Chapter 2
The Foundational Crisis of Mathematics

A paradox is a situation that involves two or more facts or qualities which contradict each other.

Abstract The need for a formal definition of the concept of algorithm was made clear during the first few decades of the twentieth century as a result of events taking place in mathematics. At the beginning of the century, Cantor's naive set theory was born. This theory was very promising because it offered a common foundation to all the fields of mathematics. However, it treated infinity incautiously and boldly. This called for a response, which soon came in the form of logical paradoxes. Because Cantor's set theory was unable to eliminate them—or at least bring them under control—formal logic was engaged. As a result, three schools of mathematical thought—intuitionism, logicism, and formalism—contributed important ideas and tools that enabled an exact and concise mathematical expression and brought rigor to mathematical research.

2.1 Crisis in Set Theory

In this section we will describe the axiomatic method that was used to develop mathematics since its beginnings. We will also describe how Cantor applied the axiomatic method to develop his set theory. Finally, we will explain how the paradoxes revealed themselves in this theory.

2.1.1 Axiomatic Systems

The basic method used to acquire new knowledge in mathematics and similar disciplines is the *axiomatic method*. Euclid was probably the first to use it when he was developing his geometry.

B. Robič, *The Foundations of Computability Theory*,
DOI 10.1007/978-3-662-44808-3_2

Axiomatic Method

When using the axiomatic method, we start our treatment of the field of interest by carefully selecting a few *basic notions* and making a few basic statements,[1] called *axioms* (see Fig. 2.1). An axiom is a statement that asserts either that a basic notion has a certain property, or that a certain relation holds between certain basic notions. We do not try to define the basic notions, nor do we try to prove the axioms. The basic notions and axioms form our *initial theory* of the field.

We then start *developing* the theory, i.e., extending the initial knowledge of the field. We do this systematically. This means that we must *define* every new notion in a clear and precise way, using only basic or previously defined notions. It also means that we must try to *prove* every new proposition,[2] i.e., *deduce*[3] it only from axioms or previously proven statements. Here, a *proof* is a finite sequence of mental steps, i.e., inferences,[4] that end in the realization that the proposition is a logical consequence of the axioms and previously proven statements. A provable proposition is called a *theorem* of the theory. The process of proving is informal, *content-dependent* in the sense that each conclusion must undoubtedly follow from the *meaning* of its premises.

Informally, the development of a theory is a process of discovering (i.e., deducing) new theorems—in Nagel and Newman's words, as Columbus discovered America—and defining new notions in order to facilitate this process. (This is the *Platonic* view of mathematics; see Box 2.1.) We say that axioms and basic notions constitute an *axiomatic system*.

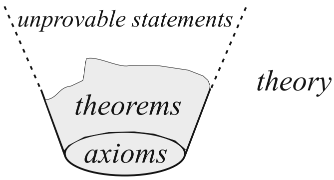

Fig. 2.1 A theory has axioms, theorems, and unprovable statements

Box 2.1 (Platonism).

According to the *Platonic view* mathematics *does not create* new objects but, instead, *discovers already existing* objects. These exist in the non-material world of abstract *Ideas*, which is accessible only to our intellect. For example, the idea of the number 2 exists per se, capturing the state of "twoness," i.e., the state of any gathering of anything and something else—and nothing else. In the

[1] A *statement* is something that we say or write that makes sense and is either true or false.

[2] A *proposition* is a statement for which a proof is either required or provided.

[3] A *deduction* is the process of reaching a conclusion about something because of other things (called premises) that we know to be true.

[4] An *inference* is a conclusion that we draw about something by using information that we already have about it. It is also the process of coming to a conclusion.

material world, *Ideas* present themselves in terms of imperfect copies. For example, the *Idea* of a triangle is presented by various copies, such as the figures $\triangle, \triangledown, \triangleleft, \triangleright$ (and love triangles too). It can take considerable instinct to discover an *Idea*, the comprehension of which is consistent with the sensation of its copies in the material world. The agreement of these two is the criterion for deciding as to whether the *Idea* really exists.

Evident Axiomatic Systems

From the time of Euclid to the mid-nineteenth century it was required that axioms be statements that are in perfect agreement with human experience in the particular field of interest. The validity of such axioms was beyond any doubt, because they were clearly confirmed by the reality. Thus, no proofs of axioms were required. Such axioms are *evident*. Euclidean elementary geometry is an example of an evident axiomatic system, because it talks of points, lines and planes, which are evident idealizations of the corresponding real-world objects.

However, in the nineteenth century serious doubts arose as to whether evident axiomatic systems are always appropriate. This is because it was found that experience and intuition may be misleading. (For an example of such a situation see Box 2.2.) This led to the concept of the *hypothetical* axiomatic system.

Box 2.2 (Euclid's Fifth Axiom).

In two-dimensional geometry a line parallel to a given line L is a line that does not intersect with L. Euclid's fifth axiom, also called the *Parallel Postulate*, states that at most one parallel can be drawn through any point not on L. (In fact, Euclid postulated this axiom in a different but equivalent form.)

To Euclid and other ancients the fifth axiom seemed less obvious than the other four of Euclid's axioms. This is because a parallel line can be viewed as a line segment that never intersects with L, even if it is extended indefinitely. The fifth axiom thus implicitly speaks about a certain occurrence in arbitrarily removed regions of the plane, that is, that the segment and L will never meet. However, since Aristotle the ancients were well aware that one has to be careful when dealing with infinity. For example, they were already familiar with the notion of asymptote, a line that "approaches a given curve but it meets the curve only in the infinity."

To avoid the vagueness and controversy of Euclid's fifth axiom, they undertook to deduce it from Euclid's other four axioms; these caused no disputes. However, all attempts were unsuccessful until 1868, when Beltrami proved that Euclid's fifth axiom *cannot be deduced* from the other four axioms. In other words, Euclid's fifth axiom is *independent* of Euclid's other four axioms.

NB *The importance of Beltrami's discovery is that it does not belong to geometry, but rather to the science* about *geometry, and, more generally, to* metamathematics, *the science about mathematics. About fifty years later, metamathematics would come to the fore more explicitly and play a key role in the events that led to a rigorous definition of the notion of the algorithm.*

Since the eleventh century, Persian and Italian mathematicians had tried to prove Euclid's fifth axiom indirectly. They tried to refute all of its alternatives. These stated either that there are no parallels, or that there are several different parallels through a given point. When they considered these alternatives they unknowingly discovered *non-Euclidean geometries*, such as elliptic and hyperbolic geometry. But they cast them away as having no evidence in reality. According to the usual experience they viewed reality as a space where only Euclidean geometry can rule.

In the nineteenth century, Lobachevsky, Bolyai, and Riemann thought of these geometries as true alternatives. They showed that if Euclid's fifth axiom is replaced by a different axiom, then a different non-Euclidean geometry is obtained. In addition, there exist in reality examples, also called *models*, of such non-Euclidean geometries. For instance, Riemann replaced Euclid's fifth axiom with the axiom that states that there is no parallel to a given line L through a given point. The resulting geometry is called *elliptic* and is modelled by a sphere. In contrast to this, Bolyai and Lobachevsky selected the axiom that allows several parallels to L to pass through a given point. The resulting *hyperbolic* geometry holds, for example, on the surface of a saddle.

NB *These discoveries shook the traditional standpoint that axioms should be obvious and clearly agree with reality. It became clear that intuition and experience may be misleading.*

Hypothetical Axiomatic Systems

After the realization that instinct and experience can be delusive, mathematics gradually took a more abstract view of its research subjects. No longer was it interested in the (potentially slippery) *nature* of the basic notions used in an axiomatic system. For example, arithmetic was no longer concerned with the question of *what* a natural number really is, and geometry was no longer interested in *what* a point, a line, and a plane really are. Instead, mathematics focused on the *properties* of and *relations* between the basic notions, which could be defined without specifying what the basic notions are in reality. A basic notion can be *any* object if it fulfills all the conditions given by the axioms.

Thus the role of the axioms has changed; now an axiom is only a *hypothesis*, a speculative statement about the basic notions taken to hold, although nothing is said about the true nature and existence of such basic notions. Such an axiomatic system is called *hypothetical*. For example, by using nine axioms Peano in 1889 described properties and relations typical of natural numbers without explicitly defining a natural number. Similarly, in 1899 Hilbert developed elementary geometry, where no explicit definition of a point, line and plane is given; instead, these are defined implicitly, only as possible objects that satisfy the postulated axioms.

Because the nature of basic notions lost its importance, also the requirement for the evidence of axioms as well as their verifiability in reality was abandoned. The obvious link between the subject of mathematical treatment and reality vanished. Instead of axiomatic evidence the fertility of axioms came to the fore, i.e., the number of theorems deduced, their expressiveness, and their influence. The reasonableness and applicability of the theory developed was evaluated by the importance of successful *interpretations*, i.e., applications of the theory on various domains of reality. Depending on this, the theory was either accepted, corrected, or cast off.

NB *This freedom, which arose from the hypothetical axiomatic system, enabled scientists to make attempts that eventually bred important new areas of mathematics. Set theory is such an example.*[5]

2.1.2 Cantor's Naive Set Theory

A theory with a hypothetical axiomatic system, which will play a special role in what follows, was the *naive set theory* founded by Cantor.[6] Let us take a quick look at this theory.

Fig. 2.2 Georg Cantor (Courtesy: See Preface)

Basic Notions, Concepts and Axioms

In 1895 Cantor defined the concept of a set as follows:

Definition 2.1. (Cantor's Set) A **set** is any collection of definite, distinguishable objects of our intuition or of our intellect to be conceived as a whole (i.e., regarded as a single unity).

Thus, an object can be any thing or any notion, such as a number, a pizza, or even another set. If an object x is in a set $\mathcal{S}$, we say that x is a *member* of $\mathcal{S}$ and write $x \in \mathcal{S}$. When x is not in $\mathcal{S}$, it is not a member of $\mathcal{S}$, so $x \notin \mathcal{S}$. Given an object x and a set $\mathcal{S}$, either $x \in \mathcal{S}$ or $x \notin \mathcal{S}$ — there is no third choice. This is known as the *Law of Excluded Middle*.[7]

---

[5] Another example of a theory with a hypothetical axiomatic system is group theory.

[6] Georg Cantor, 1845–1918, German mathematician.

[7] The law states that for any logical statement, either that statement is true, or its negation is — there is no third possibility (Latin *tertium non datur*).

Cantor did not develop his theory from explicitly written axioms. However, later analyses of his work revealed that he used three principles in the same fashion as axioms. For this reason we call these principles **Axioms of Extensionality**, **Abstraction**, and **Choice**. Let us describe them.

Axiom 2.1 (Extensionality). *A set is completely determined by its members.*

Thus a set is completely described if we list all of its members (by convention between braces "{" and "}"). For instance, $\{\diamond, \triangleleft, \circ\}$ is a set whose members are $\diamond, \triangleleft, \circ$, while one of the three members of the set $\{\diamond, \{\triangleleft, \triangleright\}, \circ\}$ is itself a set. When a set has many members, say a thousand, it may be impractical to list all of them; instead, we may describe the set perfectly by stating the *characteristic property* of its members. Thus a set of objects with the property P is written as $\{x \mid x \text{ has the property } P\}$ or as $\{x \mid P(x)\}$. For instance, $\{x \mid x \text{ is a natural number } \wedge\ 1 \leqslant x \leqslant 1000\}$.

What property can P be? Cantor's liberal-minded standpoint in this matter is summed up in the second axiom:

Axiom 2.2 (Abstraction). *Every property determines a set.*

If there is no object with a given property, the set is *empty*, that is, $\{\}$. Due to the *Axiom of Extensionality* there is only one empty set; we denote it by $\emptyset$.

Cantor's third principle is summed up in the third axiom:

Axiom 2.3 (Choice). *Given any set $\mathcal{F}$ of nonempty pairwise disjoint sets, there is a set that contains exactly one member of each set in $\mathcal{F}$.*

We see that the set and the membership relation $\in$ are such basic notions that Cantor defined them informally, in a descriptive way. Having done this he used them to define rigorously other notions in a true axiomatic manner. For example, he defined the relations $=$ and $\subseteq$ on sets. Specifically, two sets $\mathcal{A}$ and $\mathcal{B}$ are *equal* (i.e., $\mathcal{A} = \mathcal{B}$) if they have the same objects as members. A set $\mathcal{A}$ is a *subset* of a set $\mathcal{B}$ (i.e., $\mathcal{A} \subseteq \mathcal{B}$) if every member of $\mathcal{A}$ is also a member of $\mathcal{B}$. Cantor also defined the operations $\overline{\cdot}, \cup, \cap, -, 2^{\cdot}$ that construct new sets from existing ones. For example, if $\mathcal{A}$ and $\mathcal{B}$ are two sets, then also the *complement* $\overline{\mathcal{A}}$, the *union* $\mathcal{A} \cup \mathcal{B}$, the *intersection* $\mathcal{A} \cap \mathcal{B}$, the *difference* $\mathcal{A} - \mathcal{B}$ and the *power set* $2^{\mathcal{A}}$ are sets.

Application

Cantor's set theory very quickly found applications in different fields of mathematics. For example, Kuratowski used sets to define the *ordered pair* (x,y), i.e., a set of two elements with one being the first and the other the second in some order. The definition is $(x,y) \stackrel{\text{def}}{=} \{\{x\},\{x,y\}\}$. (The ordering of $\{x\}$ and $\{x,y\}$ is implicitly imposed by the relation $\subseteq$, since $\{x\} \subseteq \{x,y\}$, but not vice versa.) Two ordered pairs are equal if they have equal first elements and equal second elements. Now the *Cartesian product* $\mathcal{A} \times \mathcal{B}$ could be defined as the set of all ordered pairs (a,b), where $a \in \mathcal{A}$ and $b \in \mathcal{B}$. The sets $\mathcal{A}$ and $\mathcal{B}$ need not be distinct. In this case, $\mathcal{A}^2$ was used to denote $\mathcal{A} \times \mathcal{A}$ and, in general, $\mathcal{A}^n \stackrel{\text{def}}{=} \mathcal{A}^{n-1} \times \mathcal{A}$, where $\mathcal{A}^1 = \mathcal{A}$.

Many other important notions and concepts, which were in common use although informally defined, were at last rigorously defined in terms of set theory, e.g., the concepts of function and natural number. For example, a *function* $f : \mathcal{A} \to \mathcal{B}$ is a set of ordered pairs (a,b), where $a \in \mathcal{A}$ and $b = f(a) \in \mathcal{B}$, and there are no two ordered pairs with equal first components and different second components. Based on this, set-theoretic definitions of injective, surjective, and bijective functions were easily made. For example, a bijective function is a function $f : \mathcal{A} \to \mathcal{B}$ whose set of ordered pairs contains, for each $b \in \mathcal{B}$, at least one ordered pair with the second component b (surjectivity), and there are no two ordered pairs having different first components and equal second components (injectivity).

Von Neumann used sets to construct *natural numbers* as follows. Consider number 2. We may imagine that it represents the state of "twoness," i.e., the gathering of one element and one more different element—and nothing else. Since the set $\{0,1\}$ is an example of such a gathering, we may define $2 \stackrel{\text{def}}{=} \{0,1\}$. Similarly, if we imagine 3 to represent "threeness," we may define $3 \stackrel{\text{def}}{=} \{0,1,2\}$. Continuing in this way, we arrive at the general definition $n \stackrel{\text{def}}{=} \{0,1,2,\ldots,n-1\}$. So a natural number can be defined as a *set* of all of its predecessors. What about the number 0? Since 0 has no natural predecessors, the corresponding set is empty. Hence the definition $0 \stackrel{\text{def}}{=} \emptyset$. We can now see that natural numbers can be constructed from $\emptyset$ as follows: $0 \stackrel{\text{def}}{=} \emptyset$; $1 \stackrel{\text{def}}{=} \{\emptyset\}$; $2 \stackrel{\text{def}}{=} \{\emptyset,\{\emptyset\}\}$; $3 \stackrel{\text{def}}{=} \{\emptyset,\{\emptyset\},\{\emptyset,\{\emptyset\}\}\}$; ... ; $n+1 \stackrel{\text{def}}{=} n \cup \{n\}$; ... Based on this, other definitions and constructions followed (e.g., of rational and real numbers).

NB *Cantor's set theory offered a simple and unified approach to all fields of mathematics. As such it promised to become the foundation of all mathematics.*

But Cantor's set theory also brought new, quite surprising discoveries about the so-called cardinal and ordinal numbers. As we will see, these discoveries resulted from Cantor's *Axiom of Abstraction* and his view of infinity. Let us go into details.

Cardinal Numbers

Intuitively, two sets have the same "size" if they contain the same number of elements. Without any counting of their members we can assert that two sets are *equinumerous*, i.e., of the same "size," *if* there is a bijective function mapping one set onto the other. This function pairs every member of one set with exactly one member of the other set, and vice versa. Such sets are said to have the same *cardinality*. For example, the sets $\{\diamond,\triangleleft,\circ\}$ and $\{a,b,c\}$ have the same cardinality because $\{(\diamond,a),(\triangleleft,b),(\circ,c)\}$ is a bijective function. In this example, each of the sets has cardinality ("size") 3, a *natural* number. We denote the cardinality of a set $\mathcal{S}$ by $|\mathcal{S}|$.

Is cardinality always a natural number? Cantor's *Axiom of Abstraction* guarantees that the set $\mathcal{S}_P = \{x \mid P(x)\}$ exists for *any* given property P. Hence, it exists also for a P which is shared by infinitely many objects. For example, if we put

$P \equiv$ "is natural number," we get a set of *all* natural numbers. This set is not only an interesting and useful mathematical object, but (according to Cantor) it also *exists* as a perfectly defined and accomplished unity. Usually, it is denoted by $\mathbb{N}$. It is obvious that the cardinality of $\mathbb{N}$ cannot be a natural number because any such number would be too small. Thus Cantor was forced to introduce a new kind of number and designate it with some new symbol not used for natural numbers. He denoted this number by $\aleph_0$ (read *aleph zero*[8]). Cardinality of sets can thus be described by the numbers that Cantor called *cardinal numbers*. A *cardinal number* (or *cardinal* for short) can either be *finite* (in that case it is natural) or *transfinite*, depending on whether it measures the size of a finite or infinite set. For example, $\aleph_0$ is a transfinite cardinal which describes the size of the set $\mathbb{N}$ as well as the size of any other infinite set whose members can all be listed in a sequence.

Does every infinite set have the cardinality $\aleph_0$? Cantor discovered that this is not so. He proved (see Box 2.3) that the cardinality of a set $\mathcal{S}$ is strictly less than the cardinality of its power set $2^{\mathcal{S}}$ — *even when $\mathcal{S}$ is infinite!* Consequently, there are larger and larger infinite sets whose cardinalities are larger and larger transfinite cardinals—and this never ends. He denoted these transfinite cardinals by $\aleph_1, \aleph_2, \ldots$. Thus, there is no largest cardinal.

Cantor also discovered (using diagonalization, a method he invented; see Sect. 9.1) that there are more real numbers than natural ones, i.e., $\aleph_0 < \mathsf{c}$, where c denotes the cardinality of $\mathbb{R}$, the set of real numbers. (For the proof see Example 9.1 on p. 193.) He also proved that $\mathsf{c} = 2^{\aleph_0}$. But where is c relative to $\aleph_0, \aleph_1, \aleph_2, \ldots$? Cantor conjectured that $\mathsf{c} = \aleph_1$, that is, $2^{\aleph_0} = \aleph_1$. This would mean that there is no other transfinite cardinal between $\aleph_0$ and c and consequently there is no infinite set larger than $\mathbb{N}$ and smaller than $\mathbb{R}$. Yet, no one succeeded in proving or disproving this conjecture, until Gödel and Cohen finally proved that *neither* can be done (see Box 4.3 on p. 59). Cantor's conjecture is now known as the *Continuum Hypothesis*.

Box 2.3 (Proof of Cantor's Theorem).

Cantor's Theorem states: $|\mathcal{S}| < |2^{\mathcal{S}}|$, *for every set $\mathcal{S}$.*

Proof. (a) First, we prove that $|\mathcal{S}| \leq |2^{\mathcal{S}}|$. To do this, we show that $\mathcal{S}$ is equinumerous to a subset of $2^{\mathcal{S}}$. Consider the function $f : \mathcal{S} \to 2^{\mathcal{S}}$ defined by $f : x \mapsto \{x\}$. This is a bijection from $\mathcal{S}$ onto $\{\{x\} | x \in \mathcal{S}\}$, which is a subset of $2^{\mathcal{S}}$. (b) Second, we prove that $|\mathcal{S}| \neq |2^{\mathcal{S}}|$. To do this, we show that there is no bijection from $\mathcal{S}$ onto $2^{\mathcal{S}}$. So let $g : \mathcal{S} \to 2^{\mathcal{S}}$ be an *arbitrary* function. Then g cannot be surjective (and hence, neither is it bijective). To see this, let $\mathcal{N}$ be a subset of $\mathcal{S}$ defined by $\mathcal{N} = \{x \in \mathcal{S} \,|\, x \notin g(x)\}$. Of course, $\mathcal{N} \in 2^{\mathcal{S}}$. But $\mathcal{N}$ is not a g-image of any member of $\mathcal{S}$. *Suppose* it were. Then there would be an $m \in \mathcal{S}$, such that $g(m) = \mathcal{N}$. Where would be m relative to $\mathcal{N}$? If $m \in \mathcal{N}$, then $m \notin g(m)$ (by definition of $\mathcal{N}$), and hence $m \notin \mathcal{N}$ (as $g(m) = \mathcal{N}$)! Conversely, if $m \notin \mathcal{N}$, then $m \notin g(m)$ (as $g(m) = \mathcal{N}$), and hence $m \in \mathcal{N}$ (by definition of $\mathcal{N}$)! This is a contradiction. We conclude that g is not a surjection, and therefore neither is it a bijection. Since g was an arbitrary function, we conclude that there is no bijection from $\mathcal{S}$ onto $2^{\mathcal{S}}$. □

[8] $\aleph$ is the first symbol of the Hebrew alphabet.

Ordinal Numbers

We have seen that one can introduce order into a set of *two* elements. This can easily be done with other finite and infinite sets, and it can be done in many different ways. Of special importance to Cantor was the so-called *well-ordering*, because this is the way natural numbers are ordered with the usual relation $\leqslant$. For example, each of the sets $\{0,1,2\}$ and $\mathbb{N}$ is well-ordered with $\leqslant$, that is, $0<1<2$ and $0<1<2<3<\cdots$. (Here $<$ is the strict order corresponding to $\leqslant$.) But well-ordering can also be found in other sets and for relations other than the usual $\leqslant$. When two well-ordered sets differ only in the naming of their elements or relations, we say that they are *similar*.

Cantor's aim was to *classify* all the well-ordered sets according to their similarity. In doing so he first noticed that the usual well-ordering of the set $\{0,1,2,\ldots,n\}$, $n \in \mathbb{N}$, can be represented by a single natural number $n+1$. (We can see this if we construct $n+1$ from $\emptyset$, as von Neumann did.) For example, the number 3 represents the ordering $0<1<2$ of the set $\{0,1,2\}$. But the usual well-ordering of the set $\mathbb{N}$ cannot be described by a natural number, as any such number is too small. Once again a new kind of a "number" was required and a new symbol for it was needed. Cantor denoted this number by ω and called it the *ordinal number*.

Well-ordering of a set can thus be described by the *ordinal number*, or *ordinal* for short. An ordinal is either *finite* (in which case it is natural) or *transfinite*, depending on whether it represents the well-ordering of a finite or infinite set. For example, ω is the transfinite ordinal that describes the usual well-ordering in $\mathbb{N}$. Of course, in order to use ordinals in classifying well-ordered sets, Cantor required that two well-ordered sets have the same ordinal *iff* they are similar. (See details in Box 2.4.) Then, once again, he proved that there are larger and larger transfinite ordinals describing larger and larger well-ordered infinite sets—and this never ends. There is no largest ordinal.

NB *With his set theory, Cantor boldly entered a curious and wild world of infinities.*

2.1.3 Logical Paradoxes

Unfortunately, the great leaps forward made by Cantor's set theory called for a response. This came around 1900 when logical paradoxes were suddenly discovered in this theory. A *paradox* (or *contradiction*) is an unacceptable conclusion derived by apparently acceptable reasoning from apparently acceptable premises (see Fig. 2.3).

Burali-Forti's Paradox. The first logical paradox was discovered in 1897 by Burali-Forti.[9] He showed that in Cantor's set theory there exists a well-ordered set Ω whose ordinal number is *larger than itself*. But this is a contradiction. (See details in Box 2.4.)

[9] Cesare Burali-Forti, 1861–1931, Italian mathematician.

Cantor's Paradox. A similar paradox was discovered by Cantor himself in 1899. Although he proved that, for any set $\mathcal{S}$, the cardinality of the power set $2^{\mathcal{S}}$ is strictly larger than the cardinality of $\mathcal{S}$, he was forced to admit that this cannot be true of *the set* $\mathcal{U}$ *of all sets*. Namely, the existence of $\mathcal{U}$ was guaranteed by the *Axiom of Abstraction*, just by defining $\mathcal{U} = \{x \,|\, x = x\}$. But if the cardinality of $\mathcal{U}$ is less than the cardinality of $2^{\mathcal{U}}$, which also exists, then $\mathcal{U}$ is not the largest set (which $\mathcal{U}$ is supposed to be since it is the set of *all* sets). This is a contradiction.

Russell's Paradox. The third paradox was found in 1901 by Russell.[10] He found that in Cantor's set theory there exists a set $\mathcal{R}$ that *both is and is not* a member of itself. How? Firstly, the set $\mathcal{R}$ defined by

$$\mathcal{R} = \{\mathcal{S} \mid \mathcal{S} \text{ is a set} \wedge \mathcal{S} \text{ does not contain itself as a member}\}$$

must exist because of the *Axiom of Abstraction*. Secondly, the *Law of Excluded Middle* guarantees that $\mathcal{R}$ either contains itself as a member (i.e., $\mathcal{R} \in \mathcal{R}$), or does not contain itself as a member (i.e., $\mathcal{R} \notin \mathcal{R}$). But then, using the definition of $\mathcal{R}$, each of the two alternatives implies the other, that is, $\mathcal{R} \in \mathcal{R} \Longleftrightarrow \mathcal{R} \notin \mathcal{R}$. Hence, each of the two is both a true and a false statement in Cantor's set theory.

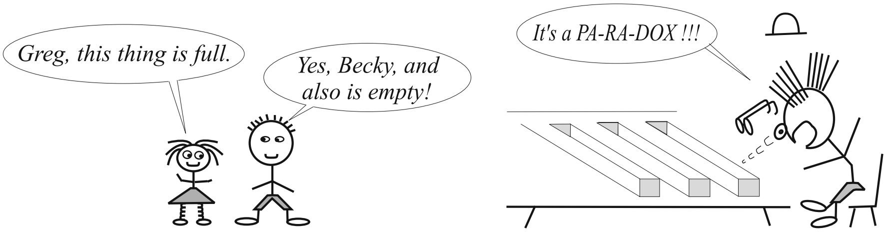

Fig. 2.3 A paradox is an unacceptable statement or situation because it defies reason; for example, because it is (or at least seems to be) both true and false

Why Do We Fear Paradoxes?

Suppose that a theory contains a logical statement such that both the statement and its negation can be deduced. Then it can be shown (see Sect. 4.1.1) that *any* other statement of the theory can be deduced as well. So in this theory everything is deducible! This, however, is not as good as it may seem at first glance. Since deduction is a means of discovering truth (i.e., what is deduced is accepted as true) we see that in such a theory every statement is true. But a theory in which everything is true has no cognitive value and is of no use. Such a theory must be cast off.

---

[10] Bertrand Russell, 1872–1970, British mathematician, logician, and philosopher.

Box 2.4 (Burali-Forti's Paradox).

A set $\mathcal{S}$ is *well-ordered* by a relation $\prec$ if the following hold: 1) $a \not\prec a$; 2) $a \neq b \Rightarrow a \prec b \vee b \prec a$; and 3) every nonempty $\mathcal{X} \subseteq \mathcal{S}$ has $m \in \mathcal{X}$, such that $m \prec x$ for every other $x \in \mathcal{X}$. For example, $\mathbb{N}$ is well-ordered with the usual relation $<$ on natural numbers. Well-ordering is a special case of the so-called *linear ordering*, i.e., a well-ordered set is also linearly ordered. For example, $\mathbb{Z}$, the set of integers, is linearly ordered by the usual relation $<$.

Suppose we do not want to distinguish between two linearly ordered sets that differ only in the naming of their elements or relations. We want to consider such sets as being similar, because they obviously share the same "type of order."

Let us define the notion "type of order" precisely. Let two sets $\mathcal{A}$ and $\mathcal{B}$ be linearly ordered with relations $\prec_{\mathcal{A}}$ and $\prec_{\mathcal{B}}$, respectively. We say that $\mathcal{A}$ and $\mathcal{B}$ are *similar* if there is a bijection $f : \mathcal{A} \to \mathcal{B}$ such that $a \prec_{\mathcal{A}} b \Longleftrightarrow f(a) \prec_{\mathcal{B}} f(b)$. The function f renames the elements of $\mathcal{A}$ to the elements of $\mathcal{B}$ while respecting both relations. We can easily prove that similarity is an equivalence relation between linearly ordered sets. So we can define the *order type* to be an equivalence class of similar, linearly ordered sets. Informally, an order type is the feature shared by all linearly ordered sets that differ only in the naming of their elements and relations.

Having defined the order types we might want to compare them. Unfortunately, they may not be comparable. It can be shown, however, that *order types of well-ordered sets* are themselves *linearly* ordered by some relation $\prec_o$. (Actually, $\prec_o$ is the usual set-membership relation $\in$.) Because such order types are ordered in a similar way to integers, we call them *ordinal numbers* (or *ordinals* for short). Hence, the definition: An *ordinal* is an equivalence class of similar well-ordered sets. For example, sets similar to $\{0, 1, \ldots, n\}$ have the same ordinal; we denote it by the natural number $n+1$. This cannot be done with sets similar to $\mathbb{N}$, so we use ω to denote their ordinal.

For each ordinal α there is exactly one ordinal $\alpha' \stackrel{\text{def}}{=} \alpha \cup \{\alpha\}$ that is the $\prec_o$-successor of α. (We also denote α' by $\alpha + 1$.) It follows that there is no $\prec_o$-largest ordinal.

This is where Burali-Forti entered. He proved that Cantor's set theory allows the construction of a set Ω of *all* the ordinals. He also showed that such an Ω leads to a paradox. Namely, Ω would not only be linearly ordered by $\prec_o$, but also well-ordered by $\prec_o$. As such, Ω would be associated with the corresponding ordinal, say α_Ω. But where would α_Ω be relative to Ω? Since Ω is the set of *all* the ordinals, it must be that $\alpha_\Omega \in \Omega$. On the other hand, α_Ω must be $\prec_o$-larger than any member of Ω, and therefore larger than itself.

2.2 Schools of Recovery

In this section we will describe the three main schools of mathematical thought that significantly contributed to the struggle against paradoxes in mathematics. These are *intuitionism*, *logicism*, and *formalism*. We will show how their discoveries synthesized in the concept of a formal axiomatic system and then in a clear awareness that a higher, metamathematical language is needed to investigate such systems.

2.2.1 Slowdown and Revision

The discovery of the paradoxical sets Ω, $\mathcal{U}$, and $\mathcal{R}$ was shocking, because Cantor's set theory was supposed to become a firm foundation for all other fields of mathematics and should, therefore, have been free of paradoxes. But the simplicity of Russell's Paradox, which used only two basic notions of set and membership relation, revealed that paradoxes originated deep in Cantor's theory, in the very definition of the concept of a set. It was this definition of a set and the unrestricted use of the *Axiom of Abstraction* that allowed the existence of the sets $\Omega, \mathcal{U}$, and $\mathcal{R}$ that, in the end, caused and revealed paradoxical situations. So it was clear that objects like $\Omega, \mathcal{U}$, and $\mathcal{R}$ should *not* be recognized as existing sets.

Therefore, Cantor's *naive* definition of the concept of a set (Sect. 2.1.2) should be restricted somehow. But this was easier said than done. Namely, Cantor's definition of a set was so natural and of such common sense that it was far from clear how to restrict it and, at the same time, retain all the sound parts of the theory. If a set is not what Cantor thought about, then what was it? And what was it not?

This once again triggered a critical reflection about the basic concepts, notions, principles, methods, and tools of set theory and logic, which might be sources of paradoxes. The aim was to make the necessary corrections to them, so that they, as a whole, would again act as a foundation for the development of mathematics and other axiomatic areas of science, but this time *safe from all paradoxes*. It turned out that no universally accepted resolutions could be expected. The critiques and proposals went in several directions, of which the three mainstream directions were called *intuitionism*, *logicism*, and *formalism*. Because they all contributed to future events, we briefly review them.

2.2.2 Intuitionism

Intuitionism argued for greater mathematical rigor in the process of proving and it advocated a non-Platonic view that the existence of a mathematical object is closely connected to the existence of its mental construction.

Fig. 2.4 Jan Brouwer
(Courtesy: See Preface)

The school was initiated by Brouwer[11] and then further developed by his student Heyting.[12] Brouwer was critical of the way in which Cantor's mathematics viewed the existence of *infinite* sets, and of the way in which mathematics was using the *Law of Excluded Middle.* He proposed a thorough change of this view as well as severe restrictions on the use of the law. Specifically, unlike Cantor, who considered infinite sets as *actualities*, i.e., accomplished objects, intuitionism advocated the classical point of view that infinite sets are no more than *potentialities*, i.e., objects that are *always* under construction, making it possible to construct as many members as needed, but *never all*. This view called for a change in the way that the *existence* of objects in infinite sets should be proven: an object is recognized as a member of an infinite set *if and only if* the object has been *constructed* or the existence of such a construction is beyond doubt. We give the details in Box 2.5.

Using these principles, intuitionism reconstructed several parts of classical mathematics and showed that such intuitionistic mathematics is free of all *known* paradoxes. Unfortunately, the price for this was rather high: large parts of mathematics had to be cast off, because it seemed impossible to reconstruct them according to intuitionistic principles. In addition, in the reconstructed mathematics, surprising changes occurred; for example, every (constructed) function is continuous.

It turned out that only a few researchers were willing to make such radical sacrifices.

NB *Nevertheless, the intuitionistic demand for mathematical rigor survived and was partially taken into account in the events to follow.*

Box 2.5 (Intuitionism).

This school argued for greater mathematical rigor in several ways.

View of Infinity. Since Aristotle, mathematics understood infinity only as the potentiality (i.e., possibility), never as the actuality (i.e., accomplishment). For instance, it is true that natural numbers $0,1,2,\ldots$ continue endlessly, yet up to any natural number there are only finitely many of them, and when we say that what remains is infinite we only mean that the rest, although growing ever larger, remains never accomplished. So, in the classical view infinity is by nature never accomplished, never actual. In contrast, Cantor's view of infinity was different, indeed radically Platonic: *"Any set, regardless of its size, is as much real as its members are real,"* he boldly advocated. To Cantor the set $\{0,1,2,\ldots\}$ was an actual, accomplished mathematical object.

Intuitionism returned to the classical view of infinity as potentiality. According to this view, using an appropriate procedure, we can find in an infinite set as many members as we wish, but never all of them. To treat infinite sets as actual, accomplished unities, is wrong, said intuitionists, and may lead to paradoxes (as shown by Russell and others).

But there are also differences between classical mathematics and intuitionism.

[11] Luitzen Egbertus Jan Brouwer, 1881–1966, Dutch mathematician and philosopher.

[12] Arend Heyting, 1898–1980, Dutch mathematician and logician.

Existence of Objects. Intuitionism treats the existence of mathematical objects differently from classical mathematics. In classical mathematics, mathematical objects exist *per se*, as Platonic ideas (see p. 10). Consequently, statements about mathematical objects are either true or false. Intuitionism does not accept this view. Instead, it advocates that the only things that exist *per se* are mental, mathematical constructions, while the existence of an object that has *not* been constructed remains *dubious*. For intuitionism, *to exist is the same as to be constructed*.

For instance, in classical mathematics, given a set $\mathcal{S}$ and a property P sensible for the members of $\mathcal{S}$, we are always allowed to *indirectly* prove that there exists a member of $\mathcal{S}$ with the property P, i.e., that the statement $\exists x \in \mathcal{S} : P(x)$ is true. To do this, we first make the hypothesis $H \equiv \neg\exists x \in \mathcal{S} : P(x)$, stating that such a member *does not* exist. Then we try to deduce from H a contradiction. If we succeed in this, we conclude that H is false. Now comes the critical step: since classical mathematics fully accepts the *Law of Excluded Middle*, there can be no other alternative but to conclude that $\neg H \equiv \exists x \in \mathcal{S} : P(x)$ is true, i.e., that such a member of $\mathcal{S}$ *exists*. But note that, generally, we have no idea about this member, or how to find it.

Intuitionism does not accept such an indirect proof of existence when the set $\mathcal{S}$ is *infinite*. Indeed, it rejects any proof of existence that neither constructs the alleged object, nor describes how to construct it at least in principle.

Use of Logic. The intuitionistic point of view was also reflected in the use of logic. For example, because of the *Law of Excluded Middle*, the classical mathematics takes for granted that, for *any* statement F, either F or $\neg F$ is true. Hence, the statement $F \vee \neg F$ is *a priori* true, even though we may never determine the truth-values of F and $\neg F$. Intuitionism, in contrast, treats the truth-values of F and $\neg F$ as *dubious*, until they are actually determined in some indisputable way.

To explain the reasons for such caution, let $\mathcal{S}$ be a set, P a property sensible for the members of $\mathcal{S}$, and F the statement $\forall x \in \mathcal{S} : P(x)$. So F conjectures that every member of $\mathcal{S}$ has the property P. How can we indisputably determine whether or not F is true? Can we always do this?

First, we can try to prove in one sweep that *all* the members of $\mathcal{S}$ have the property P. (We can use various techniques, such as mathematical induction.) If it turns out that we are unable to construct a proof that works for every member of $\mathcal{S}$, it might be that F is *false*. However, it might also be that F is *true*, where $P(x)$ holds for every $x \in \mathcal{S}$, but for a *different* reason in each case. That is, our inability to construct a one-sweep proof might be due to the lack of a recognizable pattern, i.e., a common reason for which different members of $\mathcal{S}$ share the property P. In this case, neither can we prove F (because the "proof" would be infinitely long) nor refute it (because F is true).

We must therefore resort to some other method to settle the conjecture F. If $\mathcal{S}$ is *finite*, we can in principle check, for each $x \in \mathcal{S}$ individually, whether or not $P(x)$ holds. When the checking is finished, we know either that F is true, or that it is false. But what if $\mathcal{S}$ is *infinite*? We can still do the checking, but we must be aware of the following. We may check as many members of $\mathcal{S}$ as we like, say 10^{18}, and find that each of them has the property P — but, generally, there is *no* way of knowing whether, for a member yet to be checked, P holds or not. So we keep checking in the hope that such a member will be reached soon. But, if in truth F is true, the checking will continue indefinitely, and we will never find out whether F is true or false. (By the way, this is the present situation with *Goldbach's Conjecture*; see Box 5.4 on p. 91.)

Finally, we may try to prove $F \equiv \forall x \in \mathcal{S} : P(x)$ by contradiction. As usually, we assume the converse, i.e., that $\neg\forall x \in \mathcal{S} : P(x)$ is true. In classical mathematics, where the equivalence $\neg\forall x \in \mathcal{S} : P(x) \Longleftrightarrow \exists x \in \mathcal{S} : \neg P(x)$ holds for arbitrary $\mathcal{S}$, we would try to deduce a contradiction from the more promising right-hand side of the equivalence. In intuitionism, however, the equivalence does not *a priori* hold; namely, if $\mathcal{S}$ is infinite, the statement $\exists x \in \mathcal{S} : \neg P(x)$ is *dubious* until we have *constructed* an $x \in \mathcal{S}$ for which $\neg P(x)$ holds. (As we have seen above, this may not be easy.) As long as $\exists x \in \mathcal{S} : \neg P(x)$ is dubious, it cannot be used to deduce a contradiction, and our proving by contradiction is stalled.

So in some situations the truth-value of a statement F cannot be indisputably determined.

2.2.3 Logicism

Logicism aimed to found mathematics on pure logic. As a side-effect it developed the notation by which mathematics was at last given concise and precise expression. The main contributions to this school were made by Boole, Frege, Peano, Russell, and Whitehead.

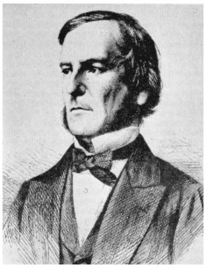

Fig. 2.5 George Boole (Courtesy: See Preface)

Fig. 2.6 Gottlob Frege (Courtesy: See Preface)

Fig. 2.7 Giuseppe Peano (Courtesy: See Preface)

Boole

In the middle of the nineteenth century scientists noticed that, from Aristotle onward, logical deduction had been using various *self-evident* rules of inference that, surprisingly, had never been rigorously analyzed and written down.

Boole[13] was among the first to become aware of the pitfalls of this. He embarked on the question of how to express logical statements by means of algebraic expressions (containing the operations "and," "or" and "not"), and then algebraically manipulate these expressions to pursue logical deduction. He described his discoveries in the book *The Laws of Thought* (1854) and thus founded *algebraic logic*. His logic was further developed by Peirce and others in the early twentieth century to become what is now known as *Propositional Calculus* **P** (see Appendix A, p. 297). Since then a more precise and clear expression of logical statements has been possible.

Frege and Peano

Frege[14] was aiming even higher. His goal was to show that arithmetic can be deduced from pure logic. In particular, he planned to define number-theoretic notions (i.e., numbers, relations, and operations on numbers) by pure logical notions, and to deduce arithmetical axioms from logical axioms.

[13] George Boole, 1815–1864, English mathematician and philosopher.

[14] Friedrich Ludwig Gottlob Frege, 1848–1925, German mathematician, logician, and philosopher.

Like Boole, Frege was well aware that a natural language, such as German, has structural, rhetorical, psychological, and other characteristics that often blur the meaning of its own statements and, consequently, the argumentation of the deductions. This required the introduction of a new, formal notation by which mathematics and logic could be given concise and precise expression. In particular, such a notation should be able to isolate all the important logical principles of inference while throwing off the lumber of natural language. In other words, the notation should be able to support purely logical deduction. So, in 1879, Frege proposed his *Begriffsschrift*, a "conceptual notation," capable of giving mathematics and logic better expression. Begriffsschrift was based on an alphabet of *symbols*, from which mathematical and logical expressions were constructed using *rules of construction*. An important innovation of Frege was that these rules directed the purely *mechanical manipulation* of symbols, without appealing to intuition or to the (possible) meaning of symbols. In addition, Frege introduced *quantified variables* and thus laid the foundations of the *First-Order Logic* (which we will describe later). The inferences were described diagrammatically, so they were in this respect somewhat unusual. Nevertheless, Begriffsschrift was capable of precisely and concisely representing the inferences that involved arbitrary mathematical statements.

At the same time, Peano[15] developed another symbolic language for expressing mathematical statements. He used innovative logical symbols (e.g., $\in, \Rightarrow$) in order to distinguish between logical and other operations. In 1895, he published a book *Fomulario Mathematico* where he expressed fundamental theorems of mathematics in his symbolic language. Peano's notation proved to be more practical than Frege's notation and is after having gone through further development in common use today.

In short, among Frege's and Peano's contributions to logic were the analysis of logical concepts, the foundation of the *First-Order Logic* **L** (see Appendix A, p. 298), and the introduction of a standard formal notation.

Russell and Whitehead

Russell's goal was even more ambitious than Frege's. He wanted to deduce *all* mathematics from logic. Namely, at the end of the nineteenth century it had already been shown that many concepts of algebra and analysis can be defined by means of number-theoretic notions, which, in turn, can be defined with purely logical notions.

To avoid his own paradox, Russell invented the *Theory of Types*. There are three requirements in this theory: 1) A *hierarchy of types* must be established. A *type* can be a member of any well-ordered set, e.g., a natural number. 2) Each mathematical object must be assigned to a type. 3) Each mathematical object must be constructed exclusively from objects of lower types in the hierarchy. As a result, the set $\mathcal{U}$ of all sets cannot exist in this theory (because $\mathcal{U} \in \mathcal{U}$), and neither does Russell's Paradox (for if $\mathcal{R}$ existed, we would have $\mathcal{R} \notin \mathcal{R}$ because of its type, and consequently $\mathcal{R} \in \mathcal{R}$ due to its definition: a contradiction). Similarly, Ω would not exist (as $\Omega \in \Omega$).

[15] Giuseppe Peano, 1858–1932, Italian mathematician.

These ideas were described in the 1910–13 book *Principia Mathematica* (*PM*) by Whitehead[16] and Russell. Using symbolic notation based on Peano's work, they developed from logical and three additional axioms the theory of sets and cardinal, ordinal, and real numbers, while avoiding all *known* paradoxes. The deductions were long, even cumbersome, yet many shared the opinion that the remaining fields of mathematics could also be deduced (at least in principle).

Fig. 2.8 Alfred Whitehead (Courtesy: See Preface)

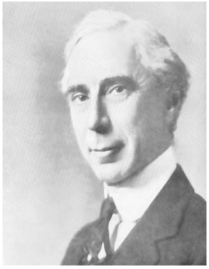

Fig. 2.9 Bertrand Russell (Courtesy: See Preface)

Did *Principia Mathematica* put an end to the crisis in mathematics? Not really. There were imperfections in *PM*. First of all, there was a kind of aesthetic flaw in the set of *PM*'s axioms, because in addition to logical axioms there were three axioms not recognized as purely logical. One of these was Cantor's *Axiom of Choice*. More importantly, it remained unclear as to whether *PM* is *consistent*, i.e., it avoids, besides all *known* paradoxes, also *all the other* paradoxes that may still be hidden in various fields of mathematics, patiently awaiting their discovery. This question became known as the *Consistency Problem* of *PM*. In addition, it was not clear whether *PM* was *complete*, i.e., whether exactly true statements are provable within *PM*. This was the *Completeness Problem* of *PM*. Consequently, *PM* was not widely accepted.[17]

NB *Nevertheless, PM was all important for future events, because it finally developed 1) a* symbolic language *for the concise and precise expression of mathematical statements from an arbitrary field of mathematics; and 2) a concise formulation of all the* rules of inference *used in the deduction of mathematical theorems. In addition, PM led to a clear formulation of the problems of* consistency *and* completeness *of a particular axiom system, the PM.*[18]

[16] Alfred North Whitehead, 1861–1947, British mathematician and philosopher.

[17] In addition, it would soon turn out that Russell's Paradox, as well as other paradoxes stemming from Cantor's liberal *Axiom of Abstraction*, can be eliminated just by a *two-level hierarchy of sets and classes*, instead of the complicated infinite hierarchy of types. (See Box 3.5 on p. 45.)

[18] As we will see in Chap. 4, the two problems were later solved in general by Gödel.

The concepts and tools developed by intuitionism and logicism were used by *formalism*, the third of the schools that attempted to resolve the crisis in mathematics.

2.2.4 Formalism

Formalism could not accept the radical measures suggested by intuitionism. It wished to keep all classical mathematics. After all, classical mathematics had been proving its immense usefulness from the very beginning.

To achieve this, formalism focused on a radically different aspect of human mathematical activity. Instead of being the *meaning* (i.e., semantics, contents) of mathematical expressions and inferences, the subject of the formalist's research was their *structure* (i.e., syntax, form). Formalism focused on the formal-language formulation of human mathematical activities and their results, as well as on the relations between these formulations.

This school was initiated by Hilbert[19] and then developed in close collaboration with Ackermann,[20] Bernays,[21] and others.

Fig. 2.10 David Hilbert (Courtesy: See Preface)

Syntax vs. Semantics

Hilbert became fully aware that it is sensible to draw a distinction between syntactic notions (i.e., notions referring to the structure of mathematical expressions) and semantic notions (i.e., notions referring to the meaning of mathematical expressions). For instance, the interpretation of a theory is a semantic notion. Recall that interpretation gives a meaning to a theory developed in a hypothetical axiomatic system (see p. 12). To describe the interpretation one needs to describe its domain, which is, mathematically, a set. But the concept of a set was not clear at that time. So Hilbert advocated a focus on syntactic notions, as the research of these seemed to require

[19] David Hilbert, 1862–1943, German mathematician.

[20] Wilhelm Friedrich Ackermann, 1896–1962, German mathematician.

[21] Paul Isaac Bernays, 1888–1977, Swiss mathematician.

only the non-problematic parts of mathematics, that is, basic logic and some basic arithmetic.

Let us describe these ideas in greater detail. Because it proved that mathematical concepts, such as that of the set, may be vague, also inference incorporating such concepts may be false, and may, eventually, lead to paradoxes. On the other hand, mathematical notions are always expressed in the *words* of some language, either natural, such as English, or symbolic, designed just for this purpose. A word is a finite sequence of *symbols* from some finite alphabet. Formalism noticed that every symbol is perfectly clear *per se*, that is, a symbol is comprehended as soon as it is recognized as a discrete part of the reality, without any further intuitive or logical analysis. This comprehension of symbols is independent of their intended meaning, which might previously be associated with them (such as the operation of addition with the symbol "+"). So why not comprehend words in that manner as well? One should only ignore the intended meaning of the word at hand and comprehend and treat it simply as a finite sequence of symbols. *Expressions*, i.e., sequences of words, could also be treated in the same fashion and, finally, the *sequences of expressions* too.

After the banishment of meaning from language constructs, one would be free to focus on their structure (syntax). But why do that? The reason is that one could found mathematical inference on a clear and precise structure (syntax) of language constructs, instead of on their (sometimes) unclear meaning (semantics). The syntax is always clear, provided it is rigorously and precisely defined (as was the case with logicism). As a result, a proof (deduction) would simply be a finite sequence of language constructs (expressions), built according to a finite number of rules. The gain would be improved control over the process of deduction and, finally, the elimination of paradoxes.

Formal Axiomatic Systems

In order to implement these ideas, formalism invented *formal axiomatic systems*. Each such system offers 1) a rigorously defined *symbolic language*; 2) a set of *rules of construction*, i.e., syntactic rules that are used to build well-formed expressions, called *formulas*, of the language; and 3) a set of *rules of inference* that are used to build well-formed sequences of formulas, called *derivations* or *formal proofs*. Each formula or derivation is viewed and treated exclusively as a finite sequence of symbols of the language. Hence, though each formula has a definite structure, no meaning is to be seen or searched for in it. Some of the formulas are proclaimed as *axioms*. Given a finite set of formulas, one may *infer* a new formula by applying a rule of inference. Formulas that can be derived by a finite sequence of inferences from axioms only are called *theorems*. Axioms, theorems, and other formulas make up the *theory* belonging to the formal axiomatic system at hand. A detailed discussion of formal axiomatic systems and their theories will appear in the next chapter.

Interpretation

Let us stress that formalists were aware of the fact that there was a limit to neglecting the semantics. After all, their ultimate goal was to establish conditions for the development of sound, safely *applicable* theories. They were aware that each theory, developed in a formal axiomatic system, should eventually be given some meaning; otherwise it would be of no use. In other words, the theory should be *interpreted*. Informally, an interpretation of a theory in a field of interest maps formulas of the theory into statements about (some) objects of the field. We will discuss interpretation again shortly.

NB *Formalism cast out the issues of meaning from the development of a theory, and shifted them to a later interpretation. What were the expected benefits of this? Such a theory could clearly show the syntactic properties of its expressions and expose various relations between these properties. Laid bare, the whole theory could be examined by* metamathematics *and subjected to its judgement.*

Metamathematics

When a theory is developed in a formal axiomatic system, the only things that can be examined within or about it are its *expressions*, the *syntactic properties* of expressions, and the *relations* between them. All these are unambiguously determined by the formal system (i.e., its language and rules of construction and inference). Thus, syntactic aspects of the theory can be systematically analyzed without the interference of semantic issues. Only now can one raise well-defined questions *about the theory* and propose answers to such questions.

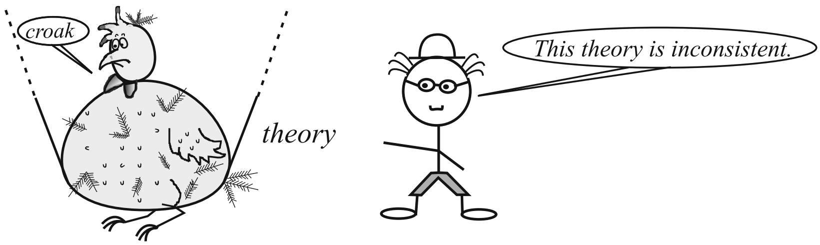

Fig. 2.11 A statement about the theory belongs to its metatheory

But questions and statements about the theory are no longer part of the theory. Instead, they belong to the higher "theory about the theory," which is called *metatheory*, or, more generally, *metamathematics*.[22] Thus, the subject matter of a metatheory is some other theory.

[22] meta- (Greek $\mu\varepsilon\tau\acute{\alpha}$) = after, beyond, about

Metamathematical statements are formulated in a natural language that can be augmented with additional symbols. (We will explain two such symbols, i.e., $\vdash$ and $\models$, shortly.) The proving of metamathematical statements is still necessary, but it is not formal, in contrast to proving within the formal system. Instead, the usual (i.e., semantic, informal) proving is used, where each inference in a proof must be grounded in the *meaning* of its premises. Of course, premises are metamathematical statements so that they can refer only to syntactic aspects of the theory. In addition, to avoid any doubts that might arise because of the use of infinity, only *finite* objects and techniques are allowed in metamathematical proofs. Such a cautious and indisputable way of reasoning is called *finitism*.

Goals of Formalism

Formalism harbored hopes that the analysis of formal systems would provide answers to many important metamathematical questions about the theories of interest. Specifically, these were the two well-known questions concerning mathematics developed in *Principia Mathematica* (see p. 25):

- The *Consistency Problem* of *PM* $\equiv$ "Is the math developed in *PM* consistent?"
- The *Completeness Problem* of *PM* $\equiv$ "Is the math developed in *PM* complete?"

But the ultimate goals of formalism were even more ambitious. Specifically, formalists intended to:

1. develop *all* mathematics in *one* formal axiomatic system;
2. *prove* that such mathematics is free of *all* known and unknown paradoxes.

2.3 Chapter Summary

The axiomatic method was used to develop mathematics since its beginnings. The evident axiomatic system required that basic notions and axioms be clearly confirmed by the reality. Since it was found that human experience and intuition may be misleading, the hypothetical axiomatic system was introduced. Here, axioms are only hypotheses whose fertility is more important than their link to reality. Such axiomatic systems offered more freedom in the search for interesting and useful theories. This approach was taken by Cantor when he developed his *Set Theory*. Because Cantor naively treated the existence of infinite sets, this resulted in several paradoxes in his theory.

Intuitionism, logicism, and formalism were three schools that reflected critically on the mathematical and logical notions and concepts that might be the cause of paradoxes.

Intuitionism advocated for greater rigor in the process of proving and for the non-Platonic view that the existence of mathematical objects is closely connected to the existence of their mental constructions. Intuitionism reconstructed several parts of classical mathematics that were free of all known paradoxes. But, at the same time, large parts of mathematics had to be cast off, as it seemed impossible to reconstruct them in the intuitionistic manner. Few researchers were willing to make such a sacrifice.

Logicism, the second school, developed a formal notation by which mathematics was given concise and precise expression. It also bore *Principia Mathematica*, a book that finally developed a symbolic language of mathematics and concisely formulated its rules of inference. In addition, it brought an awareness of the importance of the problems of consistency and completeness of axiomatic theories.

The third school, formalism, built on the ideas and tools developed by intuitionism and logicism, and aspired to retain all mathematics. Formalism acknowledged that the syntax and semantics of mathematical expressions should be clearly separated and dealt with in succession. It introduced the concept of the formal axiomatic system, i.e., an environment for the mechanical, syntax-oriented development of a theory. In addition, it introduced a clear distinction between a theory and its metatheory.

http://www.springer.com/978-3-662-44807-6

The Foundations of Computability Theory
Robič, B.
2015, XX, 331 p. 109 illus., Hardcover
ISBN: 978-3-662-44807-6