions (1) and (2); and (3) holds by construction. To prove (4), observe that the diagram of homomorphisms is such as to force that every geometric sentence true on $\mathcal{N}$ is true on $\mathcal{M}$.[36] The commutativity of the diagram $\pi_n \circ g = pi_n^{'}$ is established as before. Now suppose there are $b, b' \in \mathcal{N}$ such that $g(b) = g(b')$. Then for all n, $\mathcal{M}_n \models x = y[\pi_n(g(b)), \pi_n(g(b'))]$, whence $\mathcal{M}_n \models x = y[\pi_n^{'}(b), \pi_n^{'}(b')]$. Since $x = y$ is geometric, this means that $b = b'$. □

Both for the purpose of the proof theory given below and for good Kantian reasons, we need a slight generalization of the preceding theorem:

THEOREM 11.4. *If* Γ *is a set of sentences and* φ *a sentence such that* $\Gamma \Vdash_{ov} \varphi$, *then there are finitely many geometric implications* $\gamma_1, \ldots, \gamma_n$ *such that* $\Gamma \models \gamma_1 \wedge \ldots \wedge \gamma_n$ *and* $\gamma_1 \wedge \ldots \wedge \gamma_n \models \varphi$.

Proof. Analogous to the proof of Theorem 11.1, the only change is that Σ is defined as $\{\gamma \mid \gamma$ is geometric implication $\forall\exists, \Gamma \models \gamma\}$ and we again show that $\Sigma \models \varphi$. □

The sentences in Γ specify a theory that every indexed model in the inverse system under consideration must satisfy. This is related to Kant's remarks in A573–4/B601–2, which point to the fact that concepts are not independent, but carry analytic structure:

> [W]e nevertheless find that this idea [of the sum total of all possible predicates in general], as an original concept, excludes a multiplicity of of predicates, which, as derived through others, are already given, or cannot coexist with one another [...]

Theorem 11.4 then says that the only consequences of this analytic structure that have objective validity are of the geometric implication form. Longuenesse (1998) discusses this point on p. 87 and again on p. 107, both times with reference to Section 36 of the *Jäsche Logik*. The point is that Kant also introduces the variable x in the representation of *analytic* judgements:

> An example of an *analytic* proposition is, To everything x to which the concept of body $(a + b)$ belongs, belongs also *extension* (b). An example

[36] This is the only place where saturation of $\mathcal{M}$ is used.

> of a *synthetic* proposition is, To everything x to which the concept of body belongs, belongs also *attraction* (c).

Longuenesse (1998, p. 87) comments

> Kant makes the presence of the x to which the two concepts are attributed explicit *for analytic as well as for synthetic judgements*. This is because in both cases concepts have meaning only if they are "universal or reflected representations" of singular objects [...] For all judgements, even when they are analytic, what ultimately makes the combination of concepts possible is relation to an "x of judgement".

Since this 'x of judgement' is the same for both analytic and synthetic judgements, the same notion of objective validity must apply.

§12. Transcendental logic as a 'logic of truth'. Theorem 11.4 leads us back to the original purpose of 'the logic of truth':

> The part of transcendental logic that expounds the elements of the pure cognition of the understanding and the principles without which no object can be thought at all, is the transcendental analytic, and at the same time a logic of truth. For no cognition can contradict it without at the same time losing all content, i.e. all relation to any object, hence all truth. (A62–3/B87)

In this section we present a sound and complete proof system for the 'logic of truth', which more or less does what Kant wants it to do. We will see that the main step in the completeness proof is Theorem 11.4.

We claim here that the 'logic of truth' is determined by the notion of validity $\Vdash_{ov}$.[37] The logical form underlying $\Vdash_{ov}$ involves the Kantian notion of object as formalized through inverse systems and their (generalized) inverse limits. In terms of this notion of validity, that φ is objectively valid can be stated as '$\varphi \Vdash_{ov} \varphi$', or if we restrict ourselves to inverse systems that satisfy a given theory Γ, as '$\Gamma, \varphi \Vdash_{ov} \varphi$'. That a 'cognition' φ contradicts the 'logic of truth' can then be formalized as either $\varphi \nVdash_{ov} \varphi$ or $\Gamma, \varphi \nVdash_{ov} \varphi$. In the first case this means concretely that there is an inverse system $\{\mathcal{M}_s\}$ with generalized inverse limit $\mathcal{M}$ such that all $s \in T$, $\mathcal{M}_s \models \varphi$ but $\mathcal{M} \not\models \varphi$. We would like this to be the case if and only if the sequent $\varphi \Rightarrow \varphi$ is not derivable in a certain sequent calculus.

12.1. A sequent calculus for the 'logic of truth'. Let $\Rightarrow_{ov}$ be the syntactic counterpart of $\Vdash_{ov}$, to be used in the geometric implication of the sequent calculus. The peculiarity of this calculus is that, whereas usually $\varphi \Rightarrow \varphi$ (or more generally $\Gamma, \varphi \Rightarrow \varphi$) is taken as an axiom(-schema), here it is one purpose of the calculus *to determine for which φ, the sequent $\varphi \Rightarrow_{ov} \varphi$ is derivable*. One special case of the schema can be taken as axiomatic, since for atomic formulas A we do have $A \Rightarrow_{ov} A$. In the usual systems one would derive $\varphi \Rightarrow_{ov} \varphi$ (for arbitrary φ) from this by induction, but with respect to $\Rightarrow_{ov}$ the rules should be such that the induction is blocked beyond the geometric implication stage. We achieve this introducing an auxiliary sequent arrow $\Rightarrow_{pov}$ which has the same semantic interpretation

[37] In the sense of Definition 10.3.

as $\Rightarrow_{ov}$, but for which the rules are such that $\varphi \Rightarrow_{pov} \theta$ is derivable only if φ, θ are geometric.

The proposed proof system thus has three sequent arrows: $\Rightarrow$, whose interpretation is as in classical logic, and both $\Rightarrow_{ov}$ and $\Rightarrow_{pov}$ whose interpretation is given by $\Vdash_{ov}$. The purpose of the calculus is to derive sequents of the form $\Gamma \Rightarrow_{ov} \varphi$; $\Rightarrow$ and $\Rightarrow_{pov}$ are auxiliary notions. We will have the following types of rules: rules for classical logic, formulated using $\Rightarrow$, rules for the objectively valid formulas involving only $\Rightarrow_{pov}$, rules for the objectively valid formulas holding for both $\Rightarrow_{pov}$ and $\Rightarrow_{ov}$ (to avoid duplication we will write these rules with an arrow $\Rightarrow_{(p)ov}$ which means 'either $\Rightarrow_{pov}$ or $\Rightarrow_{ov}$'), and rules that mix the various arrows. The system will be called $\mathcal{LT}$ in a reference to Kant's term 'logic of truth', and comprises the following rules:

(1) The standard sequent rules for $\Rightarrow$.

(2) Rules valid for $\Rightarrow_{pov}$ only

$$A \Rightarrow_{pov} A \quad \text{Axiom } (A \text{ an atomic formula})$$

$$\frac{\varphi \Rightarrow_{pov} \varphi \qquad \psi \Rightarrow_{pov} \psi \qquad \varphi \Rightarrow_{pov} \theta \qquad \psi \Rightarrow_{pov} \theta}{\varphi \vee \psi \Rightarrow_{pov} \theta} \text{ Left } \vee$$

$$\frac{\varphi \Rightarrow_{pov} \theta}{\varphi \Rightarrow_{pov} \exists x \theta} \text{ Right } \exists$$

$$\frac{\varphi \Rightarrow_{pov} \theta}{\exists x \varphi \Rightarrow_{pov} \theta} \text{ Left } \exists \ (x \text{ not free in } \theta)$$

(3) Rules valid for both $\Rightarrow_{pov}$ and $\Rightarrow_{ov}$

$$\frac{\varphi \Rightarrow_{(p)ov} \theta}{\varphi \wedge \psi \Rightarrow_{(p)ov} \theta} \text{ Left } \wedge$$

$$\frac{\varphi \Rightarrow_{(p)ov} \theta \qquad \varphi \Rightarrow_{(p)ov} \tau}{\varphi \Rightarrow_{(p)ov} \theta \wedge \tau} \text{ Right } \wedge$$

$$\frac{\varphi \Rightarrow_{(p)ov} \theta}{\varphi \Rightarrow_{(p)ov} \theta \vee \tau} \text{ Right } \vee$$

(4) Rules valid for $\Rightarrow_{ov}$

$$\frac{\varphi \Rightarrow_{ov} \theta}{\forall x \varphi \Rightarrow_{ov} \theta} \text{ Left } \forall$$

$$\frac{\varphi \Rightarrow_{ov} \theta}{\varphi \Rightarrow_{pov} \forall x \theta} \text{ Right } \forall \ (x \text{ not free in } \varphi)$$

(5) Rules involving $\Rightarrow_{pov}$, $\Rightarrow_{ov}$ and/or $\Rightarrow$

$$\frac{\varphi \Rightarrow_{pov} \varphi}{\varphi \Rightarrow_{ov} \varphi}$$

$$\frac{\varphi \Rightarrow_{pov} \varphi \qquad \Gamma \Rightarrow_{ov} \varphi \qquad \Gamma, \psi \Rightarrow_{ov} \theta}{\Gamma, \varphi \to \psi \Rightarrow_{ov} \theta} \text{ Left } \to$$

$$\frac{\varphi \Rightarrow_{pov} \varphi \qquad \Gamma, \varphi \Rightarrow_{ov} \psi}{\Gamma \Rightarrow_{ov} \varphi \rightarrow \psi}\ \text{Right} \rightarrow$$

$$\frac{\Gamma \Rightarrow \varphi \qquad \varphi \Rightarrow_{ov} \psi \qquad \psi \Rightarrow \theta}{\Gamma \Rightarrow_{ov} \theta}\ ov\text{–Cut}$$

The *ov*-Cut rule plays a special role in $\mathcal{LT}$: one could say it connects 'general logic' (represented by the classical sequent arrow $\Rightarrow$) with transcendental logic proper (represented by $\Rightarrow_{ov}$). It says that cut-formulas must be of a specific form (geometric implications). It cannot be eliminated, because it expresses that inferences from inverse system to inverse limit have to proceed via objectively valid formulas, and cannot proceed more directly.
Our next aim is to prove soundness and completeness of $\mathcal{LT}$ with respect to $\Vdash_{ov}$.

LEMMA 12.1. $\dfrac{\varphi_1 \Rightarrow_{ov} \varphi_1 \qquad \dots \qquad \varphi_n \Rightarrow_{ov} \varphi_n}{\varphi_1 \wedge \dots \wedge \varphi_n \Rightarrow_{ov} \varphi_1 \wedge \dots \wedge \varphi_n}$

Proof. By repeated application of the left and right $\wedge$ rules. □

LEMMA 12.2. $\dfrac{\varphi \Rightarrow_{pov} \varphi \qquad \psi \Rightarrow_{ov} \psi}{\varphi \rightarrow \psi \Rightarrow_{ov} \varphi \rightarrow \psi}$

Proof.

$$\dfrac{\dfrac{\varphi \Rightarrow_{pov} \varphi \qquad \psi \Rightarrow_{ov} \psi}{\varphi \rightarrow \psi, \varphi \Rightarrow_{ov} \psi}\ \text{Left} \rightarrow \qquad \varphi \Rightarrow_{pov} \varphi}{\varphi \rightarrow \psi \Rightarrow_{ov} \varphi \rightarrow \psi}\ \text{Right} \rightarrow$$

As a consequence we have

LEMMA 12.3. $\dfrac{\varphi \Rightarrow_{pov} \varphi \qquad \psi \Rightarrow_{ov} \psi}{\forall x(\varphi \rightarrow \psi) \Rightarrow_{ov} \forall x(\varphi \rightarrow \psi)}$

LEMMA 12.4. *$\mathcal{LT}$ is sound for the intended interpretation: $\Gamma \Rightarrow_{ov} \psi$ implies $\Gamma \Vdash_{ov} \psi$.*

THEOREM 12.5. *$\varphi \Rightarrow_{ov} \varphi$ is derivable iff φ is provably equivalent (in classical logic) to a conjunction of geometric implications.*

Proof. For the right to left direction, it follows from Lemmas 12.3 and 12.1 that for every conjunction ψ of geometric implications, $\psi \Rightarrow_{ov} \psi$ is derivable. The *ov*-Cut rule extends this to the case of φ provably equivalent to such ψ. For let $\gamma_1, \dots, \gamma_n$ be such that for, $\varphi \models \gamma_1 \wedge \dots \wedge \gamma_n$, and $\gamma_1 \wedge \dots \wedge \gamma_n \models \varphi$, then *ov*-Cut applied to the cut formula $\gamma_1 \wedge \dots \wedge \gamma_n$ yields $\varphi \Rightarrow_{ov} \varphi$.[38] The left to right follows from Theorem 11.1 by soundness (Lemma 12.4). □

THEOREM 12.6. *$\mathcal{LT}$ is complete: $\Gamma \Vdash_{ov} \psi$ implies $\Gamma \Rightarrow_{ov} \psi$.*

Proof. If $\Gamma \Vdash_{ov} \psi$, then by Theorem 11.4 there is a conjunction $\gamma_1 \wedge \dots \wedge \gamma_n$ of geometric implications γ_i such that $\Gamma \models \gamma_1 \wedge \dots \wedge \gamma_n$ and $\gamma_1 \wedge \dots \wedge \gamma_n \models \psi$. By completeness for classical first-order logic it follows that $\Gamma \Rightarrow \gamma_1 \wedge \dots \wedge \gamma_n$ and $\gamma_1 \wedge \dots \wedge \gamma_n \Rightarrow \psi$ are derivable. By Theorem 12.5 $\gamma_1 \wedge \dots \wedge \gamma_n \Rightarrow_{ov} \gamma_1 \wedge \dots \wedge \gamma_n$ is derivable. The proof is completed by an application of *ov*-Cut:

[38] This explains the peculiar 'symmetric' form of the cut rule.

$$\frac{\Gamma \Rightarrow \gamma_1 \wedge \ldots \wedge \gamma_n \qquad \gamma_1 \wedge \ldots \wedge \gamma_n \Rightarrow_{ov} \gamma_1 \wedge \ldots \wedge \gamma_n \qquad \gamma_1 \wedge \ldots \wedge \gamma_n \Rightarrow \psi}{\Gamma \Rightarrow_{ov} \psi}$$

□

It follows that $\varphi \Rightarrow_{ov} \varphi$ is not derivable in $\mathcal{LT}$ ('contradicts the logic of truth') if and only if there is an inverse system $\{\mathcal{M}_s\}$ with generalized inverse limit $\mathcal{M}$ such that all $s \in T$, $\mathcal{M}_s \models \varphi$ but $\mathcal{M} \not\models \varphi$ ('loses any relation to the object', in the sense that what a sentence says about objects of experience need not be related to what it says about the transcendental object).

Our claim is not that the completeness theorem shows that Kant's transcendental logic has now been formalized in its full extent. For example, we did not at all consider his treatment of causality. If we were to do so, the structure of the inverse systems would have to be considerably enriched: the indexed models would have to contain events, and would have to support predicates describing instantaneous and continuous causation,[39] perhaps governed by axioms such as the event calculus (van Lambalgen & Hamm , 2004). Our main purpose here has been to introduce a formal semantics for judgements based on Kant's notion of object, and in doing so to establish a framework in which further specifications of transcendental logic can be situated.

§13. Back to the Table of Judgements. With these formal results established, we may now consider their implications for Kant's Table of Judgements and its role in the argumentative structure of *CPR*. It was Kant's intention to enumerate primitive logical forms which each express a particular function of cognition as it attempts to construct objects out of sensory manifolds. Strawson (1966) claimed this was a misconceived enterprise, since (a) there are no primitive logical forms due to interdefinability (e.g., the hypothetical judgement is definable using negation and the disjunctive judgement), and (b) based on an (arbitrary) choice of primitives, indefinitely many logical forms can be derived, and one would not want to say that each of these corresponds to a particular function of cognition.

The answer to this is that Kant selected 'primitive' logical forms of judgement already with a particular transcendental purpose in mind, a purpose that renders the classical semantics underlying Strawson's objection inapplicable. We have seen on p. 258 that according to Longuenesse (1998, p. 78) Kant selected forms of judgement for inclusion in the Table, not because these were sanctioned by traditional logic, but because they play a role in "bring[ing] given cognitions to the objective unity of apperception," that is, to relate our representations to objects.[40] Thus, there is first a (largely suppressed) general argument selecting particular judgement forms, followed by detailed arguments (mostly in the Transcendental Deduction of the B edition) aligning particular logical forms with particular cognitive functions directed toward the constitution of objects.

The results of this paper pertain to the first step in the argumentation. We have seen that in B141–2, Kant provides a *functional* characterisation of judgement in terms of objective validity:

> If, however, I investigate more closely the relation of given cognitions in every judgement [...] then I find that a judgement is nothing other than

[39] Watkins (2004) argues that it is especially continuous causation as it occurs in Newtonian mechanics through the concept of force, that led to Kant's particular theory of causality.

[40] There is one awkward aspect of this interpretation though, as we have seen negative particular judgements are not objectively valid.

> the way to bring given cognitions to the **objective** unity of apperception. That is the function of the copula **is** in them: to distinguish the objective unity of given representations from the subjective. Only in this way does there arise from this relation [between given cognitions] a **judgement**, i.e. a relation that is **objectively valid** [...] (B141–2).

Theorem 11.1 can be viewed as the derivation of a structural characterization—deriving specific syntactic forms—from Kant's functional characterisation of judgement.

This characterization is very closely related to the Table of Judgements. On the one hand, it is easily seen that affirmative categorical judgements are included in our class, as are disjunctive judgements.[41] The extended discussion of hypotheticals in Section 4.1 shows that their intrinsic logical complexity can be as high as geometric implication, hence they fit in our class. On the other hand, our class does not include negative particular judgements, but then these types of judgements were probably not intended to be candidates for objective validity, given Kant's views on negation. Furthermore, our characterization limits the complexity of objectively valid judgements to geometric implication, that is, the logical complexity of hypotheticals. Thus we have in a sense proved the completeness of the Table of Judgements.

This is perhaps the place to address what may be a lingering worry in the reader's mind. Friedman (1992) has extensively argued that it is precisely Kant's logic's *lack* of ability to express the $\forall\exists$ quantifier combination that drove his philosophy of mathematics. For example, because his logic does not allow him to express that a certain domain is infinite (this needs the $\forall\exists$ quantifier combination), Kant was forced to adopt a 'constructive' approach to mathematics that acknowledges potential infinity only. By contrast, we have identified the geometric implication judgement forms as those judgements that should be of interest to Kant, since only they can claim objective validity. Shouldn't we say with Friedman that although Kant should have considered geometric implication judgement forms, he was in effect barred from doing so because the inferences contained in his general logic are too weak? The answer must be the following. It is clear that Kant did not conceive of 'all' in the universal categorical judgement and 'some' in the particular categorical judgements as universal and existential quantifiers that can be freely combined among themselves and with other logical operators, for example, to generate a $\forall\exists$ judgement. It is also clear that Kants general logic is not complete for geometric logic. This however is not an argument against the claim that Kant's judgement forms are geometric. Even the fact that Kant takes categorical judgements in subject–predicate form is not a counterargument, since predicates may well contain existential quantifiers (as Kant's examples of synthetic a priori principles show), and once this is allowed, geometric implications look like bona fide subject–predicate judgements. Lastly, Kant has ways to deal with the $\forall\exists$

[41] With the proviso that one may question whether the meaning given to the disjunctive judgement here exhausts Kant's meaning: his talk of parts and wholes, A73–4/B99, and more elaborately in the *Jäsche Logik*, Section 27, Section 28) suggests his disjunction has some substructural aspects. For example, the usual introduction rule for $\vee$ fails for his intended meaning. The simultaneous presence of the disjuncts can perhaps be captured by the *par* connective of linear logic (see the rules for 'par logic' in Hyland & de Paiva, 1993, p. 282), or the structural connective "," in Boricic' (1985) multiple-conclusion natural deduction system . Note that in these cases the elimination rule for disjunction is not the one familiar from classical (or intuitionistic) logic, but the disjunctive syllogism, as in Kant. We conjecture that the connection between disjunctive judgements and the Category of Community that Kant discusses in the Third Analogy can be made much more transparent by giving disjunction one of the substructural readings.

quantifier combination in geometric implications; as the argument of the Second Analogy of Experience amply shows, he interprets the quantifier combination as a rule. Thus for Kant, the inferences necessary for dealing with geometric implications are to be found in transcendental logic. His way of proceeding here is consistent, because geometric logic is intrinsically intuitionistic, so that we may think of the $\forall\exists$ quantifier combination as given by a rule.

§14. Acknowledgements. We thank Béatrice Longuenesse for enlightening correspondence, Pablo Cubiles Kovacs for discussions about the model-theoretic part of the paper, and the anonymous referee for pointing out flaws in the exposition. The first author is supported by the NWO project 360-20-150: 'The origins of truth'.

BIBLIOGRAPHY

Avigad, J., Dean, E., & Mumma, J. (2009). A formal system for euclid's Elements. *Review of Symbolic Logic*, **2**(4), 700–768.

Bartlett, F. C. (1968). *Thinking: An Experimental and Social Study*. London: Allen and Unwin.

Boricic, B. R. (1985). On sequence-conclusion natural deduction systems. *Journal of Philosophical Logic*, **14**, 359–377.

Bruner, J. S. (1973). *Beyond the Information Given*. New York, NY: W. W. Norton.

Chang, C. C., & Keisler, H. J. (1990). *Model Theory*, *Studies in Logic and the Foundations of Mathematics*, Vol. 73 (third edition). Amsterdam, The Netherlands: North-Holland, First edition: 1973, second edition: 1977.

Coquand, T. A completeness proof for geometric logic. Technical report, Computer Science and Engineering Department, University of Gothenburg. Available from: http://www.cse.chalmers.se/ coquand/formal.html. Retrieved September 29, 2010.

Friedman, M. (1992). *Kant and the Exact Sciences*. Cambridge, MA: Harvard University Press.

Goldblatt, R. (2006). *Topoi. The Categorial Analysis of Logic*. Mineola, NY: Dover.

Hodges, W. (1993). *Model Theory*. Cambridge, MA: Cambridge University Press.

Hyland, M., & de Paiva, V. (1993). Full intuitionistic linear logic (extended abstract). *Annals of Pure and Applied Logic*, **64**, 273–291. doi:10.1016/0168-0072(93)90146-5.

Kant, I. (1992). *Lectures on Logic; Translated from the German by J. Michael Young*. The Cambridge edition of the works of Immanuel Kant. Cambridge, MA: Cambridge University Press.

Kant, I. (1998). *Critique of Pure Reason; Translated from the German by Paul Guyer and Allen W. Wood*. (The Cambridge edition of the works of Immanuel Kant). Cambridge, MA: Cambridge University Press.

Kant, I. (2002). *Theoretical philosophy after 1781; Edited by Henry Allison and Peter Heath*. The Cambridge edition of the works of Immanuel Kant. Cambridge, MA: Cambridge University Press.

Kitcher, P. W. (1990). *Kant's Transcendental Psychology*. New York: Oxford University Press.

Longuenesse, B.(1998). *Kant and the Capacity to Judge.* Princeton, NJ: Princeton University Press.

MacFarlane, J. (2000). What does it mean to say that logic is formal? PhD Thesis, University of Pittsburgh.

Palmgren, E. (2002). An intuitionistic axiomatisation of real closed fields. *Mathematical Logic Quarterly*, **48**(2), 297–299.

Posy, C. J. (2003). Between Leibniz and Mill: Kant's logic and the rhetoric of psychologism. In Jacquette, D., editor. *Philosophy, Psychology and Psychologism*, Dordrecht: Kluwer, pp. 51–79.

Reich, K. (1932). *Die Vollstaendigkeit der kantischen Urteilstafel*. Berlin: Schoetz. Translated as *The completeness of Kant's Table of Judgements* transl. by Kneller, J., and Losonsky, M. (1992). Stanford University Press.

Rosenkoetter, T. (2009). Truth criteria and the very project of a transcendental logic. *Archiv fuer Geschichte der Philosophie*, **61**(2), 193–236.

Steinhorn, C. I. (1985). Borel structures and measure and category logics. In Barwise, J., and Feferman, S., editors. *Model-theoretic Logics*, chapter 16. New York, NY: Springer-Verlag, pp. 579–596.

Strawson, P. F. (1966). *The Bounds of Sense: An Essay on Kant's "Critique of Pure Reason"*. London: Methuen.

Stuhlmann-Laeisz, R. (1976). *Kants Logik*. Berlin: De Gruyter.

Thompson, M. (1953). On Aristotle's square of opposition. *The Philosophical Review*, **62**(2), 251–265.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Clarendon Press*, **12**, 97–136.

van Lambalgen, M., & Hamm, F. (2004). *The Proper Treatment of Events*. Oxford: Blackwell.

Watkins, E. (2004). *Kant and the Metaphysics of Causality*. Cambridge: Cambridge University Press.

Wolff, M. (1995). *Die Vollstaendigkeit der kantischen Urteilstafel*. Frankfurt am Main: Vittorio Klostermann.

INSTITUTE FOR LOGIC,
LANGUAGE AND COMPUTATION/DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF AMSTERDAM
OUDE TURFMARKT 141, 1012GC
AMSTERDAM
E-mail: m.vanlambalgen@uva.nl