Synthese (2008) 162:1–13
DOI 10.1007/s11229-007-9166-9

ORIGINAL PAPER

Categories for the working mathematician: making the impossible possible

Jessica Carter

Received: 1 June 2006 / Accepted: 12 February 2007 / Published online: 16 March 2007
© Springer Science+Business Media, Inc. 2007

Abstract This paper discusses the notion of necessity in the light of results from contemporary mathematical practice. Two descriptions of necessity are considered. According to the first, necessarily true statements are true because they describe 'unchangeable properties of unchangeable objects'. The result that I present is argued to provide a counterexample to this description, as it concerns a case where objects are moved from one category to another in order to change the properties of these objects. The second description concerns necessary 'structural properties'. Although I grant that mathematical statements could be considered as necessarily true in this sense, I question whether this justifies the claim that mathematics as a whole is necessary.

Keywords Philosophy of mathematical practice · Necessity

The history of mathematics abounds with results that could be described as 'making the impossible possible'. For example, introducing the imaginary unit makes it possible that the equation $x^2 + 1 = 0$ has a solution; working in $\mathbb{Z}_3$ makes it possible that $2 + 2$ is equal to 1; by rejecting Euclid's fifth postulate, it becomes possible that the sum of the angles of a triangle is not equal to two right angles. On the other hand, it is a common view that mathematical statements are necessarily true. Necessary, true statements are often described as 'statements that are true in all possible worlds'. But on considering examples like the ones above, one could question whether it makes sense to claim that mathematical statements are necessarily true. In order to

A version of this paper was presented at the meeting, 'Impact des Categories: Aspects Historiques et Philosophiques', held in Paris, Autumn 2005. I thank the audience for valuable comments and discussion. I especially wish to thank Colin McLarty for his invaluable support. Also thanks to the two anonymous referees for their helpful comments.

J. Carter (✉)
Department of Mathematics and Computer Science,
University of Southern Denmark,
Campusvej 55, Odense 5230, Denmark
e-mail: jessica@imada.sdu.dk

address this question two descriptions of necessity will be discussed: one in terms of unchangeable properties of unchangeable objects leading to the notion of possible worlds, and another in terms of propositions following necessarily from certain assumptions. I will consider these descriptions of necessity in the light of an example from contemporary mathematics. The example concern mathematics using category theory. In the light of the fact that category theory has today become an integrated part of many branches of mathematics, note that the use of category theory in this example does not make it exceptional.

When talking about truth in terms of possible worlds one needs to consider how to conceive of possible worlds in mathematics. I suggest that the result presented from category theory is different than certain other examples that could be taken as providing different worlds in mathematics. Thus, in view of the case study presented here, I will suggest that not all mathematical statements are necessarily true in the sense that they describe properties of unchangeable objects. Depending on how one understands the notion of possible worlds in mathematics, it could also be the case that mathematical statements are not true in all possible worlds. But it shall be admitted that, once appropriate assumptions have been laid out, certain statements follow necessarily. In the last part of the paper, I will discuss whether mathematics as a whole is necessary. Here my view will be that both mathematical thinking and concept formation depend on sources that admit the possibility that they 'must not necessarily be so'.

1 Necessity in mathematics

A number of philosophers of mathematics claim that mathematical statements are necessarily true.[1] In 'Realistic Rationalism' 1998, J. Katz argues for a realist position in the philosophy of mathematics, claiming that this is the only philosophy that can explain the special certainty of mathematical statements. The special certainty of mathematical statements is then explained by the necessary existence of the objects that the statements are about. Katz explains necessity as follows: "Necessary formal truths are necessary because they describe unchangeable properties and relations of unchangeable objects. Given that abstract objects can have neither spatial nor temporal location, no property of their intrinsic structure and no relation they bear to one another solely in virtue of their intrinsic structure is subject to change. ... Since the predication expressed by a formal truth holds in virtue of the fact that the property or relation that it attributes to the abstract objects it is about is one of their intrinsic (structural) properties and relations, there is no possibility that the objects do not have the property or relation"(pp. 78–79).

Even philosophers who criticize Katz's work approve of his claim that mathematical statements are necessarily true. In a paper commenting on Katz's book, A. Olivier writes "I do not deny that mathematical truths are necessary. I only claim that Katz has not *explained* their necessity, because he has not explained why mathematical objects exist necessarily, nor has he explained why mathematical objects have their intrinsic properties and relations necessarily" (Oliver 2000, p. 134, italics in original).

Not all philosophers claim that mathematical statements are necessary. In an insightful paper, Kraut points out that we ought to consider which data need to

[1] In addition to those mentioned here, see for example, Potter (2000), arguing for the necessity of mathematical statements, as well as Tennant (1997).

be explained: The necessity of mathematical statements or the *seeming* necessity of mathematical statements:

> *Even at the outset there is room for dispute. It is not clear whether the datum to be explained is mathematical necessity or the appearance of mathematical necessity. It is not clear whether we need a theory of non-spatial objects, or rather an explanation of why the objects of mathematical inquiry appear, from within the phenomenology of mathematical practice, to lack spatial position. It is not clear that the truths which appear to be graspable by intuitive reason alone are not somehow grounded in stimulatory inputs and computational processes explicable in naturalistic terms. Crudely speaking, it is not clear precisely what to count as the data. Nor is it clear what would count as an explication of the data, or what would make one explanation superior to another, or whether an explanation had accommodated the data or somehow lost sight of it.* (Kraut 2001, p. 154)

This is an important point that we shall bear in mind in this paper. Before I go on to discuss different notions of necessity we need to take a closer look at what are considered to be the data, namely mathematical practice.[2] In doing so, it is also the aim of this paper to throw some light on aspects of the practice of mathematics that consist in 'making the impossible possible'.

2 A result in algebraic topology

The result that we shall consider is, "A Mayer–Vietoris sequence constructed from controlled geometric modules is not exact. But it can be placed in a category of fractions in which it becomes exact."

The Mayer–Vietoris sequence has the following form:

$$Wh_0 \xrightarrow{d_1} Wh_1 \xrightarrow{d_2} Wh_2 \xrightarrow{d_3} K_0 \xrightarrow{d_4} K_0'$$

To claim that a sequence, $A \xrightarrow{f} B \xrightarrow{g} C$, is exact at B is to claim that the image of the map f ($\mathrm{im}(f)$) is equal to the kernel of the map g ($\ker(g)$). Set theoretically these are understood as follows. The kernel of a map is the set of objects that is mapped to zero, i.e., if $g\colon B \to C$ then

$$\ker(g) = \{b \in B | g(b) = 0\}.$$

The image of a map $f\colon A \to B$ is the set

$$\{b \in B | \exists a \in A \text{ and } f(a) = b\},$$

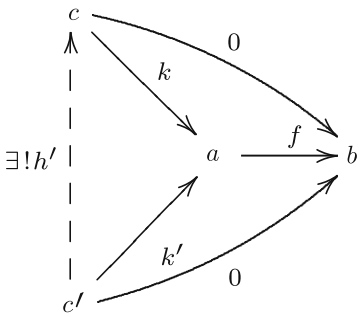
i.e., all the objects that are in the range of f. Both of these sets are subsets of B, and if they are equal, the sequence is said to be exact. In category theory, the definitions of kernel and image are made in terms of maps and without references to sets of objects. This entails—and this is one of the main points of this example—that the kernel and image of a map depend on in which category the map is placed.

[2] Recently a growing number of philosophers have begun to practice what has been denoted 'Practice oriented philosophy of mathematics', where the starting point for philosophical inquiry is mathematics as it is practiced. See, for example, Ferreirós and Gray (2006) and Larvor (2001).

For details of this study refer to (Carter (1994) or Carter (forthcoming). Here I present a few details that are relevant for this exposition.

A category $\mathcal{C}$ consists of a certain collection of objects $Ob(\mathcal{C})$ and for any two objects $b, c \in Ob(\mathcal{C})$, there is a set, morph(b, c), of morphisms (functions) between b and c. This collection may be empty, but an identity morphism 1_b must be contained in morph(b, b). Furthermore, if there are morphisms morph(b, c) and morph(a, b), then their composition must be in morph(a, c). Given two categories, $\mathcal{C}$ and $\mathcal{D}$, there then can be defined a map between these, a so-called functor, $F\colon \mathcal{C} \to \mathcal{D}$. A functor sends objects of $\mathcal{C}$ to objects of $\mathcal{D}$ and morphisms in $\mathcal{C}$ to morphisms in $\mathcal{D}$ subject to certain conditions. Furthermore, it is possible to define maps between functors, the so called natural transformations.

The category theoretical way of defining a kernel is as follows: Given a morphism $f\colon a \to b$, then the kernel of f is a morphism $k\colon c \to a$ such that the composition $f \circ k = 0$.[3] Furthermore, if there is another morphism $k'\colon c' \to a$ with this property, i.e., $f \circ k' = 0$, then there is a unique morphism $h'\colon c' \to c$ such that $k' = k \circ h'$ (k' factors uniquely through k):



The Mayer–Vietoris sequence that we consider is constructed from so called controlled geometric modules.[4] It is possible to prove that it is not exact.[5] The sequence is not exact because, although it does fulfil that $\mathrm{im}(d_i) \subseteq \ker(d_{i+1})$, the other inclusion does not hold, i.e., we do not have that $\ker(d_{i+1}) \subseteq \mathrm{im}(d_i)$. It can be proven, however, that if we compose with certain so called stabilization maps s, s', then we obtain the desired inclusion: $s \circ \ker(d_{i+1}) \subseteq s' \circ \mathrm{im}(d_i)$.

To make the sequence exact, one first puts it into an appropriate category $\mathcal{A}$ in which the structure of kernels and images is maintained such that the above inclusions still hold. To make the sequence exact, one needs to get rid of the stabilization maps in the inclusion $s \circ \ker(d_{i+1}) \subseteq s' \circ \mathrm{im}(d_i)$. The category in which this is possible is a so-called category of left fractions: Given a category $\mathcal{C}$ and a set of morphisms Σ in $\mathcal{C}$, fulfilling certain conditions, then there exists a category $(\Sigma^{-1})\mathcal{C}$ and a functor

[3] The zero morphism is defined in terms of a null object that, if it exists, is unique up to isomorphism. For further details, we recommend Mac Lane (1998).

[4] Geometric modules were originally introduced by Connell and Hollingsworth (1969) in order to prove a conjecture made by Whitehead (1939). The conjecture was that simple homotopy equivalence is a topological invariant: Given that two complexes A and B are topologically equivalent (i.e., that there exists a homeomorphism between them), then they are also simple homotopy equivalent. Control on geometric modules introduces a notion of size in the setting, i.e. on the modules and maps between them.

[5] This construction is made by Ranicki and Yamasaki (1995). This paper proves the before mentioned conjecture made by Whitehead.

$P\colon \mathcal{C} \to (\Sigma^{-1})\mathcal{C}$ such that P is the identity on objects and $P(s)$ is an isomorphism for $s \in \Sigma$. Simple examples of the construction of categories of left fractions are the construction of the category $\mathbb{Z}$ as the category of fractions of the category $\mathbb{N} \cup \{0\}$ and the construction of $\mathbb{Q}$ from $\mathbb{Z}$. In both cases the categories contain only one object and the morphisms are the numbers. In the case of $\mathbb{N}$, composition of two numbers is addition, and in the case of the construction of the rational numbers, composition is multiplication. What happens when constructing the category of left fractions is that additive and multiplicative inverses, respectively, are added.

It is possible to construct a set Σ in the category of geometric modules, $\mathcal{A}$, that corresponds to the stabilization maps, forming the category of left fractions, $(\Sigma)^{-1}(\mathcal{A})$. We have remarked that when moving from one category to another, one changes the structure of the morphisms. By forcing certain maps to become isomorphisms one also changes the structure of monic morphisms and epi-morphisms and thus the kernels as well as the images.[6] As the notion of exactness is defined in terms of kernels and images, the structure of exact sequences will also change in the new category. In this case, it is possible to describe the exact sequences in the category of fractions, $(\Sigma^{-1})\mathcal{A}$ in relation to the category $\mathcal{A}$. It turns out to be the case that $\ker(d_{i+1}) \subseteq \mathrm{im}(d_i)$ holds in the category of fractions when $s \circ \ker(d_{i+1}) \subseteq s' \circ \mathrm{im}(d_i)$ for $s, s' \in \Sigma$. Furthermore, the inclusions $\mathrm{im}(d_i) \subseteq \ker(d_{i+1})$ still hold in the category of fractions. Thus in the category of fractions the Mayer–Vietoris sequence becomes exact.

3 Making the impossible possible?

At a first glance, the result above could be described as 'Making the impossible possible'. In the set theoretical setting, the Mayer–Vietoris sequence is not exact and it is impossible that it could be exact. This result carries over when the sequence is placed in the category $\mathcal{A}$. But when forming the category of left fractions, one obtains the result that the sequence is exact.

This observation is not in any way unique to mathematical practice. It is a general feature of mathematics. Many examples can be found of something that could first not be done, but that mathematicians later somehow found a way of making possible. This was one purpose of Hilbert's method of ideal elements. Hilbert characterized this method as "extremely important and fertile" (Hilbert 1967, p. 372). The point of introducing new elements into a system is to make the propositions of the system as simple and perspicuous as possible. An important requirement was, of course, that extensions were consistent. Hilbert's favorite examples were the addition of the imaginary unit, i, to make all polynomial equations have a root and the addition of a line at infinity to obtain projective geometry.

In the present case, it is not the addition of new objects that does the trick. It is the change of setting that forces the morphisms to behave differently. Putting things this way almost makes it sound as if we move the objects from one world into another.

This brings us back to the discussion of necessity. According to one description of necessity, usually attributed to Leibniz, necessarily true means 'true in all

[6] In this case we are working in a so-called abelian category, where there is some extra structure imposed. In an abelian category, to be an isomorphism is equivalent to being both a monic morphism and an epi-morphism. Furthermore one of the conditions to be an abelian category is that every monic morphism is a kernel.

possible worlds'.[7] If it could be made out that different categories represented different worlds, then we would have encountered a result from mathematics that is not true in all possible worlds. However, it may not be clear that we can conceive of different categories as being different worlds. Furthermore one might object that much simpler examples from mathematics provide us with different worlds. For example, adding the imaginary unit i to the real numbers.

To respond to these objections it is necessary to think about how to conceive of 'possible worlds' in mathematics. A criterion that merely states, we are dealing with different possible worlds whenever we have two opposing statements in mathematics, seems too vague. There are a number of different descriptions of possible worlds in the literature; Kripke (1980) and Lewis (1986) present two quite different views. Here we compare (some of) Kripke's descriptions of possible worlds with different cases from mathematics in order to determine when it is appropriate to denote something 'a possible world'.

The cases to be considered are the following:

(i) The polynomial $p(x) = (x^2 + 1)(x^2 - 5) \in \mathbb{Q}[x]$ has no roots over $\mathbb{Q}$. But if it is placed in $\mathbb{C}$ (or just a field containing i and $\sqrt{5}$), it has a root. Should we consider $\mathbb{Q}$ and $\mathbb{C}$ as two different possible worlds?

(ii) A polynomial of the form $p(x) = x^3 + ax^2 + bx + c \in \mathbb{R}[x]$ always has a real root.

(iii) The Mayer–Vietoris sequence is not exact in the category, $\mathcal{A}$, but it is exact in the category of left fractions $(\Sigma^{-1})\mathcal{A}$. Do these categories form different worlds?

In 'Naming and Necessity', Kripke writes that a possible world is something that is stipulated: "A possible world isn't a distant country that we are coming across, or viewing through a telescope. . . . A possible world is *given by the descriptive conditions we associate with it*"(p. 44, italics in original). In relation to a theory of reference, one of Kripke's main points is that a name is a rigid designator which means that "in every possible world it designates the same object"(p. 48). One of his examples concerns Nixon and statements about whether Nixon could have lost the election. On this Kripke writes: "Why can't it be part of the *description* of a possible world that it contains *Nixon* and in that world *Nixon* didn't win the election. It might be a question, of course, whether such a world *is* possible. . . . the man who might have lost the election or did lose the election in this possible world is Nixon, because that's part of the description of the world"(p. 44).

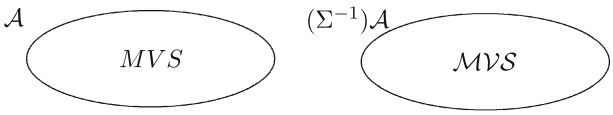
Let us consider the above examples in relation to Kripke's description of possible worlds. In the first case, it makes most sense to claim that the polynomial in question is the same polynomial, regardless of whether it is considered as a polynomial over $\mathbb{Q}$ or over $\mathbb{C}$. The question is whether we should regard $\mathbb{Q}$ and $\mathbb{C}$ as different possible worlds. Similarly, since we have a functor from $\mathcal{A}$ to $(\Sigma^{-1})\mathcal{A}$ that is the identity on objects, it also makes sense here to claim that the objects in the Mayer–Vietoris sequence are identical in the two categories. Can we then conceive of these categories as different possible worlds? I propose that $\mathbb{Q}$ and $\mathbb{C}$ do not constitute different worlds. $\mathbb{Q}$ is merely a subset of $\mathbb{C}$. But I believe the case of the two categories is different as here it makes sense to claim that they can be taken to be different worlds. In the case of the polynomial, it appears as if we move it from one world $\mathbb{Q}$ to another world $\mathbb{C}$. But

[7] See for example, Hacking (2000, p. 111).

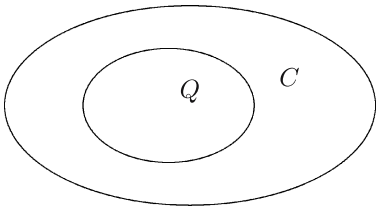
this view will not do when considering case (ii) above. To show that a polynomial of uneven degree has at least one real root, one regards the polynomial over the complex numbers and shows that complex roots always come in conjugate pairs. One may then conclude the stated theorem about real polynomials. If the complex numbers were not simply an extension of the real numbers, it would be hard to explain how this fact about a polynomial over the complex numbers could be transferred to polynomials over the real numbers. If one considers the complex numbers as a different world than the real numbers one has to come up with an explanation of why properties of the polynomial in this world also hold when it is placed in the world of real numbers. Even in the real world it is sometimes the case that we obtain counter-factual statements when talking about different domains: It is the case that my grandmother does not have any offspring in England, but she does have two children in Europe. This example illustrates the case when we consider the polynomial over $\mathbb{Q} \subset \mathbb{C}$.

The category theory case may be viewed such that we stipulate a new setting in which to place the objects from the Mayer–Vietoris sequence. Note also that via the functor that takes objects from the category $\mathcal{A}$ to the category of fractions, it is made explicit that the objects of the two categories are the same. The new setting is formed by stipulating that certain morphisms become isomorphisms. In the category theory case, we do not have the situation that certain theorems about, for example, the Mayer–Vietoris sequence in one category can be transferred to the category of fractions. What we do is prove conditional statements like 'if something is a kernel in $(\Sigma^{-1})\mathcal{A}$ then it will have a certain form in $\mathcal{A}$'. These theorems make it possible to deduce things about the objects in the category of fractions. We may picture the two cases as follows:

The category case:



The case with polynomials over different fields:



It is relevant to remark on Lewis's position on possible worlds. In his view the 'corresponding' objects in different worlds are never identical. Taking this view entails that we regard the polynomial from case (i) above as different when considered as an object over $\mathbb{Q}$ or over $\mathbb{C}$. This position is part of S. Shapiro's structuralism. In a previous paper (Carter (2005)), I have argued that this view is incompatible with mathematical practice.

Thus, I claim that there are differences between cases (i) and (iii), and I have argued, using Kripke's description of possible worlds, that it makes most sense to

consider case (iii) as giving us different possible worlds. I anticipate different kinds of objections to this claim. Note first that I do not exclude the possibility that there are cases from mathematics similar to case (iii) that do not use category theory. But in some cases that would appear to be similar, I doubt whether one could state that it is the same objects that are put in different worlds. Take, for example, the number 2 in $\mathbb{Z}$ and in $\mathbb{Z}_3$. In the first group, we have that $2+2=4$, but in the second, we have that $2+2=1$. In the group $\mathbb{Z}_3$, the objects can be construed in different ways. One way to think of them is as equivalence classes of the integers. Another is to describe the group as the set $\{0,1,2\}$ and specify the operation $+$. But the group could equally be described by any three-membered set with an appropriate relation. Thus I claim that it is not obvious that the number 2 is the same in the two groups. One could say that the description of $\mathbb{Z}_3$ one chooses is dependent on the context.

There may also be philosophers who would insist that we are dealing with possible worlds in all cases. Even if this is so, I still claim that there are important differences in how those different 'worlds' are obtained. It might also be altogether rejected that it makes sense to speak of different worlds in mathematics.[8] But no matter how it is put—in terms of different worlds or not—part of mathematical practice consists in constructing new settings in various ways to change properties of mathematical objects. My point remains that mathematicians constantly find new ways to make the impossible possible and that mathematics itself can be described as the science of making the 'impossible possible'.

Let us return to Katz's explanations of necessity. There are two statements:

(i) Necessary formal truths are necessary because they describe unchangeable properties and relations of unchangeable objects (Katz, p. 78).
(ii) ... formal truth holds in virtue of the fact that the property or relation that it attributes to the abstract objects it is about is one of their intrinsic (structural) properties and relations, there is no possibility that the objects do not have the property or relation (p. 79).

If we read 'unchangeable properties and relations of unchangeable objects' in terms of possible worlds, such that objects do not change when moved between different worlds, the first explanation resembles the definition of necessarily true as 'true in all possible worlds'. Again, if we grant that the objects appearing in the Mayer–Vietoris sequence are the same regardless of in which category they are placed, the case study provides a contradiction with respect to the explanation of necessity in terms of unchangeable properties of unchangeable objects. Properties of mathematical objects can change when the objects are placed in different settings. Even if the reading of 'unchangeable objects' in terms of possible worlds is rejected, I maintain that this conclusion holds. In recent literature, category theorists stresses the importance of context in mathematics. The traditional view about foundations is rejected: the view of a single foundation that suffices to formulate mathematics in: "No one doing category theory thinks we are someday going to find the one 'true topos', in which all mathematics happens"(Awodey 2004, p. 55). Instead Awodey points out that "the 'categorical-structural' [foundation] we advocate is based instead on the idea of specifying, for a given theorem or theory only the required or relevant degree

[8] One of the referees of this paper suggested that it only makes sense to talk about modality in certain cases. The reviewer proposed that it would make sense to talk about possible worlds from inside of certain 2-categories.

of information or structure, the essential features of a given situation, for the purpose at hand, without assuming some ultimate knowledge, specification, or determination of the 'objects' involved"(Awodey 2004, p. 56).

It is suggested in this quotation that, when using categeries in mathematics, what matters are relations between objects, or the structures that objects can be placed in, rather than 'intrinsic properties of objects'.[9]

When using only the first description of necessity, it seems that mathematical statements are not necessary. But this goes against a strong feeling that most people have when dealing with mathematics. Also the result described here does seem to follow necessarily, once the appropriate assumptions have been made.

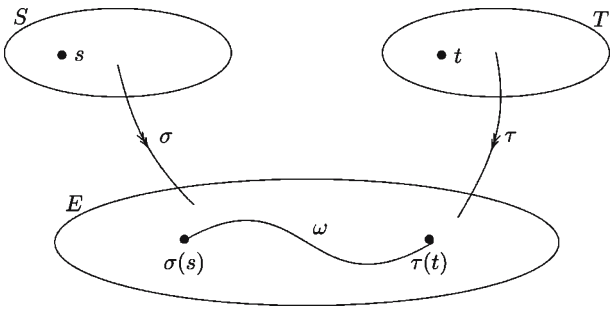
So could there be another way of describing the necessity of mathematical statements? This leads to Katz's second explanation of the necessity of formal truth. Here necessity is described in terms of intrinsic structural properties and relations. This opens up the possibility described in this paper, that the result depends on in which *structure* or context we place the objects, and that once these assumptions have been laid out, the result follows. Note that we here should read 'intrinsic structural properties' in relation to the context or structure in which an object is placed. Another way to address the differences between the descriptions of necessity is in terms of 'lack of specification'. For example the property of a polynomial of 'having a root' is not determined until it is specified in which domain the polynomial is contained. Similarly, one could claim that properties that vary across different categories fail to be completely specified.[10]

So we concede that once the appropriate (formal) definitions, structures, and relations have been given, then the result follows necessarily. The question that remains is: Is mathematics as a whole necessary?

My general view about this, based on mathematics as practiced, is as follows:

1. There is more to mathematical thought than formal definitions and formal settings.
2. Mathematical thought comes first, formalization comes second. Axioms or definitions are chosen to enable rigorous reasoning, i.e., such that results follow necessarily from axioms.

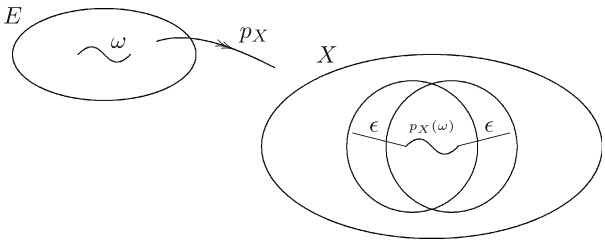
One could illustrate the first point by presenting the category of geometric modules. It would help readers to think about these entities by providing pictures. For example, a function $f\colon (S,\sigma) \to (T,\tau)$ between two geometric modules is composed of so-called Moore paths. A Moore path ω is easily illustrated by the picture:



[9] For examples of such category theoretical structures appearing in mathematics refer to Street (1995).

[10] Thanks to one of the referees for making this point.

The picture is not itself a formal definition of geometric modules or the maps between them. The picture merely represents these definitions. Nevertheless, pictures may be closer to what mathematicians think of when they look for the properties of geometric modules. A simple example: A morphism between two modules S and T is composed by paths starting in S and ending in T. Paths are composed by concatenation. It is possible to introduce a notion of size on paths by the following definition. We say that a path ω has size ϵ if the image of the path under a map to a metric space, p_X, is contained in the intersection of the closed ball with center in the image of the starting point of the path and radius epsilon and the closed ball with center in the image of the ending point and radius epsilon, i.e., in mathematical terms $p_X(w[0, l_w]) \subseteq B(p_X(w(0)), \epsilon) \cap B(p_X(w(l_w)), \epsilon)$. This is much easier to illustrate in a picture:



A proof of the fact that the composition of an ϵ path with a δ path yields an $\epsilon + \delta$ map is very easily found by drawing a picture like the one above. It is not at all apparent that one might realize the result holds simply by looking at the formal definition of the size of paths.[11] An even simpler example of when a picture can be helpful in seeing why a result holds consists of representing the algebraic equality $(a + b)^2 = a^2 + b^2 + 2ab$ by a square.

Regarding the second claim, I hold that neither mathematical thought nor the formation of axioms are necessary. The point that mathematical thought comes first and axiomatization second can be illustrated by some examples from the history of mathematics.

The first axiomatization of mathematics in European thought was made by Euclid in his Elements. It is usually claimed that what Euclid did was to take the already known theorems of, for example, geometry and order them into a theory, built up from five axioms.[12] Thus, in this case, the theory was already there and what Euclid did was extract the appropriate axioms.

Much later, in the 19th century, a number of problems had accumulated, threatening the status of mathematics as the science that provides secure knowledge. This led to the foundational crisis and the different foundational schools. As we are talking about axioms, it is natural to bring forth Hilbert's axiomatic method, which was designed to reinstate security in mathematics. Also in Hilbert, we find claims that axioms come only after the theory has been developed. In 'Axiomatic Thought' Hilbert

[11] In 'Proof and progress in mathematics', Thurston (1994, pp. 164–165) makes a similar claim regarding mathematical reasoning.

[12] Scholars are presently discussing whether Euclid did in fact come up with the axioms that can be found in the Elements. It has been argued that these axioms were added later. See, for example, Mueller (2003), who argues that some of the axioms were available in Euclid's period. In any case, my point remains that the axioms are found after the general theory.

expresses this as "When we assemble the facts of a definite, more or less comprehensive field of knowledge, we soon notice that these facts are capable of being ordered. . . .This framework of concepts is nothing other than the *theory* of the field of knowledge. . . . If we consider the theory more closely, we always see that a few distinguished propositions of the field of knowledge underlie the construction of the framework of concepts, and these propositions then suffice by themselves for construction, in accordance with logical principles, of the entire framework" (Hilbert, 1996, pp. 1107–1108, italics in original). Later in the paper Hilbert describes what he denotes as a deepening of the foundation, that consists of finding more 'fundamental' axioms for the theory. From this quote, as well as Hilbert's many examples, it is clear that Hilbert thinks of the axiomatic method as something that comes *after* a certain theory (or field of knowledge) has been developed. It is also worth noting that Hilbert's description of mathematical objects supports the point that there is more to mathematical thought (and mathematical objects) than formal definitions. As expressed in 'Natur und Mathematisches Erkennen': "With the example of the Continuum, we notice that the formation of concepts in mathematics is incited by thought and led by experience"[13](Hilbert 1992, p. 11, my translation). Regarding Hilbert's views on existence, reference is commonly made to the slogan 'Consistency implies existence' that Hilbert expressed on certain occasions. However, this slogan does not seem to encompass Hilbert's view on mathematical objects.[14] In 'On the infinite' 1967, Hilbert also wrote about success of procedures (p. 370) and, inspired by Kant, he claimed that some things are "given to our faculty of representations, certain extralogical concrete objects that are intuitively present as immediate experience prior to all thought"(Hilbert, 1967, p. 376). The notion of 'consistence implies existence' would presumably apply to ideal objects that are added to finitary mathematics.

Leibniz also defined absolute necessity in terms of logical necessity, such that denial of a necessary, true statement entails that the concepts involved dissolve. According to Leibniz mathematics deals with absolute, necessary truths, in the sense that if certain mathematical concepts, for example a circle and its radius, have been defined, then it will be a necessary truth that the radius has the same length no matter where it is measured. Denying this statement would entail that the concept of a circle is violated. However, one may question whether all mathematical concepts are necessary concepts. Leibniz is renowned for inventing calculus, which contained the infinitesimals. Leibniz's description of these as *useful fictions* bears similarities to Hilbert's descriptions of mathematical objects: "this study has provided a preliminary definition of infinitely small quantities as being useful, being created by the human mind, and being inaccessible to the senses. Furthermore, they represent measurable quantities without being measurable themselves" (Hedegaard forthcoming). Thus a comparison of Hilbert with Leibniz reveals certain common features. In both there is the possibility of rigorous mathematical reasoning. Hilbert pictures mathematics as a whole as a system containing certain concepts that are tied together by an inner necessity. Leibniz describes mathematical reasoning as dealing with combinatorial proofs by which all truths, by a chain of reductions analysing the concepts involved, can be transformed to certain basic truths. To Leibniz these basic truths consist of

[13] The quote is: So sehen wir an dem Beispiel des Kontinuums, wie die mathematische Begriffsbildung durch die Anschauung angeregt und von der Erfahrung geleitet wird.

[14] Peckhaus (2004, pp. 111–127) also makes a similar remark.

statements of identity (in terms of concepts containing each other). Thus mathematics is portrayed as a rigorous system containing necessary truths in the sense that 'statements are true in terms of their formal or intrinsic structural properties'. However, when it comes to the description of certain mathematical objects or the process of concept formation, both Leibniz and Hilbert give descriptions that deviate from this description of necessity. When writing about concept formation, Hilbert claims that mathematical objects are guided by our experience, whilst Leibniz describes the infinitesimals as being created by the human mind and as being useful fictions. Thus both descriptions involve an element of something external: experience, creations of the human mind and usefulness. That these external sources contribute to concept formation raises the possibility that the concepts are not formed by necessity.

For the final example, let us return to the case study presented above. This case study has roots in Euler's theorem about polyhedra. The notion of a polyhedron clearly has its origin as a formalization of "real" physical polyhedra. Gradually the notion of a polyhedron was changed. As is described in Lakatos's (1976) famous book, the notion was first developed in an interplay with the proof of Euler's theorem. Later, the definition of a polyhedron was turned into a definition of a complex. This definition was motivated partly by a wish to make the definition of a polyhedron more rigorous or formal, and partly by a desire to find a more general definition that could be used in arbitrary dimensions. This suggests that concept formation in mathematics sometimes consists of a long process, starting by extracting certain features from our physical reality, where the goal is to end up with a formal definition of a concept from which rigorous reasoning can take place. So perhaps this is the ultimate goal of mathematics, to obtain formal definitions of concepts so that we can obtain necessary knowledge from these definitions.[15] But it must be emphasized that we also regard the process that precedes the rigorous definition of mathematical concepts as part of mathematical practice.

Perhaps one might expect, when using category theory as a case study for discussing necessity in mathematics, that it would be possible to argue for necessity in a strong sense as 'true in all possible worlds'. Especially in light of the fact that category theory has been proposed as a foundation for mathematics. But our result has shown the contrary. All that category theory allows us to do is to move objects and maps between different categories, and in that respect, to change their properties. We could once again refer to Awodey's description of category theorists' conception of foundations for mathematics as something that is not fixed once and for all, but that provides frameworks that are extremely fruitful for current mathematical practice.

Summing up, we have considered two principal definitions of necessity in mathematics. One describes necessity in terms of 'unchangeable properties of unchangeable objects'. This description seems inadequate in the light of the case study presented here. A second description of necessity as 'intrinsic structural properties following from formal definitions' is, we conjecture, the ultimate goal of mathematics and perhaps what drives mathematicians, although it certainly does not capture all of mathematical practice.

[15] In the last century a number of books have been written in order to provide a certain area with a firmer foundation. We may refer to Eilenberg and Steenrod (1952), A. Weil's 'Foundations of Algebraic Geometry' (1946) and Weyl (1923).

References

Awodey, S. (2004). An answer to Hellman's question: Does Category theory provide a framework for mathematical structuralism? *Philosophia Mathematica, 12*, 54–64.

Carter, J. (1994). *e-controlled K-theory as an inverse system.* Master Thesis at the Department of Mathematics and Computer Science. University of Southern Denmark, Denmark.

Carter, J. (2005). Individuation of objects—A problem for structuralism? *Synthese, 143,* 291–307

Carter, J. (forthcoming). Structuralism as a Philosophy of Mathematical Practice, *Synthese.*

Connell, E., & Hollingsworth, J. (1969). Geometric groups and Whitehead torsion. *Trans. American Math. Soc., 140*, 161–181.

Eilenberg, S., & Steenrod, N. (1952). *Foundations of algebraic topology*. Princeton: Princeton University Press.

Ferreirós, J., & Gray, J. (2006). *The architecture of modern mathematics. Essays in history and philosophy*. New York: Oxford University Press.

Hacking, I. (2000). What mathematics has done to some and only some philosophers. In T. Smiley (Ed.), *Mathematics and necessity* (pp. 83–138). Oxford: Oxford University Press.

Hedegaard, L. (forthcoming). Fictions. On the relation between infinitesimals and monads in the leibnitian system. *Centaurus*.

Hilbert, D. (1992). *Natur und mathematisches Erkennen. Vorlesungen gehalten 1919–1920 in Göttingen*. Basel: Birkhäuser Verlag.

Hilbert, D. (1996). Axiomatic thought. In W. B. Ewald (Ed.), *From Kant to Hilbert. A sourcebook in the foundations of mathematics*. New York: Oxford University Press.

Hilbert, D. (1967). On the infinite. In J. Van Heijenoort (Ed.), *From Frege to Gödel*. Cambridge: Harvard University Press.

Katz, J. (1998). *Realistic rationalism*. Cambridge: MIT Press.

Kraut, R. (2001). Metaphysical explanation and the philosophy of mathematics: Reflections on Jerrold Katz's realistic rationalism. *Philosophia Mathematica, 9*, 154–183.

Kripke, S. (1980). *Naming and necessity*. Oxford: Basil Blackwell Publisher.

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery*. Cambridge: Cambridge University Press.

Larvor, B. (2001). What is dialectical philosophy of mathematics? *Philosophia Mahtematica, 9*, 212–229.

Lewis, D. (1986). *On the plurality of worlds*. Oxford: Basil Blackwell.

Mac Lane, S. (1998). *Categories for the working mathematician* (2nd ed.). New York: Springer Verlag.

Mueller, I. (2003). Remark on Euclid's Elements I, 32 and the parallel postulate. *Science in Context, 16*, 287–297.

Oliver, A. (2000). A realistic rationalism. *Inquiry, 43*, 111–136.

Peckhaus, V. (2004). Garantiert Widerspurchsfreiheit existenz? In B. Christiansen & U. Scheffler (Eds.), *Was folgt? Themen Zu Wessel (Logische philosophie)* (pp. 111–127). Berlin: Logos Verlag.

Potter, M. (2000). *Reason's nearest kin: Philosophy of arithmetic from Kant to Carnap.* Oxford: Oxford University press.

Ranicki, A., & Yamasaki, M. (1995). Controlled K-Theory. *Topology and its Applications, 61*, 1–59.

Street, R. (1995). Higher categories, strings, cubes and simplex equations. *Applied Categorical Structures, 3*, 29–77.

Tennant, N. (1997). On the necessary existence of numbers. *Nous, 31*, 307–336.

Thurston, W. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society, 30*(2), 161–177.

Weil, A. (1946). *Foundations of algebraic topology* (Vol. 29). New York: American Mathematical Society Colloquium Publications.

Weyl, H. (1923). *Die Idee der Riemannschen Fläche*. Berlin: B.G. Teubner.

Whitehead, J. H. C. (1939). Simplicial spaces, nuclei and *m*-groups. *Proceedings of London Mathematical Society, 45*, 243–327.