Construction and Schemata in Mathematics

Klaus Frovin Jørgensen
Section for Philosophy and Science Studies
Roskilde University

Recently I read Patrick Suppes claiming that "the most important open scandal in philosophy is the problem of free will" [17, p. 205]. I very much agree with Suppes that the problem of the free will is a major puzzle, which we should try to get a better understanding of by examining the deeper issues connected with the free will. This essay, however, does not treat the problem of the free will. It concerns the problems of the ontology and epistemology of mathematics. In genereal, the problems of the philosophy of mathematics are just as old and—if it makes sense to talk about solvability of such problems—perhaps just as unsolved as the problem of the free will.

Mathematics is a very important ingredient of knowledge. In its most simple form mathematics plays a necessary role in our understanding of the surrounding world and is necessary for solving simple problems of ordinary life. At the other end of the simplicity-scale we find the mathematics as used in science. Also here mathematics has a necessary role in our descriptions of nature and the way in which we are involved with it and each other. It truly amazes me that there seems to be very little consensus with respect to the ontology and epistemology of mathematics.

There can be no doubt that the natural numbers exit; I mean, we use them all the time. The question is only: In which sense? This question I have treated this question elsewhere [13], thus I will here be more interested in the existence of mathematical objects whose nature seems more abstract.

Platonism is an important part of the philosophy of mathematics. The upshot of Platonism is its simplicity with respect to existence and truth: Mathematical objects are existing non-temporally and non-spatially in a universe which we can access only by rational thinking. They exist eternally and independently of human beings and are perfectly real. Mathematical statements are therefore objectively true or false and their truthvalue is also independent from humans.[1] Within Platonism one introduces in this way a dualism[2] and from which it follows that all kinds of constructivism are incompatible with Platonism. As Plato himself expresses it:

> [...] no one who has even a slight acquaintance with geometry will

[1] Kurt Gödel, for instance, expressed views in this directions [8].

[2] It is well-known what the general problem of dualisms is: How are the dual worlds connected?

> deny that the nature of this science is in flat contradiction with the absurd language used by mathematicians, for want of better terms. They constantly talk of "operations" like "squaring", "applying", "adding", and so on, as if the object were to *do* something, whereas the true purpose of the whole subject is knowledge – knowledge, moreover, of what eternally exists, not of anything that comes to be this or that at some time and ceases to be. (Republic 527a)

I find this element of Platonism rather problematic and in the following I would like to outline a philosophy of mathematics which on the one hand takes the language used by mathematicians as serious as possible: When mathematicians talk, for instance, about constructing the real numbers what can they possibly mean? On the other hand I have no intensions towards 'taking the fists from the boxer'. On the contrary, I think to much energy has been put into producing normative philosophy of mathematics in discussing what is 'good' or what is 'bad' mathematics.

I want to do a certain kind of naturalized philosophy of mathematics. We see mathematics is being done in a great many aspects of the human life. I would like to know *what* is going on—the interest is not in the direction of what should be going on.

1 Equivalence classes and quotient structures

Mathematicians often talk about constructions in connection with quotient structures.

Let X be some set and suppose $\sim$ is an equivalence relation on X.[3] Given any element x of X we now construct another object $[x]$ which is the collection of all elements which are equivalent to x, i.e.,

$$[x] = \{y \in X \mid x \sim y\}.$$

$[x]$ is called the equivalence class generated by x under $\sim$ and the question now is, in which sense the equivalence class is *constructed* and in which sense it *exists*. Given such an X and $\sim$ we quite often collect all the equivalence classes and form a new set $X/{\sim}$ consisting of all the equivaence classes, in other words:

$$a \in X/{\sim} \quad \text{if and only if,} \quad \begin{array}{l} a \text{ is an equivalence class generated} \\ \text{by some element of } X \text{ under } \sim. \end{array}$$

By extensionality of sets we can lift the equivalence relation $\sim$ on X to an identity relation on $X/{\sim}$.

$$[x] = [y] \text{ if and only if } x \sim y.$$

For at very simple example let X be the set of all cars. If $x \sim y$ obtains just in case x has the same color as does y, then $\sim$ is an equivalence relation.[4] We

[3] That $\sim$ is an equivalence relation on X means that tree conditions obtain for $\sim$ with respect to elements of X. Firstly, any element x is related to itself, i.e., $x \sim x$; secondly, $x \sim y$ implies $y \sim x$; and thirdly, $x \sim y$ and $y \sim z$ implies $x \sim z$.

[4] Leaving the problem of Sorites aside for a moment.

can then form $X/\sim$ which consists of all car-colors, so to speak. Philosophers would tend to think of $X/\sim$ as a collection of concepts, whereas mathematicians would think of it as a collection of new objects.

We can also construct $\mathbb{Q}$ as the collection of equivalence classes of ordered pairs of whole numbers (a, b), with $b \neq 0$. The equivalence relation is defined by

$$(a, b) \sim (c, d) \text{ if and only if } ad = bc.$$

The equivalence class $[(a, b)]$ is then understood as the rational number a/b.

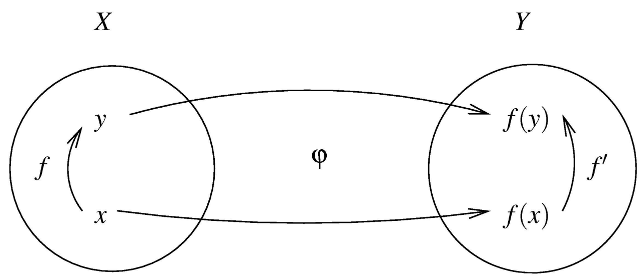
More generally, we can construct equivalence classes in the following way. Let f be a function form X to Y. Then f defines an equivalence relation on X by

$$x \sim y \text{ if and only if } f(x) = f(y).$$

Here $[x]$ is the set of all elements y, which are mapped to $f(x)$. The equivalence relation is the kernel of f:

$$\ker f = \{(x, y) \in X \times X \mid f(x) = f(y)\}.$$

Now suppose that we are given even more. Let us look at the case where we are given a *homomorphism* $\varphi : X \to Y$ where X and Y are structures, e.g., groups or rings, of the same type.[5] Now, a homomorphism $\varphi : X \to Y$ is a certain mapping which preservers the structure of X to Y, in the sense that within Y there is a sub-structure which in a certain sense imitates X. If for instance there is a relation R belonging to X then $(x, y) \in R$ implies that in Y there is an R' such that $(\varphi(x), \varphi(y)) \in R'$. Also, mappings are preserved in the sense that if a function f in X maps x to y, then there is a function f' in Y such that $f'(\varphi(x)) = \varphi(y)$. In the remaining part of this essay I tacitly assume that if φ is a homomorphism from one structure to another, then the two structures are of the same type.



An injective homomorphism φ is called a monomorphism and a surjective homomorphism is called an epimorphism. If φ is bijective it is an isomorphism.

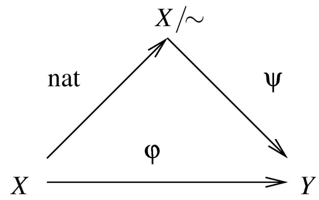
[5]That X and Y are structures of the same type means that we can pair relations from X with relations from Y such that the arity of the two relations in each pair is the same. Likewise for functions. Constants are understood as 0-ary functions.

Moreover, suppose we have a structure X, then a *congruence relation* $\sim$ on X is an equivlence relation on X which respects the relations and functions of X. That means, for instance, if $(x, y) \in R$ and $x \sim x'$ and $y \sim y'$ then $(x', y') \in R$. Thus, given a structure X and a congruence relation we can naturally construct the *quotient structure* $X/\sim$ by letting the domain of $X/\sim$ be the collection of all equivalence classes and if R is a relation on X we define $R_{/\sim}$ by

$$([x], [y]) \in R_{/\sim} \text{ if and only if } (x, y) \in R.$$

Given a congruence relation $\sim$ on a set X we call the epimorphism nat from X to $X/\sim$ defined $x \mapsto [x]$ the *natural epimorphism.*

Theorem. Suppose $\varphi : X \to Y$ is a homomorphism from a structure X to a structure Y. Then there is a monomorphism $\psi : X/\sim \to Y$ such that $\varphi = \psi \circ \text{nat}$. This simply means that the following diagram commutes:



Let me give a very simple example. Let $\mathbb{Z}$ be the structure of whole numbers with the ordinary operations for plus and times, i.e., $\mathbb{Z} = (Z, +, \cdot)$, where Z is the set of whole numbers. Moreover,

$$\mathbb{Z}_2 = (Z_2, \oplus, \odot)$$

with $Z_2 = \{0, 1\}$ and $\oplus$ and $\odot$ defined as:

$\oplus$	0	1
0	0	1
1	1	0

$\odot$	0	1
0	0	0
1	0	1

Now define $\varphi : Z \to Z_2$ in the following way:

$$\varphi(n) = \begin{cases} 0, & \text{if } n \text{ is even,} \\ 1, & \text{otherwise.} \end{cases} \tag{1}$$

Now it is easily seen that $\varphi(n + m) = \varphi(n) \oplus \varphi(m)$ and $\varphi(n \cdot m) = \varphi(n) \odot \varphi(m)$. Therefore, φ respects the functions and φ is an epimorphism. The equivalence relation $\sim_\varphi$ defined by φ gives rise to two equivalence classes [0] and [1] containing the even whole numbers and the uneven whole numbers, resp. And in fact, as the monomorphism from $\mathbb{Z}/\sim_\varphi$ to $\mathbb{Z}_2$, given by the theorem, is surjective the two simple structures $\mathbb{Z}/\sim_\varphi$ and $\mathbb{Z}_2$ are isomophic.

We have now seen different ways to construct new structures, namely quotient structures. There are many other ways of constructing new structures such as product structures and ultra-product structures. I will not, however, analyze those more advanced methods for constructing new objects here. Now I think it is important to think philosophically about what is going on when we form equivalence classes and quotient structures.

2 Kant on the construction of objects

I propose that generalizing Immanuel Kant's notion of schemata is a very fruitful approach to the epistemology of mathematical objects and concepts. Now, it is very well-known that Kant had a rather antiquated view on mathematics and the so-called a priori. He claims to prove that the Euclidean space is the one and only structure that the empirical space can have and he also claims that finger counting validates the laws of arithmetic and he talks about drawing figures in thought or on paper and proving theorems on the basis of these figures. In the "Transcendental Aesthetics" in the first *Critique* he says that space is infinite whereas in the "Dialectics" he says it is neither finite nor infinite. I must be aknowledged that it is from time to time difficult to make sense out of his, sometimes, rather confusing writings. Nevertheless, I claim that his theory of knowledge can be used for something reasonable. I very much agree with Clark Glymore:

> Kant's three works on epistemology are filled with arguments. Unfortunately, the major arguments appear either invalid or too obscure to assess with confidence [...] Unlike Descartes and Hume, Kant's importance does not rest with his arguments themselves. Instead, the value of Kant's work for the theory of knowledge rests in the *kind* of argument he thought he could give for his theory, and in the general picture of knowledge presented in that theory. [7]

In particular I will here examine Kant's theory of schemata.

In general schemata are a mediating "third thing" (A138/B177) which makes it possible that objects can be subsumed under concepts. The latter is somehing purely intellectual and therefore non-sensible, the former on the other hand is something sensible. There must therefore be something which mediates between the senses and the intellectual. In fact, any concept is useless without a schema. As there are basically three different types of concepts—empirical, pure sensible, transcendental—there are also three different types of schemata:

1. Empirical schemata,
2. Geometrical schemata,
3. Transcendental schemata.

By way of example, let us take a look at Kant's own example of empirical schemata. An empirical concept to Kant is a concept which we have derived from experience. Take the concept of dog, for instance. Typically we would tend to think of a dog as an animal which has four feet (and is inclined towards barking). But we are not able to *refer* to anything, unless the concept is used *together* with its schema. When this is the case, then:

> The concept of a dog [together with its schema] signifies a rule in accordance with which my imagination can specify the shape of a four-footed animal in general, without being restricted to any single particular shape that experience offers me of any possible image that I can exhibit *in concreto*. (A141/B180)

The schema of dog is a rule-governed procedure by which the imagination can produce a *paradigmatic mental* image, which functions as a prototype or a representative example giving me a *figurative representation* of a typical dog.[6] By using this I can think of a dog (as such) even though no dogs are present, and when there is a dog present I can subsume this animal under my concept dog, since I find the characteristics to be present in the dog, because of essential similarities between the dog and my mental image of a dog in general. I can compare dogs, and I can count collection of dogs.

On the face of it, the empirical schemata may all seem unproblematic and straight-forward. But there is more to it. The problem is that empirical concepts are dynamical, since "**empirical** concepts cannot be defined at all but only explicated", where "**to define** properly means just to exhibit originally [*ursprünglich*] the exhaustive concept of a thing with its boundaries" (A727/B755). Therefore, to give a definition of the concept "dog", would be to give a priori (*originally*) the necessary and sufficient characteristics for objects to be subsumed under dog. But empirical concepts are derived from experience and thus different persons may understand different things under the concept.

In general, one of the the essential properties of a schema is that it enables *universal* reasoning. From the concept and the construction of a paradigmatic we should be able to reason due to the schema generally about the concept. This is problematic in the case of empirical concepts as these—according to Kant—are not definable. Be that as it may, in the case of mathematical concepts we are en a better position, as these indeed are definable.

There is no doubt that Kant has a special interest in geometry. And in connection with the theory of schemata this is particularly clear. I think there are three reasons for this.

The first reason: The challenge from mathematics. Kant understands mathematics and mathematical knowledge as based on constructions taking place in time. Therefore mathematics is synthetic. Moreover, the propositions of mathematics are necessary, therefore a priori. Simultaneously, Kant aims at giving a general account of human knowledge and here the sciences play a central role.

[6]Thus Kant's theory of schemata are a clear predecessor of a contemporary theory of prototypes found in cognitive psychology as given by Rosch.

So Kant's theory intends at being a framework for understanding the human activity of science. Mathematics is a special kind of science: The objects are in a certain sense non-empirical; nevertheless mathematical knowledge is synthetic. Thus, Kant wants to develop a theory in which we are 'guaranteed' that there really *are* meaningful mathematical objects, fulfilling the criteria concerning constructibility and necessity. The theory of schemata is a basic pillar in explaining this.

The second reason: The possibility of abstract concepts. Kant's more general problem is to explore the possibility of subsuming objects under concepts. If Kant can explain how geometrical concepts, such as triangle, 'function' then, perhaps, he can lift this to a *general* theory of the relation between concepts and objects? In fact, I think this is Kant's strategy

The third reason: The source of schematism. As a Kantian unification of the two foregoing I find it most likely that Kant got the idea about schematism from his studies of Euclid. Lisa Shabel sharpens this claims by saying that Euclidean reasoning through diagrams "provides an interpretive model for the function of a transcendental schema" ([16], 109).

Here we are especially interested in the schematism of mathematical concepts, thus we should take a look at Kant's theory of geometrical schemata. This has, as point of departure, Euclid's geometry.

3 Geometrical schematism

Euclid's *Elements* comprises 13 books with content ranging from basic plane geometry, over arithmetic and incommensurables to solid geometry. A list of definitions, postulates and common notions open Book I. Whereas the postulates and common notions remain the same throughout the 13 books, each of the subsequent books add new definitions to this list. The first book concerns basic geometry and Kant often mentions or refers to the first book in the *Critique*, either the propositions or the postulates[7] or paraphrasings hereof.[8]

Kant was more than inspired by Euclid. He wants to give an epistemological foundation of the *constructive* reasoning style that he sees in Euclid. Let us therefore look at one example from Eulid's first book [5]:

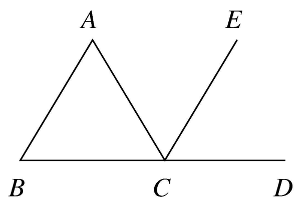
PROPOSITION 32.

> *In any triangle, if one of the sides be produced, the exterior angle is equal to the two interior and opposite angles, and the three interior angles of the triangle are equal to two right angles.*
>
> Let ABC be a triangle, and let one side of it BC be produced to D;
>
> I say that the exterior angle ACD is equal to the two interior and opposite angles CAB, ABC, and the three interior angles of the triangle ABC, BCA, CAB are equal to two right angles.

[7]But note, in case of the postulates he only mentions the first three(!)

[8]See (A24/B39; B154; A163/B204; A234/B287; A239/B299; A261/B317; A300/B356; A511/B539; A716/B744).

For let CE be drawn through the point C parallel to the straight line AB. [I. 31]



Then, since AB is parallel to CE, and AC has fallen upon them, the alternate angles BAC, ACE are equal to one another. [I. 29]

Again, since AB is parallel to CE, and the straight line BD has fallen upon them, the exterior angle ECD is equal to the interior and opposite angle ABC. But the angle ACE was also proved equal to the angle BAC;

therefore the whole angle angle ACD is equal to the two interior and opposite angles BAC, ABC.

Let the angle ACB be added to each;

therefore the angles ACD, ACB are equal to the three angles ABC, BCA, CAB.

But the angles ACD, ACB are equal to two right angles; [I. 13]

therefore the angles ABC, BCA, CAB are also equal to two right angles.

Therefore etc.

The use of diagrams is central in Euclid's reasoning. Here this is already seen in the very formulation of the proposition. The angles are classified as either interior or exterior. These terms cannot be understood unless one uses the diagram—"exterior" and "interior" are only defined implicitly through diagrams. Now, the diagram of I.32 depicts what we are given: a triangle ABC. But more is shown in the diagram. Out of the given triangle we construct line BD (by postulate 2). The extremity of any line is a point, thus D is constructed and exists. Furthermore (due to proposition 31) a line CE can be drawn parallel to AB.[9]

The diagram used in the proof of proposition 32 shows properties about line CE which can *only* be inferred by use of the diagram: From the text we do not know the *direction* in which the construction of CE goes. But from the diagram we see that it is drawn upwards and thus splits ACD in two angles ACE and

[9]Proposition 31 is, in turn, proved by appealing to postulate 1 and 2 and propositions 23 and 27, and the construction taking place in this proof is also visualized by a diagram (as all the proofs in *Elements* generally are).

ECD. Euclid's fifth common notion states that "[t]he whole is greater than the part"—but it is through the diagram we learn what the whole is (ACD) and what the parts are (ACE and ECD).

Now that we have determined relations between the lines and we know what is interior and exterior we see (due to proposition 29) that BAC and ACE are equal (in size). In fact this is the only non-obvious part of the proof.[10] The rest of the proof of proposition 32 is more or less trivial.

This type of reasoning is very paradigmatic to Kant and therefore he wants to give this account because he takes it to be the most important mathematical discipline.[11]

3.1 Constructibility and schemata

It is "schemata that ground our pure sensible concepts" (A140/B180). This grounding is deeply connected with the so-called "**construction** of concepts". What it means to "**construct** a concept" is expressed by Kant on pages A713–4/B741–2. Here we learn that to construct a concept means to construct in accordance with rules a non-empirical intuition, which should represent the concept universally. The production of the intuition—which is Kant's epistemological pendant to Euclidean diagrams—can be done either purely by the imagination or it can be a figure drawn on paper. In the latter case the empirical intuition functions as a symbol which refers by analogy to the pure intuition. The *universality* of the image arises when the particular image is taken *together* with the procedure generating the image. Such a procedure is deeply connected with the Euclidean style of reasoning:

> Give a philosopher the concept of a triangle, and let him try to find out in his way how the sum of its angles might be related to a right angle. He has nothing but the concept of a figure enclosed by three straight lines, and in it the concept of equally many angles. Now he may reflect on this concept as long as he wants, yet he will never produce anything new. He can analyze and make distinct the concept of a straight line, or of an angle, or of the number three, but he will not come upon any other properties that do not already lie in these concepts. But now let the geometer take up this question. He begins at once to construct a triangle. Since he knows that two right angles together are exactly equal to all of the adjacent angles that can be drawn at one point on a straight line, he extends one side of his triangle, and obtains two adjacent angles that together are equal to two right ones. Now he divides the external one of these angles by drawing a line parallel to the opposite side of the triangle, and sees that here there arises an external adjacent angle which is equal to an

[10] Proposition 29 is proved—through a diagram—by appealing essentially to postulate 5 and proposition 15 (which relies proposition 13 and postulate 4).

[11] Moreover, the Euclidean model of construction provides a general model for Kant's notion of *constructibility*.

> internal one, etc. In such a way, through a chain of inferences that is always guided by intuition, he arrives at a fully illuminating and a the same time general solution of the question. (A716–7/B744–5)

Geometrical knowledge evolves in precisely the manner that we have seen the production of mathematical knowledge as given in the case of proposition 32. And here Kant in fact explicitly refers to the proof that proposition. The geometer *constructs* a triangle. But he already knows something in advance, namely (a generalization of) proposition 13. *Therefore* he constructs—in accordance with postulate 2—an extension of one of the lines. Then he divides (a construction validated by proposition 31) the external angle, and *so on*. But it "is always guided by intuition" by the construction of an image (purely mental or empirical with symbolic reference to the pure image) in accordance with some basic postulates, definitions and common notions; after wards discoveries are realized through the constructed image.[12] Then a 'general solution' can be found. I will treat the problem concerning universality in section 3.2. Here I concentrate on what precisely the schemata are.

It is a central thesis of that any meaningful concept has a schema. For instance, "[t]he schema of the triangle [...] signifies a rule of the synthesis of the imagination with regard to pure shapes in space" (A141/B180). So the natural question is: What does this rule consist of? The schema for the concept "triangle" must be a rule by which we can construct any (image of a) triangle. Kant writes: "Thus we think of a triangle as an object by being conscious of the composition of three straight lines in accordance with a rule according to which such an intuition can always be exhibited" (A105). And later in the *Critique*;

> "three lines, two of which taken together are greater than the third, a triangle can be drawn," then I have here the mere function of the productive imagination, which draws the lines greater or smaller, thus allowing them to abut at any arbitrary angle. (A164/B205)

These two quotations show that the "function of the productive imagination" which Kant is referring to is the function defined in the proof of Euclid's proposition 22:

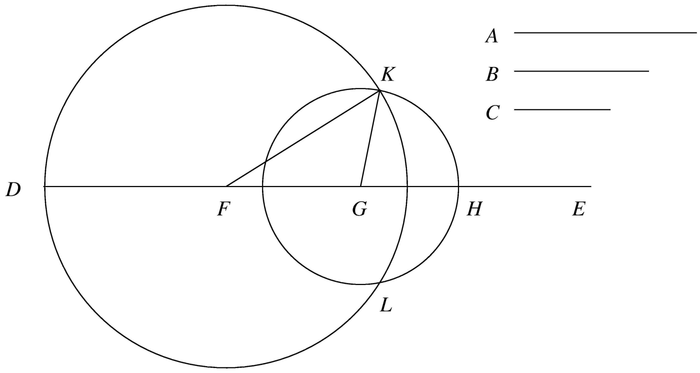
PROPOSITION 22.

> *Out of three straight lines, which are equal to three given straight lines, to construct a triangle: thus it is necessary that two of the straight lines taken together in any manner should be greater than the remaining one.* [I.20]

The proof is of course given by *constructing* a triangle out of the three given lines A, B and C, which fulfill the requirement that any two are greater than the

[12]Thus my interpretation differs from Friedman's [6, p. 90], where it is rejected that intuition "enable us to "read off" the side-sum property [in the case of proposition I.32]". Of course the 'reading of'-procedure is not just a simple visual inspection, as Kant also rejects in the case of Euclid's propostion 5 as referred to on page Bxi-xii.

remaining one. The requirement was proved in proposition 20 to be a property of the collection of the three sides of any triangle. On a line DE constructed to be long enough (postulate 2) we construct DF to be equal (in length) to A; we construct FG to be equal (in length) to B and construct GH equal (in length) to C. This can be done by the procedure given in the proof of proposition 3. Now describe two circles with center F and diameter FD and center G and diameter GH, respectively (postulate 3). The circles intersect (as seen from the diagram) and thus they meet in K. Connect K with F and G (postulate 1) and the triangle is constructed.



Now, the schema for the concept "triangle" is, according to Kant, a rule-governed operation which "draws the lines greater or smaller, thus allowing them to abut at any arbitrary angle" (A164/B205) (postulate 2). Out of these three lines a triangle is constructed.[13] The rule-governed operation producing all triangles must, however, also conform to the insight expressed in the *assumption* given in the formulation of proposition 22: That any two of the lines are greater than the third.

Let us understand the concept "triangle" as similar to a definition. In modern terms this will sound like the following:

x is a triangle, iff, x is a polygon with three vertices
and three sides which are straight lines.

Now, the concept is like a (Peircean) type, whereas the individually constructed images are the tokens falling under the type. The schema corresponding to the concept is, following this line of thought, a *complete* method which exhaust the relation between a type and the tokens falling under that type. By the schema we can construct all triangles, but we can also decide whether a given figure is

[13]This construction is with respect to postulates ultimately founded on the basic postulates. Thus the schema and the meaning of a concept as understood by Hilbert (c.f. the classic discussion between Hilbert and Frege on how concepts are defined in (formal) mathematical theories are strikingly close.

a triangle or not, by examining whether it can it be constructed by the schema. Thus the concept itself is something passive, whereas the schema amounts to certain active construction procedures.

3.2 Schemata and universality

Central to the schematism is that the schematic rules are universal: "[T]his representation of a general procedure of the imagination for providing a concept with its image is what I call the schema for this concept" (A140/B180–1). In the case of geometrical schemata the problem is specifically the problem of drawing a universal conclusion based on a singular instance.

My understanding of Kant is that when we argue for a universal statement from a singular instance we use schemata in the following way. We want to prove that some property belongs generally to a mathematical concept:

1. We recognize a particular constructed figure—presented as a mental image or a figure drawn on paper—as a token of a type. We use schema or schemata for recognizing this.

2. Then we prove that the property we are after holds for this particular token.

3. We recognize that in this proof we have not used anything about that particular token which would not hold for any other token of that type. We use the schema in recognizing this.

4. Therefore, the type has the property, i.e., any token of that type has the property.

This is precisely what happens in the proof of I.32; "we have taken account only of the action of constructing the concept", and therefore we have not used properties such as "the magnitude of the sides and the angles" and since "we have abstracted from these differences" the figure represents the concept universally.[14] Now this is of course nothing but a very typical way of proving universal statements in mathematics. We prove that some property A holds for a, then we recognize that none of our assumptions concerns a, except that a is of a certain type, say, the natural numbers; therefore we can conclude that every x of that type must have property A. Precisely this way of reasoning is codified by the natural deduction rule:

$$\frac{\begin{array}{c}\vdots\\ A(a)\end{array}}{\forall x A(x)}$$

Let me give a simple example, where I assume that the concept of a natural number can be schematized (see my [12] for a detailed analysis of the schematism of the natural numbers).

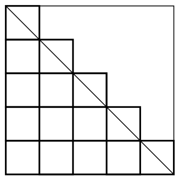
[14] All quotes are from (A713–4/B741–2).

Simple example. The equation

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}$$

holds for all natural numbers n.

Proof. We argue for the case where $n = 5$:



The diagram shows $1 + \cdots + 5$ placed in a square with area $5 \cdot 5$. To get $1 + \cdots + 5$ we divide the area of the square by 2 and add half of the elements on the diagonal. Therefore:

$$1 + 2 + \ldots + 5 = \frac{5 \cdot 5}{2} + \frac{5}{2} = \frac{5(5+1)}{2}.$$

The schematic justification of this proof—where there are at least two types of universality around— is the following. First we realize by the 'schema of number' that an empirical intuition, a square, really represents the concept of a particular number universally. Secondly, the property we prove about the particular token, is really—it is seen by geometrical schemata—about any token square. Finally, again by the 'schema of number' we realize that the property about any square can be transfered to the number any such square represents universally.[15]

3.3 Schemata and abstract concepts

One of the reasons why geometrical schemata can be successful in the way just described is that geometrical concepts are well-defined. As already mentioned this is in contrast with empirical concepts. According to Kant (A727/B755) we can give necessary and sufficient criteria for, say, triangles. Therefore, we can *decide*, in a finite amount of time, whether an intuition is an instance or not of some concept due to the schema.

The distinction between image and schema is a way of making abstract concepts possible, and thus the distinction can be seen as a refutation of a crude

[15]The example is found, for instance, in [3, 64]. Interestingly, Brown writes "The simple moral I want to draw from this example is just this: We can in special cases correctly infer theories from pictures, that is, from visualizable situations. An intuition is at work and from this intuition we can grasp the truth of the theorem." This sounds very much like Kant at A713/B742. Note, however, that Brown sees his own position as 'full-blooded Platonism'; a term which can hardly be said to characterize Kant's position.

(skeptic and) empiricist claim that mathematical concepts, as simple generalizations, are illusionary. It is in this way that schemata ground or found mathematical concepts, as claimed for instance on A141/B180, and described on A234/B287.

Generally Kant views schemata as the mediating element "which must stand in homogeneity with the category on the one hand and the appearance on the other" (A138/B177). This applies as well to empirical concepts. The schema of dog is mediating between the empirical dog and the concept of dog in that it produces a mental image which as an representation mediates between the empirical and the intellectual.

The situation is different in the case of geometrical and mathematical concepts: The objects are not empirical objects, but pure intuitions. This leads Shabel [15, p. 24] to conclude that "[i]n the case of mathematical concepts then, schemata are strictly redundant: no "third thing" is needed to mediate between a mathematical concept and the objects that instantiate it since mathematical concepts come equipped with determinate conditions on and procedures for their construction".[16] I think this is a somewhat peculiar view. Though certainly, as Shabel notes, mathematical concepts are well-defined. Thus we can define precisely what it means to be a triangle. But a definition does not necessarily include a description of the schema. As Kant writes on A716/B744 "[g]ive a philosopher the concept of a triangle"—from the concept alone, the philosopher can learn nothing about the sum of the interior angles. And, although the certain relation between a right angle and the sum of the interior angles is discoverable only by the schema, the relation itself belongs to concept, not to the schema. I therfore find it reasonable to view a geometrical concept as a (passive) type, images (or intuitions) as tokens, and the schema of the type as the constructive (i.e., active) *relation* between the type and its tokens. Such a constructive relation is to be understood, on the one hand, as our cognitive capacity for recognizing in finite time something as a token of a type; on the other hand it gives rise to a rule for constructing a paradigmatic and pure instantiation of that concept. The schema is therefore a *decidable and constructive* procedure in the sense that we can decide in finite time the type to which a token belongs, but also that we can 'go from type to token' through construction.

To sum up: The schema is the "mediating third" which connects the conceptual with the spatial. Take the example of the concept triangle; due to the schema of triangle, we can: recognize triangles as triangles; construct triangles and reason in a universal way about the full type, that is, all kinds of triangles. The ingredient is the beginning of a theory of diagrammatical reasoning, which I will no discuss here, however.

To pin it out with the help of mathematics: Let T be a concept and suppose $\overline{T}$ is the (Fregean) extension of T, i.e.,

$$\overline{T} = \{x \mid T(x)\}.$$

[16]In this respect she follows Guyer [9, p. 159].

The schema of T has (at least) four different functions: χ, Eq, f og Φ.

1. Recognition: For all x, $\chi(x) = 1$ iff $x \in \overline{T}$.

2. Representational equality: If two representations of a concept are equal, then the schema recognizes this; for all $x, y \in \overline{T}$:

$$x = y \text{ iff } \mathrm{Eq}(x, y) = 1.$$

3. Construction. For any $x \in \overline{T}$, there exists a finite amount of time, such that, $f(t) = x$.

4. Universal reasoning: Given any property P there is t and x such that $f(t) = x$ and

$$\Phi(P, x) = 1 \text{ iff } \text{ for alle } y \in \overline{T} : P(y).$$

All functions χ, Eq, f og Φ should decide in a finite amount of time. Furthermore, the functions are rule-governed and thus Kant's notion of schema actually anticipates the notion of algorithm.

I therefore interpret Kant's theory of schemata in the following way. An objective concept needs to be founded by a schema. Without such a schema a concept cannot be objective. A concept can be fully schematized by an agent, when the agent posseses some rule-governed procedure which in a finite amount of time recognizes and produces instances of the concept and allows for universal reasoning.[17] Let us return to the examples with the quotient sets and structures as given in the beginning of this essay.

4 Equivalence classes and schemata

Let us return to our example with the even and uneven whole numbers. The question now is, are these concepts schematisable in the strong Kantian sense? Given that the general concept a natural number can be schematized it is straight forward to realize that the whole numbers can be constructed—and thus schematized—from the natual numbers.

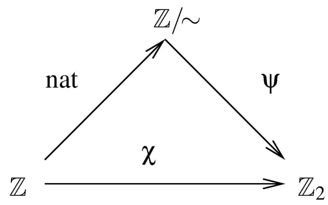
Let $=$ be equality between whole numbers,[18] and define the eveness of numbers as n is even if and only if there exists $|m| \leq |n|$ such that $n = 2m$.

$$\chi(n) = \begin{cases} 1, & \text{if } n \text{ is even,} \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Then we take $n \sim m$ iff $\chi(n) = \chi(m)$. Clearly, there is an algorithm which deciding this. Thus the concepts of even and uneven numbers can be schematised relatively to the whole numbers. Obviously we have:

[17]Thus the schema takes care of the justificational role of knowledge as found in the traditional view of knowledge as true, justified belief.

[18]This is assumed to be part of the schema of the general concept of a whole number.



The extension of the concept 'even-whole-number' is [0] and the extension of the concept 'uneven-whole-number' is [1]. Given that the natural numbers can be schematised in the strong Kantian sense, the concepts even and uneven which are concepts having $\mathbb{Z}/\!\sim$ as interpretation are schematisable.

The rational numbers $\mathbb{Q}$ can be treated in the same way. Suppose we are given $\mathbb{Z}$, then product of the whole numbers is constructable. Now the rational numbers can be constructed as equivalence classes of $\mathbb{Z} \times \mathbb{Z}$. Suppose $b \neq 0$, then

$$(a, b) \sim (c, d) \ \text{ iff } \ ad = bc.$$

Now, $[(a, b)]$ is understood as the rational number a/b. Thus the rational numbers can be schematised given that the naturals can.

5 Kantian schemata and Tait's thesis

It is certainly not trivial to see that even the natural numbers are schematisable in the strong Kantian sense, which includes that the schema enables finitary reasoning in the sense of the four functions χ, Eq, f and Φ as given on page 15. Already in the case of the natural numbers the aspect of universal reasoning can be questioned. The property 'any even number is the sum of two prime numbers' is not decided at the moment. Is it reasonable to assume that this is decidable in the strong sense, that we can produce a paradigmatic even natural number which is the sum of two prime numbers in such a way that we can produce a proof that enables generalisation to all even natural numbers? And what are the functions which are schematisable? Let us assume for a moment that the finitary functions are the functions which are schematisable in the strong Kantian sense. Tait's thesis now is, that the finitary functions are precisely the primitive recursive functions [18, 19, 20]:

> A function f is finitary, if and only if, f is primitive recursive.

The class of primitive recursive functions is the smallest class of functions:[19]

1. containing initial functions for zero, successor and projection.

[19] See [14, 22] for details.

2. closed under composition and under primitive recursion: Given primitive recursive g, h we have

$$\begin{aligned} f(x,0) &= g(x) \\ f(x,y+1) &= h(x,y,f(x,y)) \end{aligned}$$

Let us evaluate critically Tait's thesis.

The first problem we face is how to understand the expression 'the function f is finitary'. There are at least two possibilities:

1. 'The function f is finitary' means that given any x the operation f applied to x is epistemically unproblematic.

2. 'The function f is finitary' means that we have a *concept* f of a function. This concept has a corresponding schema which allows us to decide whether a given representation is an instance of f and to use this function, i.e., to compute it in the sense of 1.

I understand Tait as promoting the former understanding.[20] In fact, the former can be derived from the latter in most cases. Suppose we have a definition of a primitive recursive function f. This is given as a *finite* piece of text which is generated according to the rules given in the definition above. We can understand this as an intensional definition of f which works simultaneously as a proto type for our concept of f; any other (definition of a) primitive recursive function g is the same if it is defined in precisely the same way. Thus intentional equality between primitive recursive functions is reduced to equality between literal definitions of primitive recursive functions. As the latter equality is completely unproblematic, 2 reduces in the intensional case to 1. The extensional case is more complex and I will treat it below in the case of general recursive functions.

It seems, however, that Hilbert and Bernays perhaps had the intensional understanding of a finitary function in mind:

> A [finitary] function, for us, is an intuitive [*anschauliche*] instruction [*Anweisung*] which on the basis of a numeral, or a pair of numerals, or a triple of numerals, . . . , assigns another numeral. [10, 26]

If f is primitive recursive, then computing $f(x)$ is unproblematic with respect to 'complete surveyability in all parts', 'immediacy' and 'intuitivity':

5.1 Tait's "if"

Given a primitive recursive function f and any number x computing $f(x)$ is completely unproblematic. The only 'complicated' element in a definition of a primitive recursive function can be the two operations: substitution and iteration.

[20] "So how can the finitist understand $f : A \to B$ [. . .] he can understand it as recording the fact that he has given a specific procedure for *defining a* B *from an arbitrary* A or, we shall say, of *constructing* a B from an arbitrary A. [20, 24].

Say that g is given and that we define f by $f(0) = a_0$ and $f(x+1) = g(x, f(x))$. Suppose we want to compute $f(x)$. If x is 0 then a_0 is given. If x is not 0, then $x = 1$ or $x > 1$. In the first case we have $f(1) = g(1, a_0)$ which by assumption is finitarily computable—say the result is a_1. If $x > 1$ then $x = 2$ or $x > 2$. In the first case finitary reasoning gives us a_2. This process goes on until we reach x. The process is guaranteed to terminate as g is finitary and x is a (finitary) natural number.[21]

5.2 Tait's "only if"

Let us define a function φ in the following way. Suppose $\varphi_1(a, b) = a + b$ and $\varphi_2(a, b) = a \cdot b$ and that $\varphi_3(a, b) = a^b$. Furthermore, let $\varphi_4(a, b)$ be the b-th element of the sequence

$$a, \quad a^a, \quad a^{(a^a)}, \quad a^{(a^{(a^a)})}, \ldots$$

Continue an unfolding definition of φ_n in this way. That is, $\varphi_{n+1}(a, b)$ is an iteration of $\varphi_n(a, a)$ b times.

The Ackermann function $\varphi : \mathbf{N} \to \mathbf{N}$ is defined as $\varphi_n(n, n)$. It grows *really* fast:

$$\varphi(1) = 2, \qquad \varphi(2) = 4, \qquad \varphi(3) = 9, \qquad \varphi(4) = 4^{4294967296}.$$

The function is not a primitive recursive function as it majorizes any function which is primitive recursive. This was observed by Ackermann [2]. It is a relevant question, however, whether Hilbert considered (or should have considered) it to be finitary?

The Ackermann function is defined as a recursive function using nested recursion. This gives rise to a rule which allows that we can compute the values of φ from below. In principle we can compute any value of in the sequence $\varphi(1), \varphi(2), \ldots, \varphi(n), \ldots$. Arguments for finitariry are:

- It fulfills the 'surveyability', 'immediacy' and 'intuitivity' criteria.
- It is mentioned by Hilbert (1926) in the programmatic article.
- In volume II of *Grundlagen der Mathematik* Hilbert and Bernays write:

> Certain methods of finitistic mathematics which go beyond recursive number theory (in the original sense [i.e., primitive recursive]) have been discussed in §7 [of volume I of *Grundlagen*], namely the introduction of functions by nested recursion [e.g., the Ackermann function] and the more general induction schemata. [11, 354]

[21] The logician could make a counter argument here by asking: "How do we know that the given number x is not a non-standard number being infinitely large". This is actually a skeptic counter-argument. Well, we are not working with non-standard numbers. Our objects are not formal objects—the are contentual or semantic objects. Thus our numbers are not interpreted in some kind of model. They are numbers and not reducible to or interpretable in anything. But the problem becomes a real problem once we allow for ideal elements in the sense of Hilbert.

- Ackermann [1] uses transfinite induction up to ω^{ω^ω} for showing consistency of a second order version of PRA. In this system φ is definable . Moreover, this consistency proof was considered in the Hilbert school to be finitistic.

I find these reasons sufficient as a refutation of Tait's "only if".[22] But let me give another argument also.

Any primitive recursive function is defined by a finite piece of text. We can therefore provide an algorithm which enumerates all primitive recursive functions;

$$f_1, f_2, \ldots, f_n, \ldots$$

Based on this algorithm we can construct a finitary function U which is the universal function taking two arguments n and x and then picks the n-th primitive recursive function and applies it to x.[23]. In other words $U(n, x) = f_n(x)$. It can be argued that U is a finitary function. Thus also $U(n, n) + 1$ is finitary. It is, however, not primitive recursive. Assume it is. Then there would exists an m such that

$$f_m(n) = U(n, n) + 1. \tag{3}$$

On the other hand, because U is the universal function we have $f_m(m) = U(m, m)$, which, however, contradicts (3) if we substitute m for n. Therefore, there are more finitary functions than primitive recursive functions.

I therefore propose that the functions χ, Eq, f and Φ are not required to be primitive recursive functions, but *partial* recursive functions.

6 A modern theory of schemata and quasi-schemata

Partial recursive functions are functions from numbers to numbers. They can be computed mechanically, step by step. It is a very robust class of functions and they can be characterized in many different ways using Turing machines, Register machines, lambda calculus, domain theory and Kleene schemata. All these different formulations of the effective functions have turned out to be equivalent and this have led to the so-called Church-Turing thesis: The partial recursive functions characterize effective computability. In the following x can denote a sequence $x_1, \ldots, x_n$ of variables.

Definition. The class of partial recursive functions is the smallest class of functions:[24]

1. Containing initial functions for zero, successor and projection.

[22] A more detailed analysis is given by Richard Zach in his dissertation [21].
[23] For simplicity I have assumed that the primitive recursive functions are 1-ary.
[24] Here I use Kleene's symbol $\varphi \simeq \psi$ which is to be understood as "either φ and ψ are both undefined, or they are both defined with the same value".

2. Closed under composition and also under primitive recursion which is: Given ψ, γ we have

$$\begin{aligned}\varphi(x,0) &\simeq \psi(x)\\ \varphi(x,y+1) &\simeq \gamma(x,y,\varphi(x,y))\end{aligned}$$

3. Closed under unrestricted μ-recursion, i.e., given ψ we have

$$\varphi(x) \simeq \mu y\big(\forall z \leq y(\psi(x,z)\downarrow) \wedge \psi(x,y) \simeq 0\big).$$

The partial recursive functions are very well-behaved. They can be given on a certain normal form, and as any partial recursive function is defined by a finite piece of text, they can be enumerated. This is the Enumeration theorem: There is a sequence, $\varphi_0, \varphi_1, \ldots, \varphi_n, \ldots$ of partial recursive functions, such that any partial recursive function is within that enumeration. Let $\langle x \rangle$ be the primitive recursive encoding of the sequence x. Then there exists a universal partial recursive function $\varphi(e,x)$ such that for any partial recursive function ψ of n variables there is e such that

$$\psi(x) \simeq \varphi(e, \langle x \rangle).$$

As discussed above we could understand these finitary functions either intensionally of extensionally. The intensional understanding is simple and unproblematic, but what about the extensional notion.

Say that φ_n and φ_m are extensionally equal, in symbols,

$$\varphi_n \approx \varphi_m, \text{ if and only if, for every } x,\ \varphi_n(x) \simeq \varphi_m(x).$$

Then any definition of a partial recursive function φ_n generates an equivalence class $[\varphi_n]$. The question now is, is such a concept schematisable? It is not trivially so. If we consider $[\varphi_n]$ as a type, then membership of that type is not recursively decidable.[25]

We could also provide some basic schematic rules for *set theory*. We have a determinate concept of a set A in case we have a schema, say a partial recursive function, which constructs a representation of A and the schema also recognizes representations of the concept. Given that two sets A and B are determinate objects, then also $A \cup B$ and $A \cap B$ and (A, B) are determinate objects.

But there are also *quasi-schemata*. These are 'rules for constructing' more abstract objects, such as the first infinite number ω or a converging sequence of rational numbers taken as a completed object. Characteristic for these quasi-schemata is that we allow constructions to run in an infinite amount of time. This is in contrast with the real schemata which are rules running only in a finite amount of time, if they give output. Examples of quasi-schemata are:

[25]The question of extensionality of functions is an interesting question that I have looked at in [13].

1. Given a set A and a equivalence relation $\sim$ on A form the quotient set $A/\sim$.

2. Zorn's Lemma or any of its equivalents: The axiom of choice, the maximal chain principle or the well-ordering principle.

3. Extensionality of functions and functionals.

4. Taking limits.

5. Power set construction.

Sometimes these principles are ideal, sometimes they are not. The axiom of choice is in an intensional context with intuitionistic logic fully schematisable, whereas in a set-theoretic context we have that extensionality and the axiom of choice imply full classical logic, as shown in [4]. Power set constructions can be un-limited or limited, as they can for instance be restricted to finite subsets or only the first-order definable subsets.

A simple example of an ideal concept (in the sense of Hilbert) with schema could be the symbol ω which can be seen as denoting the object

$$\{n \mid n \text{ is a natural number.}\}$$

The schema for constructing this set is the quasi-schema: "Take the limit".

The notion of schema and quasi-schema give rise to *a relativised notion of construction* in mathematics. Here there are at least two different notions of constructions.

Definition. Suppose X is an already constructed object, whether ideal (in the sense of Hilbert) or not. Then we say that Y *is (quasi)-schematisable relatively to* X in case one of the following situations obtains, either

- Y can be constructed using a (quasi) schema on X, or
- Y is obtained by adding some already constructed (ideal) elements to X.

Thus there are basically two different types of ideal elements. Some simple examples are

1. $\mathbb{Z}$ is schematisable relatively to $\mathbb{N}$,

2. $\mathbb{Q}$ is schematisable relatively to $\mathbb{Z}$,

3. $\mathbb{R}$ is quasi-schematisable relatively to $\mathbb{Q}$,

4. $\mathbb{C}$ is schematisable relatively to $\mathbb{R}$,

5. Projective geometry is schematisable relatively to Euclidean geometry.

This notion of ideal elements is open towards some dynamical aspects of ideal elements. Our constructions and construction-procedures are not fixed once and for all. Thus it took historically a considerable amount of time to provide an interpretation of the (imaginary) complex numbers, which reduces a complex number to a point in the real plane. Also projective geometry arose by supplying Euclidean geometry with points at infinity. Later on it was discover that we can in fact give a a model of projective geometry within Euclidean geometry by moving up in dimension. There seems to be, however, a very fundamental border between that which can be finitely schematised and that which cannot. I find it very hard to believe that we can eliminate the ideal element when we construct the reals out of the rationals.

References

[1] Wilhelm Ackermann. Begründing des "tertium non datur" mittels der Hilbertschen Theorie der Wiederspruchsfreiheit. *Mathematische Annalen*, 93, 1924.

[2] Wilhelm Ackermann. Zum Hilbertschen Aufbau der reellen Zahlen. *Mathematische Annalen*, 99:118–133, 1928.

[3] James Robert Brown. Naturalism, Pictures, and Platonic Intuitions. In Paolo Mancosu, Klaus Frovin Jørgensen, and Stig Andur Pedersen, editors, *Visualization, Explanation, and Reasoning Styles in Mathematics*, Synthese Library, pages 57–73. Springer Verlag, 2005.

[4] Jesper Carlström. EM + Ext_{-} + $\mathrm{AC}_{\mathrm{int}}$ is equivalent to $\mathrm{AC}_{\mathrm{ext}}$. *Math. Log. Quart.*, 50(3):236–240, 2004.

[5] Euclid. *The Thirteen Books of Euclid's Elements Translated from the Text of Heiberg*. Dover Publications, New York, 1956. With introduction and commentary by T.L. Heath.

[6] Michael Friedman. *Kant and the Exact Sciences*. Harvard University Press, Cambridge, Massachusetts, 1992.

[7] Clark Glymore. *Thinking Things Through*. MIT, Cambridge, MA., 1992.

[8] Kurt Gödel. What is Cantor's continuum problem? In Paul Benacerraf and Hilary Putnam, editors, *Philosophy of Mathematics*, pages 470–485. Cambridge University Press, Cambridge, 2 edition, 1947/83.

[9] Paul Guyer. *Kant and the Claims of Knowledge*. Cambridge University Press, Cambridge, 1987.

[10] D. Hilbert and P. Bernays. *Grundlagen der Mathematik*, volume 1. Springer-Verlag, Berlin, 2. edition, 1934.

[11] D. Hilbert and P. Bernays. *Grundlagen der Mathematik*, volume 2. Springer-Verlag, Berlin, 2. edition, 1939.

[12] Klaus Frovin Jørgensen. Kant and the arithmetical schematism. In Henning Boje Andersen, Frederik Voetmann Christiansen, Vincent F. Hendricks, and Klaus Frovin Jørgensen, editors, *The Way through Science and Philosophy: Essays in Honour of Stig Andur Pedersen*, London. College Publications. Forthcoming, 2006.

[13] Klaus Frovin Jørgensen. *Kant's Schematism and the Foundations of Mathematics*. PhD thesis, Roskilde University, Roskilde, November 2005.

[14] Piergiorgio Odifreddi. *Classical recursion theory*, volume 125 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1989.

[15] Lisa Shabel. Kant's Philosophy of Mathematics. In Paul Guyer, editor, *The Cambridge Companion to Kant*. Cambridge University Press, 2. edition. The volume has not been published yet – page-references are to the final pre-print version.

[16] Lisa Shabel. *Mathematics in Kant's Critical Philosophy*. Routledge, 2003.

[17] Patrick Suppes. Interview on Formal Philosophy. In Vincent F. Hendricks and John Symons, editors, *Formal Philosophy*, pages 193–208. Automatic Press, 2005.

[18] W. Tait. Finitism. *Journal of Philosophy*, 78:524–56, 1981.

[19] W. Tait. Remarks on finitism. In Wilfried Sieg, Richard Sommer, and Carolyn Talcott, editors, *Reflections on the Foundations of Mathematics: Essays in Honor of Solomon Feferman*, volume 15 of *LNL*. Association for Symbolic Logic, 2002.

[20] William Tait. *The Provenance of Pure Reason*. Oxford University Press, Oxford, New York, 2005.

[21] Richard Zach. *Hilbert's Finitism*. PhD thesis, University of California, Berkeley, 2001.