# FREGE, HILBERT, AND STRUCTURALISM

Mark Burke

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the Ph. D. in Philosophy

Department of Philosophy
Faculty of Arts
University of Ottawa

# Table of Contents

## Abstract

The central question of this thesis is: what is mathematics about? The answer arrived at by the thesis is an unsettling and unsatisfying one. By examining two of the most promising contemporary accounts of the nature of mathematics, I conclude that neither is as yet capable of giving us a conclusive answer to our question. The conclusion is arrived at by a combination of historical and conceptual analysis. It begins with the historical fact that, since the middle of the nineteenth century, mathematics has undergone a radical transformation. This transformation occurred in most branches of mathematics, but was perhaps most apparent in geometry. Earlier images of geometry understood it as the science of space. In the wake of the emergence of multiple distinct geometries and the realization that non-Euclidean geometries might lay claim to the description of physical space, the old picture of Euclidean geometry as the sole correct description of physical space was no longer tenable. The first chapter of the dissertation provides an historical account of some of the forces which led to the destabilization of the traditional picture of geometry. The second chapter examines the debate between Gottlob Frege and David Hilbert regarding the nature of geometry and axiomatics, ending with an argument suggesting that Hilbert's views are ultimately unsatisfying. The third chapter continues to probe the work of Frege and, again, finds his explanations of the nature of mathematics troublingly unsatisfying. The end result of the first three chapters is that the Frege-Hilbert debate leaves us with an impasse: the traditional understanding of mathematics cannot hold, but neither can the two most promising modern accounts. The fourth and final chapter of the thesis investigates mathematical structuralism—a more recent development in the philosophy of mathematics—in order to see whether it can move us beyond the impasse of the Frege-Hilbert debate. Ultimately, it is argued that the contemporary debate between 'assertoric' structuralists and 'algebraic' structuralists recapitulates a form of the Frege-Hilbert impasse. The ultimate claim of the thesis, then, is that neither of the two most promising contemporary accounts can offer us a satisfying philosophical answer to the question 'what is mathematics about?'.

## Acknowledgements

This thesis represents a radical (and initially daunting) departure from my earlier work in philosophy. During those innumerable dark nights of the soul when I began to question the sanity of both myself and the entire history of mathematics, I have always been fortunate to find myself surrounded by an incredibly supportive and inspiring group of people.

First and foremost I would like to thank my partner in life, Anna Candido, without whom this project would unquestionably have been abandoned long ago. I've relied perhaps too much on Anna's always kind words and on her willingness to entertain the ludicrous late-night questions of a philosopher at odds with himself and his subject. I cherish every day spent with her.

I would also like to thank my loving family—my parents, Danny and Carol, and my sister Kristy. Against all reason they have supported my interest in philosophy since the very beginning. I owe to them any strength of character and persistence required for this project—and a lot more besides!

Philosophically, this thesis has benefitted greatly from my conversations with colleagues and great friends at the University of King's College in Halifax. Though only rarely touching on questions of mathematics, the hours (days, months, years…) spent arguing about general philosophical matters with Darren Dyck, Matthew Furlong, Christopher Rice, and Saša Stanković have shaped my entire approach to philosophy, as well as some of the specific contents of this thesis. Often in these conversations, it has struck me that we are all working on the same problems from different angles. I only hope, in the future, to benefit them as much as they have me.

Outside of philosophy, my close friendships with George Blott, Peru Dyer, and Scott Harber have helped me remain (if I *have* remained) sane while going through the endless ringer of academia. The hours we've spent together producing and destroying art have been some of the finest of my life.

Finally, I would like to thank my supervisor, the always patient, always critical, always kind Paul Rusnock. My first encounter with Rusnock was in a course he taught on Gottlob Frege. Prior to this encounter, I had intended to write my doctoral dissertation in the area of ethical and political philosophy. Afterwards, I became deeply engrossed in the strange forms of rigor and precision found in logic and mathematics. Without his willingness to endure my questions (which invariably missed the mark) and to educate me (in metalogic, modal logic, set theory, group theory, even category theory) this project would never have come to fruition. Of course none of my own imprecisions, and none of the faults of this thesis, can be attributed to him.

# Introduction

What is mathematics about? This seems a rather straightforward question, but even our naïve first approximations seem to miss the mark. A common response to the question is that mathematics concerns number. But then we have geometry—a venerable old branch of mathematics which needn't concern itself with number at all. And, besides, what exactly are numbers, in any case? Since its earliest development, mathematics has been a richly varied science, including at different times geometry, arithmetic, algebra, mechanics, analysis, logic, and many other branches besides. And yet, despite this rich variety, most philosophers and mathematicians have tended to treat mathematics as a unified science, sharing, in some sense, a common content. Throughout much of its history, mathematics has been seen as a bastion of stability amongst the sciences, many of which change significantly in the wake of a new discovery or mode of thinking. But mathematics, too, has faced its share of instability. Throughout the eighteenth century, European mathematicians struggled to understand the exact meaning of the calculus, and devoted a great deal of effort to the taming of the infinitesimal. In the nineteenth century, the very nature of geometry was called into question by the emergence of non-Euclidean, projective, and other geometries. The nineteenth and twentieth centuries have witnessed perhaps the most fundamental crises so far in the shape of the set-theoretic paradoxes and Gödel's incompleteness results. Each of these events (among others) has altered our conception of mathematics as a science in significant ways, such that it is highly unlikely that a contemporary mathematician would answer the question 'what is mathematics about?' in the same way that Euclid or Archimedes might have.

My goal in this dissertation is not to canvass all possible views of the nature of mathematics, but to examine a major division amongst these views through the lens of a few privileged exemplars. The division is between what I call the ontologically-fixed conception of mathematics and the relational or free-floating or purely structural conception of mathematics.

In the first chapter, I investigate a large scale shift in understanding regarding the nature of geometry from the time of Euclid to the publication of David Hilbert's *Die Grundlagen der Geometrie* in 1899. My choice in this chapter of geometry, rather than a more obviously abstract mathematical field like algebra, is a deliberate one. The reason for this choice is that geometry, of all the branches of mathematics, has always seemed the most concrete and, therefore, the least likely to give way to an abstract structural approach. Unlike algebra, whose exact subject matter is and was a matter for debate, geometry has historically been understood to have an obvious and determinate subject matter (i.e., physical space). Thus, when an abstract structural view emerges even in geometry it lends credence to the claim that mathematics as a whole is best construed as the science or study of structures. (As we will see in the fourth chapter of the dissertation, where we examine Benacerraf's non-uniqueness problem, the other bastion of determinate content within mathematics—basic arithmetic—has come to be understood structurally as well). The direct focus of this first, geometric chapter is the relatively narrower problem of understanding shifting historical attitudes toward the parallel postulate. Various attempts to situate this postulate relative to the system of Euclid's *Elements* are considered until, in the 1830s, we begin to see a radical alteration in the conception of geometry as a science. Ultimately, I will argue that, in the wake of the development of mathematically consistent and physically applicable models of non-Euclidean geometries (among other factors), there was no longer an adequate philosophical account of the exact nature of geometrical science.

In the second chapter I turn to a debate that seems to promise such an account. On the one hand, we have the philosopher and logician Gottlob Frege, who argues that mathematics has a determinate content, that it is a science whose truths are absolute, and that mathematical terms have determinate referents. On the other side of the debate we have David Hilbert, one of the most influential mathematicians of modern times, and one of the major forces behind the modern axiomatic approach to the subject. For Hilbert, mathematics does not have a determinate content in the sense that Frege demands. Instead, it concerns uninterpreted formal systems, which can be *given* content through a procedure called ‘interpretation’, but which are free-standing when left alone. For Hilbert, questions of ontology and of absolute truth and reference miss the point of (modern) mathematics entirely. At the end of the chapter, we consider some rather significant criticisms of Hilbert’s view made by Frege. The result is that profound doubts have been cast on Hilbert’s answer to the question ‘what is mathematics about’.

The third chapter proceeds by investigating the chief alternative to Hilbert’s relational view: namely, Frege’s ontologically fixed view. Frege’s interesting logicist programme involved the reduction of basic arithmetic to pure logic. For him, an important part of this project concerned fixing the references of mathematical terms in his unique ontology of functions and objects. The chapter gives an account of the motivations for and contents of Frege’s ontological hierarchy and then moves on to discuss some serious problems with his account of the nature of mathematics. More specifically, the chapter examines two problems pertaining to the fixing of references: the concept *horse* problem, which seems to leave Frege unable to refer to one of the basic elements of his ontology, and the Julius Caesar problem, which threatens to undermine his capacity to fix the references of mathematical terms as basic as ‘number’. The conclusion of the chapter shows that, while Frege can escape the concept *horse* problem reasonably well, his program ultimately founders on the Julius Caesar problem. This leaves us at an impasse: neither Hilbert’s view nor Frege’s view can satisfactorily explain what mathematics is really about.

In the fourth and final chapter I continue my investigation of this question by examining some relatively recent attempts to answer (or otherwise dissolve) the Caesar problem faced by Frege. The focus of the first half of the chapter is on a 1965 paper by Paul Benacerraf which rejects an analogue of the Caesar problem insofar as it is based upon a fundamental misunderstanding of the nature of mathematics. The positive position Benacerraf briefly develops after this dissolution of the Caesar problem is strikingly reminiscent of Hilbert’s purely relational view. His claim is that mathematics is not about the kind of ontology that Frege was interested in but instead about *abstract structures*. This view has come to be called ‘structuralism’. In the end, Benacerraf’s particular form of structuralism is shown to be subject to deep difficulties and so we must turn to more recent work in the philosophy of mathematics to answer our question. Some of the most promising contemporary work follows in the wake of Benacerraf’s programme and can be called ‘structuralist’ in much the same way. In the contemporary setting, I shall argue, two main kinds of structuralism may be distinguished. On the one hand we have assertoric structuralism (exemplified in this chapter chiefly by the work of Stewart Shapiro). Assertoric structuralism contends, like Frege, that mathematics has a determinate content and that the claims of mathematics are meaningful assertions about this content. Unlike Frege, however, assertoric structuralists claim that this content is comprised of abstract structures. I go on to argue that Shapiro's brand of assertoric structuralism ends up enmeshed in the same sort of difficulties that troubled Frege's view.

On the other hand, we have algebraic structuralism (exemplified here by Steve Awodey and John Bell). Algebraic structuralism, like Hilbert's earlier position, rejects the idea that mathematics has a single determinate content in favour of a view which relativizes this content to a particular system, framework, or structure. Algebraic structuralism in its modern form is often developed using the tools of the branch of abstract algebra called category theory. On the algebraists' view, there are multiple different local categories within which mathematics can be done, each with its own ontology and logic. Thus, for them, Frege's dream of absolutely fixing the references of mathematical terms is just that—a dream, and a misleading one. Like Hilbert, however, the algebraic structuralists are forced to embrace relativized notions of truth, reference, and meaning. In the conclusion to the chapter, I argue that neither assertoric nor algebraic structuralism offer us a significant philosophical advance over the impasse of the Frege-Hilbert debate, and we are left with no way to adjudicate between competing, and deeply opposed, conceptions of the nature of mathematics.

In short, the thesis illustrates an historical and conceptual oscillation between two distinct pictures of mathematics. One in which mathematics is understood as a science having a unique content and clear referents, the other in which both content and reference vary depending upon one's local situation, system, or framework. We are left, then, with no clear answer to the question 'what is mathematics about,' but a clearer diagnosis of our dilemma.

# Chapter 1: The development of a structural view of geometry

## §0. Introduction

The point of this chapter is to explain why a structural approach to geometry emerged. This can appear to be a purely historical question, but I believe it is crucial for the philosophical project of this thesis as well. The project of the thesis, recall, is to investigate what mathematics is about. Since the $19^{th}$ century, there have been two broad approaches to this question. The first centers on discovering exactly which objects lie at the basis of all mathematical reasoning. This approach is much older and is found already in the work of Euclid. The second approach suggests that mathematics concerns structures and attempts to explain what structures are and how they relate to one another. This approach is new—beginning in the $19^{th}$ century—but is dominant in most areas of contemporary mathematics and the philosophy of mathematics. In this chapter, I will explain why this second view emerged at all.

It is impossible to evaluate and understand the various forms of mathematical structuralism if we have no idea what problems they were originally intended to address. The success of the structural approach to mathematics since the beginning of the twentieth century has made many of these initial problems difficult even to formulate: the status of Euclid's parallel postulate, for instance, does not strike us as a pressing problem which cuts to the very heart of geometry. It is now mostly an historical curiosity. The reason why we are not deeply troubled by the parallel postulate is because a new mode of thinking about geometry has emerged which eliminates both the problem of the parallel postulate and the basis for our deep interest in it. This new mode of thinking is not particularly concerned with the specifics of Euclidean geometry because it understands this geometry as just one among infinitely many. It is the relations between geometries and other sorts of 'structures' which occupy the new mode of thinking. This approach has been wildly successful. It has been so successful, in fact, that to speak meaningfully about modern mathematics seems to require us to accept the conceptual vocabulary of structuralism. We breathe the air of structuralism almost without notice.

But, as Frege notes already in his correspondence with Hilbert and elsewhere, the 'modern' approach to mathematics is by no means clear itself. I think it is very difficult to see that this approach might be fundamentally incoherent or unclear if we already ask and answer every question by employing its conceptual vocabulary. In this first chapter, then, I want to draw attention to the peculiarity of structuralism by examining the non-structuralist questions which prompted its emergence in the first place. It is difficult to avoid speaking even about the history of mathematics in structural terms. I think a careful and detailed attention to the problems which led to structuralism and the philosophical views which generated these problems is the best approach.

The chapter begins with an examination of the conceptual structure of Euclid's influential text, the *Elements*. After examining Euclid's work directly, we then turn to a series of criticisms and concerns pertaining to the parallels axiom. This historical account of the troubles engendered by this axiom is intended to illustrate the slow shift away from the object-centered account of geometry we find in Euclid towards the structure-focussed views more characteristic of contemporary mathematics. This historical portrait culminates in an examination of the unstable philosophical position of geometry just prior to the publication, in 1899, of Hilbert's *Grundlagen der Geometrie*. The chapter should show that, at the end of the 19$^{th}$ century, an entirely new way of doing mathematics had emerged, but its philosophical basis had not been clearly articulated. Chapter two will consider Hilbert's work, which, to many, brought a new philosophical and conceptual clarity for modern mathematics.

This historical matter is relevant not merely as background information, but as a necessary element for understanding the concerns discussed by Frege and Hilbert (to be examined in subsequent chapters). Without an appreciation of both the traditional picture of geometry and the instability created by its demise, we cannot grasp many of the important subtleties encountered in the Frege-Hilbert correspondence. In particular, without grasping changes in the common understanding of geometry throughout the nineteenth century, the motivations for and advancements embodied within Hilbert's work disappear entirely. Moreover, without understanding the crucial differences between Frege's position vis-à-vis axiomatization and the traditional Euclidean picture of geometry, we cannot fully appreciate how different even Frege's supposedly traditional picture of geometry is from that of his predecessors.

The ideas of both Frege and Hilbert have been very influential over the last century. The prominence of their respective views on logic and mathematics has made it easy for many commentators to read their works somewhat anachronistically. For instance, even astute readers of Hilbert tend to attribute a fully model-theoretic understanding of axiomatics to his earlier work.[1] But, as we shall see in the second chapter, during the earlier part of his career Hilbert had not yet developed clear enough logical and semantic views for this to be the case. By illustrating the links between Hilbert's work and that of his predecessors (which is partially accomplished in this chapter), such anachronism becomes easier to avoid. In Frege's case, too, commentators have tended to misconstrue his works based on anachronistic readings. When examining the Frege-Hilbert correspondence, for instance, many commentators have found Frege's criticisms to be pedantic or old-fashioned in a way that simply misses the point of 'modern' axiomatics.[2] Careful attention to prior developments in geometry and axiomatics, however, helps to illustrate, first, that Frege's views are not quite as conservative as they appear and, second, that his criticisms of the supposedly model-theoretic work of Hilbert are more sophisticated than has usually been assumed.

A more general benefit of this thoroughly historical way of proceeding is that it allows us to see that seemingly straightforward mathematical terms mean and have meant different things for different people, and that these differences in meaning have not always been noticed. In my initial attempts to address the problem of structure in mathematics I myself rather naively presupposed that there was some widespread agreement about what 'structure' might mean. As further historical enquiry has shown the meaning of common mathematical terms like 'axiom' and 'structure' is far from settled even today (though there is much less debate about the nature of axioms today, as the question seems to many to have been settled).

[1] See, for example, [Weyl 1970, 264-268].
[2] E.g., [Torretti 1984] and [Freudenthal 1962].

## §1. Euclid and the Parallel Postulate

It is chiefly from Euclid's *Elements* that the traditional understanding of geometry as the science of space and spatial objects derives. And it is chiefly from the geometries which differ from Euclid's own that we derive the cluster of views which hold that geometry is concerned primarily with spatial structures. At the outset of this chapter, I want to establish as clearly as possible the content of the traditional Euclidean picture of geometry. I am not interested in the purely mathematical aspects of Euclid's geometrical thinking, but in showing the philosophical principles and assumptions which underlie this thinking. As our historical analysis will show, many of these principles remained tacit, vague, and unclear for hundreds and even thousands of years. Indeed, one useful way of understanding the structure-focussed view of geometry which dominates the world of mathematics today is to see it as an attempt to avoid the incoherence and limitations of the Euclidean picture of geometry. In this chapter, I intend to flesh out what I take to be the central features of this incomplete picture so that we can better understand why alternative views emerged at all.

### 1.1 The origins and importance of Euclid's *Elements*

The thirteen books of Euclid's *Elements* are one of the marvels of ancient Greek mathematics, and, indeed, of rigorous thinking more generally. For nearly two thousand years, they remained a benchmark for certitude and provided an ideal to be lived up to by all other scientific treatises. Euclid's axiomatic method has also been emulated in a variety of fields, from ethics to economics, in addition to a pervasive and massive influence within almost every sub-discipline of mathematics, from geometry to analysis. It is indeed difficult to overestimate the importance this geometrical textbook has had for the development of scientific thought over the last two and a half millennia.

But, despite its singular position within the history of thought, Euclid's most famous work was not the first of its kind, only the most successful. We know from Proclus (410-485), and obliquely from Eudemus (c. 370 BC- c. 300 BC) and Simplicius (c. 490-560), that Hippocrates of Chios (fl. latter half of the 5$^{th}$ century BC) also wrote a treatise on geometry, entitled the *Elements*, about a century before Euclid's. We also know from Proclus that the two Platonists Leon and Theudius of Magnesia (c. 4$^{th}$ century BC) composed elements, though nothing from either of these works has survived and little else is known of their authors. Despite the disappearance of these competitors, it is clear both that Euclid's *Elements* was part of a tradition of Greek mathematical textbook writing and that it was considered by many of his contemporaries to be the pinnacle of that tradition. We can surmise its position at the head of the queue both from the number of favourable references in ancient sources and from the fact that its competitors have not survived. My interest here is not exclusively in the superiority and idiosyncrasy of Euclid's work. I focus on Euclid chiefly because his work has indisputably been the most influential and because his work has been thought by many to represent the central features of the classical Greek attitude toward mathematics. This Greek attitude, or, rather, the Greek methodological and philosophical approach to mathematics, has shaped the history and content of mathematics to such an extent that many of its assumptions remained invisible for centuries. By looking closely at Euclid and the problems stemming from his work, we will gain a clearer historical and philosophical understanding of the subsequent attack on the traditional picture of mathematics as a science concerned with certain types of objects.

So, let us turn now to the thing itself and discuss the aims, format, and content of Euclid's best-known work.

### 1.2 The aims of the *Elements*.

As noted above, the *Elements* was a member of a literary tradition which likely began in the fifth century B.C. with Hippocrates of Chios. Euclid's *Elements*, presumably like the lost members of the tradition, is a compendium of mathematical results—in this case those pertaining to plane geometry, proportions, number theory, incommensurable quantities, and the geometry of solids. As Kline notes, "[t]he *Elements* [...] is as much a mathematical history of the age just brought to a close as it is the logical development of a subject".[3] Accordingly, it is likely that very few of the theorems in the *Elements* are the original production of Euclid himself; rather, Euclid is responsible chiefly for the arrangement of the work and, probably, for a few of the proofs as well. While there is still debate as to the exact status of the *Elements* within the mathematical community of Euclid's time,[4] we can, with Proclus (to whom we owe a valuable commentary on Euclid), nevertheless discern two principal aims in the work. The first of these is the investigation and development of the theory of the "cosmic bodies"[5] (i.e., the five regular or Platonic solids) in the final three books—this aspect of the work likely represented the height of Greek geometrical knowledge of the time. The *Elements* presents a clear path to the development of this branch of geometry by way of a rigorous and lucid presentation of the methods and results required for its elaboration.

The second aim of the *Elements* is more general and pedagogical in character. Of this aim, Proclus writes:

> Of the purpose of the work with reference to the student we shall say that it is to lay before him an elementary exposition [...] and a method of perfecting his understanding for the whole of geometry. If we start from the elements, we shall be able to understand the other parts of this science; without the elements we cannot grasp its complexity, and the learning of the rest will be beyond us. The theorems that are simplest and most fundamental and nearest to first principles are assembled here in a suitable order, and the demonstrations of other propositions take them as the most clearly known and proceed from them.[6]

---

[3] [Kline 1972, 56].

[4] The debate concerns Euclid's intended audience; i.e., whether he was chiefly interested in communicating to already-sophisticated mathematical peers, or if the book was intended as a sort of catechism for prospective students of mathematics. Kline writes of the matter that "The purpose of Euclid's *Elements* is in question. It is considered by some as a treatise for learned mathematicians and by others as a text for students. Proclus gives some weight to the latter belief" [Kline 1972, 57]. The block quote from Proclus on this page is partial evidence for his belief in the pedagogical intent of the *Elements*, though the inclusion of relatively complex work on the platonic solids seems to indicate that Euclid may have been writing for already-knowledgeable mathematicians.

[5] [Proclus 1970, 70].

[6] [Proclus 1970, 71].

Thus, Proclus thinks, by studying the basic elements of geometry, the aspiring geometer is led by degrees toward ever more complex theorems and a correspondingly deepened understanding of the nature of the subject. Ultimately all of these resources are marshalled in the development of the theory of the platonic solids, combining both aims. Since the structure of the *Elements* is, in large part, the source of its unique influence, I turn now to a closer examination of that structure.

## 1.3 The format of the *Elements*

Book I of the *Elements* sets the stage for the austerity and strictness of the rest of the work. It begins abruptly with a list of twenty-three definitions. There is no preamble or explanation, and these definitions are followed by two more lists: the first contains five postulates, while the second contains five 'common notions' or axioms. After these lists Euclid proceeds to prove a number of propositions by employing the common notions, postulates, and acceptable modes of reasoning (which are never enumerated within the work itself). The following twelve books are all similar in character; occasionally, new definitions are introduced, though no new postulates or common notions appear. So, in sum, we find within the work five postulates, five common notions, a relatively small number of definitions, and a large number of theorems with accompanying proofs and figures. All of this is remarkably similar to a modern mathematics textbook—deceptively so.

### 1.3.1 Definitions.

Euclid's definitions can be divided into two distinct types. Those of the first type, exemplified by his definition of the circle,[7] are directly employed within proofs. These definitions define key geometrical concepts in terms of simpler ones. Thus, the circle is defined in terms of points, straight lines, planes, etc., and is appealed to in the proof of the first theorem of Book I. The second sort of definition has a less obvious role to play. These definitions (exemplified by Euclid's definition of 'point' and 'line'), are never directly employed to prove a proposition, nor are they referred to once they have been introduced. What, then, is their function? Clearly, one role is that of a propaedeutic—they introduce the reader to the most basic sorts of objects (points, lines, planes, etc.) about which Euclid will reason. While modern mathematical definitions tend to be of the abbreviatory sort, Euclid's second group of definitions might better be aligned with what Frege, for instance, called 'elucidations' [*Erläuterungen*].[8] These are pre-systematic remarks which help guide the reader toward a proper understanding of the basic elements to be treated within the system. Once these elements become clearly fixed, the function of the elucidations is achieved and they can simply be ignored. Thus, for instance, Euclid's definitions of a point as "that which has no part"[9] or a line as "breadthless length"[10] help the reader to

[7] "A circle is a plane figure contained by one line such that all the straight lines falling upon it from one point among those lying within the figure are equal to one another" [Euclid, I, Definition 15; Heath 1956, 183]. Euclid then defines the centre of a circle as the aforementioned point [Euclid, I, Definition 16; Heath 1956, 183].

[8] Cf., e.g., [Frege 1914, 207] and [Frege 1971, 6-8]. I discuss Frege's use of the term *Erläuterung* in more detail in the second and third chapters of the thesis. Earlier in the 19th century Bernard Bolzano (1781-1848) employed the term *Verständigung* [understanding] to much the same effect as Frege's *Erläuterung*; cf. [Bolzano 1975, §9] where he lays out the rules for the use of these sorts of explications.

[9] [Heath 1956, 155].

[10] [Heath 1956, 158].

understand precisely which objects he has in mind, though the definitions themselves play no direct role when he later proves propositions which contain the terms at issue. Ian Mueller, in a paper which highlights the divergences between modern axiomatic geometry and Euclid's approach, writes that

> The definitions should be looked at as attempts to make clear the meanings of the terms to be used before argumentation begins, that is, to make clear the nature of the objects to be studied. That the most fundamental definitions (e.g. of point, line, straight line) succeed only with persons who already have some idea what the objects in question are does not really matter if these definitions are taken to represent preliminary agreements among people of presumably normal intelligence.[11]

As Mueller suggests, Euclid's second type of definition should help to focus the reader's mind on the appropriate sorts of objects so that official argumentation can begin. The reader may already possess ideas of these objects, in which case the definitions serve the simple function of directing the mind to what it already knows, or of representing this knowledge. In the case where the reader does not possess appropriate ideas, the definitions cannot fix the reference of the terms in question; they serve, again, as helpful hints which can (at best) direct sufficiently capable readers to the sorts of things that Euclid has in mind. Beyond this preliminary role in establishing the particular sorts of objects about which Euclid will reason, definitions of the second type play no direct role within the system itself. Importantly, also, these definitions do not insist upon the existence of the objects which they define; the attribution of existence is reserved for the postulates.[12]

<u>1.3.2 Postulates and common notions (axioms)</u>.

Unlike some of Euclid's most basic definitions, the postulates and common notions *are* directly employed in order to prove propositions, and their special status relative to the rest of the system is of particular importance for understanding the nature of Euclid's work. Like the use made of definitions, the division of unproved assumptions into the two categories of common notions and postulates is not unique to Euclid. There was already by Euclid's time a robust methodological tradition, headed by Aristotle, which made such a division.

In general, common notions (Aristotle called them 'common items' or 'common opinions') are primitive, unprovable principles which do not contain reference to the specific notions of a given science. Assent to these common notions is not demanded because these notions act as constraints on reasoning in general, and, thus, their rejection would be tantamount to the rejection of deductive reasoning as such. Axioms, on the contrary, are claims for which assent is demanded and which are specific to the purposes of the particular science at issue. The idea is that we must accept certain facts as basic in order to get the deductive machinery going within a

[11] [Mueller 1969, 294].

[12] This is one of the central elements of the Aristotelian conception of definition. Aristotle writes, for instance, that "*what* the primitives and what the items proceeding from them mean is assumed; but *that* they are must be assumed for the principles and proved for the rest" [Aristotle 1994, I. 10, 76a 31-35, emphasis mine].

particular area. Aristotle writes of the difference between definitions, postulates, and common notions as follows:

> An immediate deductive principle I call a posit if it cannot be proved but need not be grasped by anyone who is to learn anything. If it must be grasped by anyone who is going to learn anything whatever, I call it an axiom (there are items of this kind); for it is of this sort of item in particular that we normally use this name. A posit which assumes either of the parts of a contradictory pair—what I mean is that something is or that something is not—I call a supposition. A posit which does not I call a definition. Definitions are posits (arithmeticians posit that a unit is what is quantitatively indivisible), but they are not suppositions (for what a unit is and that a unit is are not the same).[13]

Thus, common notions are unprovable principles which guide deductive reasoning in general, which needn't be explicitly accepted for that reason. Postulates are unprovable principles, assent to which is demanded for the purposes of a particular science. Definitions, finally, are postulates which make no existence claims. They inform us what the basic elements of the system are, but leave it open whether or not these elements exist.

These divisions are found in Euclid's work. There the common notions, for example, make no mention of the special geometrical objects defined, but concern only more general concepts like 'part,' 'whole,' and 'equality'. Euclid's[14] five common notions are as follows:

1. Things which are equal to the same thing are also equal to one another.
2. If equals be added to equals, the wholes are equal.
3. If equals be subtracted from equals, the remainders are equal.
4. Things which coincide with one another are equal to one another.
5. The whole is greater than the part.[15]

As we can see, no mention is made here of the objects defined in the first twenty-three definitions (e.g., points, lines, planes, etc.). These are general principles that Aristotle and Euclid would likely have assumed to be implicitly accepted as constraints upon deductive reasoning by anyone of normal intelligence. The postulates, by contrast, are specifically geometrical in character. They are as follows:

Let the following be postulated:

1. To draw a straight line from any point to any point
2. To produce a finite straight line continuously in a straight line
3. To describe a circle with any centre and distance
4. That all right angles are equal to one another

---

[13] [Aristotle 1994, I. 2 72a 15-24].

[14] There is some debate as to the historical origin of the common notions in Euclid, with the suggestion that some or all of them were added by later commentators. For relevant discussion, see [Heath 1956, 221-222].

[15] [Heath 1956, 155].

5. That if a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than the two right angles.[16]

In contrast to the common notions, the postulates make free use of the primitive terms of geometry. Their function is to insist that we accept some basic claims from which all other geometrical theorems will be derived. These claims are not proved in the *Elements*. Ideally, too, they should not be provable using the resources which Euclid makes available there; the postulates are basic claims (accepted as true), which provide the basis for Euclid's proof system. Thus, there should be a sharp separation between the theorems of the system (which are provable using only basic principles and previously proven theorems) and the axioms themselves. In addition to the importance of their unprovability, their *truth* is crucial as well. For, in contrast to some conceptions of mathematical truth in which truth is only defined relative to the consistency of a given axiomatic scheme (and, hence, undefined for axioms in isolation),[17] the Greeks demanded that the axiomatic basis of a science be independently true as well. As we shall see, the common view for the next two millennia was, precisely, that (most of) Euclid's postulates were not only understood to be true, but necessary and indubitable as well. Given the demand for rigor which characterized early Greek geometry, it is no surprise that even early readers of the *Elements* had their interest piqued by the fifth postulate, the so called parallel postulate, which is considerably more complex than the previous four. It is by examining the parallel postulate and the philosophical problems which stem from it that we start to see problems with the traditional Euclidean picture of geometry.

## 1.4 The parallel postulate

Despite widespread agreement that the fifth postulate, like the others, was a truth, there was early on a serious debate as to its status as a postulate within the system. Despite its status as a postulate, many felt that the parallel postulate should be provable from the other four postulates, a belief which gave rise to a remarkably long and complex history, beginning as early as Ptolemy and continuing well into the 19$^{th}$ century. Even in 1767 D'Alembert called the status of the parallel postulate the "the scandal of geometry".[18] And as late as 1813 Gauss—one of the men chiefly responsible for overturning the view of Euclidean geometry as the science of space—remarked that "In the theory of parallels we are still no further along than Euclid," and that the theory of parallels was a rather "shameful part of mathematics".[19]

---

[16] [Heath 1956, 154-155].

[17] This is the case for, e.g., Hilbert's influential conception of mathematical truth. Hilbert writes in response to Frege's somewhat more traditional view of mathematical truth, that "If the arbitrarily postulated axioms [of geometry] together with all their consequences do not contradict one another, then they are true and the things defined by these axioms exist. For me, this is the criterion of truth and existence" [Hilbert, in Frege 1971, 12]. I discuss the differences between these two conceptions of mathematical truth in more detail in the second chapter.

[18] "La définition et les proprietés de la ligne droite, ainsi que des lignes parallèles, *sont donc l'écueil*, et pour ainsi dire *le scandale des élémens de géométrie*" [d'Alembert 1759-1767; reprint 1805, volume 2, 331].

[19] "In der Theorie der Parallellinien sind wir jetzt noch nicht weiter als Euklid war. Diess ist die *partie honteuse* der Mathematik, die früh oder spät eine ganz andere Gestalt bekommen muss" [Gauss, dated 27 April 1813; in Gauss 1863-1929, Band VIII, 166].

Though there was little argument over the status of the first four postulates, the provability of the fifth was, from Euclid's time onward, a matter of some contention. As Heath notes, "[f]rom the very beginning, as we know from Proclus, the Postulate was attacked as such, and attempts were made to prove it as a theorem or to get rid of it by adopting some other definition of parallels".[20] It is worth quoting the relevant passage from Proclus at some length to illustrate the troubling character of the postulate as it was understood relatively early in its history. Proclus writes that the fifth postulate

> [...] ought to be struck from the postulates altogether. For it is a theorem—one that invites many questions, which Ptolemy proposed to resolve in one of his books—and requires for its demonstration a number of definitions as well as theorems. And the converse of it is proved by Euclid himself as a theorem. But perhaps some persons might mistakenly think that this proposition deserves to be ranked among the postulates on the ground that the angles' being less than two right angles makes us at once believe in the convergence and intersection of the straight lines. To them Geminus has given the proper answer when he said that we have learned from the very founders of this science not to pay attention to plausible imaginings in determining what propositions are to be accepted in geometry. Aristotle likewise says that to accept probable reasoning from a geometer is like demanding proofs from a rhetorician. And Simmias is made by Plato to say, "I am aware that those who make proofs out of probabilities are impostors." So here, although the statement that the straight lines converge when the right angles are diminished is true and necessary, yet the conclusion that because they converge more as they are extended farther they will meet at some time is plausible, but not necessary, in the absence of an argument proving that this is true of straight lines. That there are lines that approach each other indefinitely but never meet seems implausible and paradoxical, yet it is nevertheless true and has been ascertained for other species of lines. May not this, then, be possible for straight lines as for those other lines? Until we have firmly demonstrated that they meet, what is said about other lines strips our imagination of its plausibility. And although the arguments against the intersection of these lines may contain much that surprises us, should we not all the more refuse to admit into our tradition this unreasoned appeal to probability? [...] These considerations make it clear that we should seek a proof of the theorem that lies before us and that it lacks the special character of a postulate. But how it is to be proved, and with what arguments the objections to this proposition may be met, we can only say when the author of the *Elements* is at the point of mentioning it and using it as obvious. At that time it will be necessary to show that its obvious character does not appear independently of demonstration but is turned by proof into a matter of knowledge.[21]

For Proclus, then, there are a few concerns weighing against the unprovability of the fifth postulate. Primarily, he claims that the postulate is, rather, a theorem—a theorem which Ptolemy

[20] [Heath 1956, 202].
[21] [Proclus 1970, 191-193]; in the final sentences here, Proclus alludes to what would be a common tactic used in attempts at proving the fifth postulate, namely using only the first four postulates and the first twenty-eight theorems of Euclid which occur before the first use of the parallel postulate is required.

had already attempted to prove. That it is a theorem seems to be made evident by the fact that Euclid proves a proposition which is equivalent to the postulate's converse, namely proposition 17 of Book I of the *Elements*. This proposition states that "In any triangle two angles taken together in any manner are less than two right angles".[22] Euclid proves this proposition by employing the second postulate and the previously proved proposition 13 of Book I. Accordingly, Proclus finds it suspicious that a principle whose converse is provable using only the resources of the first four postulates should not itself be provable from those resources. Nevertheless, Euclid's proof of proposition 17 of Book I rests upon the unstated further assumption that two straight lines meet, if they meet at all, in at most one point—an assumption quite as strong as the parallel postulate itself.[23] Proclus takes it to be the case that the existence of a proof of the postulate's converse speaks in favour of the provability of the postulate itself. This seems a reasonable enough assumption; nevertheless, when we take into consideration the fact that Euclid's proof requires the assumption of a principle which is equal in complexity and obscurity to the parallel postulate, it is no longer obvious that the existence of this proof can act as evidence either for or against the provability of the parallel postulate itself.

In addition to the supposed provability of its converse, Proclus also finds fault with the parallel postulate insofar as it seems to have been given the status of a postulate for rather weak epistemological reasons. He rejects the view that our immediate assent to a geometrical claim can justify us in assigning that claim the status of a postulate. For a proposition to be considered a postulate, our immediate assent, based upon its undeniable plausibility, is not enough. Beyond the mere fact of the layman's acceptance of the proposition as obvious, it must also be deemed necessary in the sense that its opposite cannot be the case. But, as Proclus rightly notes, there are species of lines other than straight lines (e.g., hyperbolas and their asymptotes) which do not obey the dictates of the parallel postulate; accordingly, our intuitive acceptance of the postulate seems to be ill-founded or, at the very least, in need of supplementation. Thus, Proclus thinks, we should expect a proof that straight lines, unlike hyperbolas and their asymptotes, do in fact obey the parallel postulate.

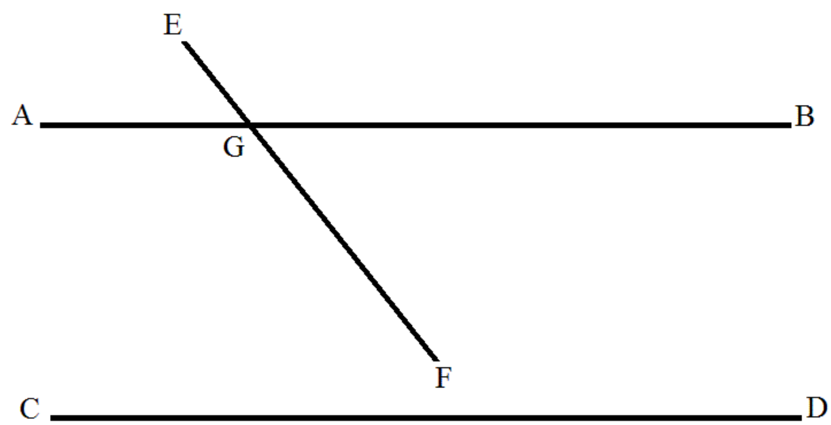
For these reasons, then, we find attempts as early as Ptolemy to prove the fifth postulate by means of the other four. In what follows, I would like to examine a few of these attempts to prove, reformulate, or resituate the parallel postulate relative to the rest of Euclid's system. My chief aim here is to illustrate that the unstable position occupied by the parallel postulate slowly led to the rise of a new understanding of the science of geometry in which 'structure' became increasingly prominent. Though I mention other figures, I will focus chiefly on Ptolemy, Proclus, Khayyam, and Saccheri, as elements of their work were central to the later emergence of non-Euclidean geometry and its destructive effect on the traditional conception of geometry as the science of space or spatial objects.

[22] [Heath 1956, 281].
[23] This assumption is provably false within elliptic geometry, for instance, unless the further condition that antipodal points are taken to be equivalent is added.

## 1.4.1 Ptolemy and Proclus

As noted above, we know of Ptolemy's attempted proof of the fifth postulate by way of Proclus. Ptolemy's assault on the parallel postulate was likely the first which attempted to prove the troublesome claim by assuming an equivalent principle. (This would turn out to be a popular tactic over the centuries). In Ptolemy's case this assumption was that "through any point only one parallel can be drawn to a given straight line".[24]

**Figure 1: Proclus and parallels**




Proclus' own proof, too, rests on a disguised equivalent of the very postulate to be proved. In Proclus' case, this is an axiom taken from Aristotle.[25] This axiom is stated as follows: "If, from a single point two straight lines making an angle are produced indefinitely, the interval between them when produced indefinitely will exceed any finite magnitude".[26] From this principle he deduces the proposition that a straight line which meets one of a pair of parallels will also meet the other. He proves this proposition as follows. Take two parallels, *AB* and *CD*, and another straight line *EF* which intersects *AB* at a point *G*. The distance between a point moving along the line segment *GF* and the line *CD* will decrease indefinitely when the distance of the moving

[24] [Heath 1956, 206].

[25] The axiom is taken from Aristotle's well-known arguments for the finitude of the universe in *De Caelo* (cf. [Aristotle 1971, I. 5, 271b 28ff.]). Guthrie's translation has: "The following arguments make it plain that every body which revolves in a circle must be finite. If the revolving body be infinite, the straight lines radiating from the centre [of the circular path of the body] will be infinite. But if they are infinite, the intervening space must be infinite. "Intervening space" I am defining as space beyond lines. This must be infinite. In the case of finite lines it is always finite, and moreover it is always possible to take more than any given quantity of it, so that this space is infinite in the sense in which we say that number is infinite, because there exists no greatest number. If then it is impossible to traverse an infinite space, and in an infinite body the space between the radii is infinite, the body cannot move in a circle. But we ourselves see the heaven revolving in a circle, and also we established by argument that circular motion is the motion of a real body" [Aristotle 1971, I.5., 271$^{b}$ 28-272$^{a}$ 7].

[26] [Proclus 1970, 371.10].

point from *G* is increased indefinitely. But, since the distance between the two parallel lines is finite, the straight line *EF* must meet *CD* (see Figure 1).

But this reasoning, too, relies upon a rather strong assumption; namely, the assumption that the distance between any two parallel lines always remains finite. And, again, this assumption is provably equivalent to the parallel postulate itself. This argument would later be criticized by Christopher Clavius (1538-1612)—whom Saccheri read and referred to[27]—on the fairly reasonable grounds that the Aristotelian axiom itself requires proof. As Heath notes, Clavius

> points out that, just as you cannot assume that two lines which continually approach one another will meet (witness the hyperbola and its asymptote), so you cannot assume that two lines which continually diverge will ultimately be so far apart that a perpendicular from a point on one let fall on the other will be greater than any assigned distance; and he refers to the *conchoid* of Nicomedes which continually approaches its asymptote, and therefore continually gets farther away from the tangent at the vertex; yet the perpendicular from any point on the curve to that tangent will always be less than the distance between the tangent and the asymptote. Saccheri supports the objection.[28]

Subsequent attempts at proof often rested on the assumption of similarly disguised equivalents of the parallel postulate. Occasionally mathematicians would explicitly offer reformulations of the postulate in seemingly simpler terms. The English mathematician, John Wallis (1616-1703), for instance, knowingly assumed that transformations which alter only the size of a figure do not affect its angles or the proportions of its sides.[29] Somewhat surprisingly, this assumption, too, is equivalent to Euclid's parallel postulate—an equivalence made apparent in more recent times by the impossibility of constructing similar triangles of different size within some of the more familiar non-Euclidean geometries.

These attempts to prove the postulate invariably end up reintroducing an equally troublesome (or equivalent) claim in its place. There are innumerable examples of this problem, a few of which are mentioned above. There was a general feeling that the parallel postulate was somehow unsatisfying, but there was no real consensus as to what feature of the postulate was at issue. As yet, there was no significant attempt to alter the Euclidean picture of geometry. Postulate-tinkerers were not revolutionaries but refiners: they sought to perfect Euclid rather than overturn him. But, over time, more and more difficulties with the Euclidean picture began to pile up.

---

[27] Saccheri discusses Clavius at [Saccheri 1733, 83ff.]; incidentally, Saccheri corrects an error in Clavius' attempt to employ the bisection of a circle by its diameter as a means of proving the parallel postulate—unbeknownst to Saccheri, he corrects Clavius in precisely the same place and manner (see [Saccheri 1733, 91]) that Leibniz had a half century earlier. On this, see [De Risi 2007, 250-252, note 109].

[28] [Heath 1956, 208].

[29] To be more precise, Wallis assumed that for any triangle *ABC* there exists a triangle *DEF* having the same angles and proportions (i.e., a triangle which is similar to *ABC*), but differing in size, or, that there exist two similar but unequal triangles. (See [Wallis 1693]). Wallis was influenced in his work on parallels by the 13th century Arab geometer, Nasir al-Din al-Tusi (1201-1274), so much so that he commissioned a translation of his work, which exploited basic facts about triangles to 'prove' the fifth postulate. [Wallis 1693] contains Wallis' own proof as well as a rendition of al-Tusi's results; cf. also [Bonola 1955, 10-12] for a brief resumé of al-Tusi's attempted proof.

1.4.2 Omar Khayyam

A rather important addition to the picture begins to emerge with the work of the Persian polymath Omar Khayyam (1050-1123). The *Commentary on the Difficulties of Certain Postulates of Euclid's Work*, Khayyam's monograph on the parallel postulate, is interestingly different than earlier work insofar as its aim was not to prove the postulate directly, but rather to derive it explicitly from an equivalent claim. In other words, where Ptolemy and Proclus, for instance, had unknowingly employed equivalent axioms to prove Euclid's, Khayyam knowingly attempted to replace the rather complex version we find in Euclid with his own, hopefully simpler and intuitively more acceptable principle. Instead of proving that the postulate was straightforwardly a theorem in Euclid's system, Khayyam sought to add a simpler version of the postulate to that system, thereby eliminating questions about the complexity and strangeness of the parallel postulate. He would then, ideally, show that Euclid's complex postulate could be derived from his simpler, more acceptable reformulation. His reformulation was as follows:

> Two convergent straight lines intersect and it is impossible for two convergent straight lines to diverge in the direction in which they converge.[30]

He also noted something which Saccheri was later to exploit to great success: namely that, if we choose to leave Euclid's fifth postulate out of the *Elements* entirely, we are presented with three possible situations. Recall that Euclid's formulation of the parallel postulate runs as follows:

> That if a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than the two right angles.[31]

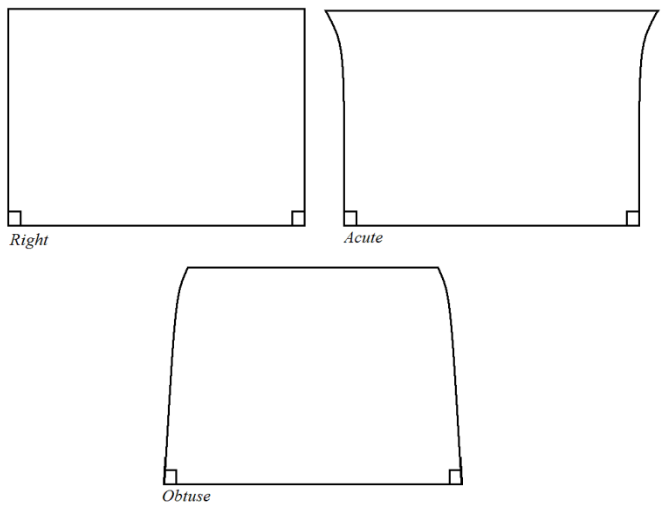
So, in the properly Euclidean case, if two straight lines, both perpendicular to a third, cross another, fourth, line, then the sum of the internal angles where the two perpendiculars meet the fourth line will be 180º. Khayyam noted two other possibilities made available if we do not accept Euclid's postulate (both of which Saccheri explored in greater detail). He noted, namely, the case in which the sum of the internal angles is greater than 180º (the obtuse case) and the case where the sum of the internal angles is less than 180º (the acute case). Indeed, it is precisely these replacements for Euclid's parallel postulate which would later provide the basis for hyperbolic geometry (the acute case) and elliptic geometry (the obtuse case). The three cases can all be represented visually (see Figure 2) by so-called Saccheri quadrilaterals (sometimes called Khayyam-Saccheri quadrilaterals), i.e., quadrilaterals which possess at least two right angles at their base and whose opposing sides are equal.

Khayyam would end up dismissing both non-Euclidean cases because he believed they led to contradictions, but his reasoning employed certain unstated assumptions provably equivalent to the parallel postulate itself. Despite his own misgivings, Khayyam's work on multiple geometric systems heralded an entirely new approach to the problem of the parallel postulate. Over time the misgivings disappeared and multiple geometries came to replace the view of Euclid's geometry

[30] [Khayyam, quoted in Rosenfeld and Youschkevitch 1996, 130].
[31] [Heath 1956, 155].

as the only one possible. But it would take a tremendous amount of time before this could happen. Why? Why were Khayyam's successors, too, reticent to embrace these other cases as consistent alternatives to Euclid's geometry? Why didn't something like the structural approach to geometry emerge in 11th century Persia?

**Figure 2: Khayyam-Saccheri quadrilaterals**




1.4.3 Girolamo Saccheri

The best place to look for an answer to this question is the Italian Jesuit Girolamo Saccheri, who seems to have possessed detailed knowledge of elements of (both) hyperbolic and elliptic geometry. And yet, unlike later geometers like Gauss, Bolyai, and Lobachevsky, he fell short of discovering these non-Euclidean geometries. An explanation of this near-success is illustrative of the small steps by which the movement away from the traditional Euclidean picture of geometry occurred. The details here are important precisely because they help to illustrate the deep hold of the Euclidean picture even in the 18th century.

In his 1733 work on Euclid (*Euclides ab omni naevo vindicatus* [Euclid Vindicated of Every Flaw]) Saccheri set out to prove the fifth postulate. His approach was similar to that of Khayyam in that it attempted, by *reductio*, to illustrate the impossibility of the two alternatives to the parallel postulate—namely the acute and obtuse cases isolated by his predecessor. His arguments against the obtuse case are rather more convincing than those against the acute. In the first instance he shows that the 'elliptical' version of the parallel postulate leads to a contradiction, namely the conclusion that all straight lines are finite (a claim which, on one reading,[32] directly

[32] As Stäckel and Engel note in their edition of Halsted's translation of Saccheri, there is insufficient precision in Saccheri's own text to decide between a few possible readings. They write, The Proposition on Exterior Angles (Eu. I. 16) that is being used here presupposes that the straight line is of *infinite length*. An alternative assumption, suggested by Riemann […] is that the straight line is *unbounded*. This property is compatible with the hypothesis that the straight line is infinite (open) as well as with the hypothesis that it is finite (closed). If the first of these two

contradicts Euclid's second postulate).[33] In the acute case, however, such a contradiction is not so easy to achieve. Saccheri's attempt to destroy the acute hypothesis (and, thus, hyperbolic geometry as a real alternative to Euclid's) is instructive. It shows the deep, and tacit, commitment to the view of geometry as the science of space and the ways in which this commitment acted as a conceptual barrier to the development of a relational or structural approach to geometry. Let us turn, briefly, to his conclusions, ignoring for the moment some of the particulars of his reasoning. Though very few geometers wrote philosophical treatises outlining their view of the nature of geometry, in arguments like Saccheri's we can tease out some of the underlying assumptions. After a long and involved chain of theorems derived from the assumption that the acute hypothesis is true,[34] Saccheri eventually arrives at the following proposition:

> Proposition XXXIII
>
> *The hypothesis of acute angle is absolutely false; because repugnant to the nature of the straight line.*
>
> PROOF. From the foregoing theorem may be established, that at length the hypothesis of acute angle inimical to the Euclidean geometry has as outcome that we must recognize two straights AX, BX, existing in the same plane, which produced *in infinitum* toward the parts of the points X must run together at length into one and the same straight line, truly receiving, at one and the same infinitely distant point a common perpendicular in the same plane with them. But since I am here to go into the very first principles, I shall diligently take care, that I omit nothing objected almost too scrupulously, which indeed I recognize to be opportune to the most exact demonstration.[35]

Saccheri here simply abdicates any further investigation of the acute-angled hypothesis. Why? Because, as he may simply have assumed from the start, this hypothesis is repugnant to the Euclidean conception of the straight line. To the modern eye, this is a rather surprising surrender after several pages of technical and sophisticated argumentation. Indeed, Bonola characterizes this move as an attempt "to come to a decision," in which we witness Saccheri "trusting to intuition and to faith in the validity of the Fifth Postulate rather than to logic".[36] George Halsted, one of Saccheri's English translators and a great populariser of early works of non-Euclidean

---

hypotheses is taken, then, as Saccheri shows, the hypothesis of the obtuse angle leads to a contradiction. If the straight line is, on the other hand, taken to be (merely) *unbounded*, which allows the second possibility – that the straight line is finite (closed) – then the Proposition on Exterior Angles (Eu. I. 16) cannot be invoked, and Saccheri's proof becomes insufficient. Indeed, if this second possibility is assumed to be a closed line of finite length, then a contradiction can be shown *not* to occur" [Saccheri 1986, 246 (editors' notes)].

[33] Accordingly, modern elliptic geometry rejects both the second and the fifth of Euclid's postulates. (The third postulate also fails to hold, unless it is slightly modified such that one can construct a circle of radius equal to any given line segment rather than the less restricted claim we find in Euclid that one can construct a circle of *any* radius).

[34] For a reasonably brief synopsis of the steps involved in this part of Saccheri's reasoning, [Bonola 1955, 38-44] is helpful.

[35] [Saccheri 1733, 173].

[36] [Bonola 1955, 43].

geometry in general, makes the following remarks on this somewhat hasty rejection of the acute angle hypothesis:

> Remember De Morgan's saying: "As to writing another work on geometry, the middle ages would as soon have thought of composing another New Testament...... [Euclid's] order of demonstration was thought to be necessary and founded in the nature of our minds;" and remember that Saccheri's book contains not merely "another *work* on geometry" but *another geometry*, a thought so tremendous, so unorthodox, that its discovery in his book by these great Church Dignitaries would have doomed Saccheri to death.[37]

Despite the rather hyperbolic tone, there is some substance to Halsted's remarks here. The notion that Euclid's venerable geometry might be wrong—or, worse, that it might simply be one geometry among many—was almost unthinkable in the early eighteenth century.[38] Bonola characterizes Saccheri's rejection here as a matter of faith in the time-tested Euclidean view of geometry. I think it is perhaps more accurate to say that Saccheri's enquiry was undertaken with the fairly reasonable assumption that geometry was the science of space and that there was no room in such a science for two valid but mutually exclusive claims. Nevertheless, Halsted seems to harbour the suspicion that Saccheri, against the spirit and limitations of his time, may very well have been a closet non-Euclidean. Pursuing this theme, he continues:

> There exists in the library of Modena, Italy, a Manuscript Biography of Saccheri. Dr. Emory McClintock, the able President of the New York Mathematical Society, writes of Saccheri: "He confessed to a distracting heretical tendency on his part in favour of the 'hypothesis anguli acuti,' a tendency against which, however, he kept up a perpetual struggle. After yielding so far as to work out an accurate theory anticipating Lobatschewsky's doctrine of the parallel-angle, he appears to have conquered the internal enemy abruptly, since, to the surprise of his commentator Beltrami, he proceeded to announce dogmatically that the specious 'hypothesis anguli acuti' is positively false." Of course, no such confession occurs or could have occurred in the book itself; for with it the book could never have been printed. If such a statement occurs in the Manuscript Biography, that must have been the work of a personal friend written after Saccheri's death. [...] The sudden dogmatic assertion above mentioned does occur, first on page 70 of the work, a quarto. But this, after seventy quarto pages of rigorous logic and elegant demonstration to establish a non-Euclidean geometry, may be looked upon as something like the stucco for the king's inspection with which the immortal architect in Egypt covered the stone bearing his own name.[39]

We may wish, with Halsted and McClintock, to read into Saccheri a more tolerant conception of the nature of geometry than is present in his work. However, such tolerance was not merely a

---

[37] [Halsted 1894, 150].

[38] Even early forms of (what would later be called) projective geometry, or Lambert's startlingly abstract 'spherical geometry', were not considered true *alternatives* to traditional Euclidean geometry, but special cases of it. In the case of projective geometry, its potentially 'extra-Euclidean' quality only became exploited in the works of Arthur Cayley in the mid-19th century.

[39] [Halsted 1894, 150-151].

matter of submission to Euclid as an unquestioned authority. Saccheri, like his contemporaries, worked from within a tradition in which the indubitable, intuitive foundation of Euclidean geometry was taken as a matter of course. Indeed, the very title and intent of the work—to vindicate Euclid of all possible blemishes—is illustrative of Saccheri's acceptance of the traditional view of Euclidean geometry as *the* correct science of physical space. Without replacing this belief in the certainty of Euclidean geometry as a description of physical space, it is highly unlikely that anyone, even a geometer as astute as Saccheri, would have been capable of accepting hyperbolic geometry as anything more than a confusingly resistant collection of falsehoods, or perhaps an amusing formal game. All of this, combined with Saccheri's text itself, speaks against any hidden affinity for hyperbolic geometry as a real alternative to Euclidean geometry. And so it is that, after seventy quarto pages, we find an exasperated Saccheri throwing up his hands at the futility of pursuing the acute angle hypothesis any further.

Nevertheless, Halsted is correct when he finds the end of this lengthy and elegant chain of reasoning somewhat abrupt: Saccheri simply refuses to accept the possibility that a geometry based on the acute angle hypothesis (what we would call 'hyperbolic geometry') can be taken seriously, and he rejects the acute angle hypothesis out of hand. All of this seems rather puzzling from a modern point of view in which the notion of multiple geometries with numerous different applications is a commonplace. But we must consider the traditional view of geometry as the science of physical space, a science whose absolute certitude was thought to have been unassailable because grounded in our direct experience and intuition of that space. To help explain Saccheri's rather abrupt rejection of a potential alternative to Euclidean geometry, I shall turn in the following section to a more focussed discussion of the philosophical and conceptual background which framed the tacit basis of Euclidean geometry as late as the mid-nineteenth century. It is only when the assumptions underlying this picture became visible that they were called into question. And it is only when the philosophical assumptions of Euclidean geometry were called into question that the structural approach to geometry began to emerge as a philosophical and mathematical alternative.

## §2. Geometric intuition and the emergence of non-Euclidean geometry

Now, as we've just seen, Saccheri's attempts to prove the parallel postulate brought him extraordinarily close to the discovery of hyperbolic geometry; a discovery which would, a century later, make Lobachevsky and Bolyai deservedly famous. Given his proximity to revolutionary geometric results, it is worth pausing to discuss the source of Saccheri's 'failure', though we may call it such only in hindsight. What prevented Saccheri from exploiting his potential discoveries in the same way that these men would do so well—why did he halt at the cusp of non-Euclidean geometry? What was the origin of his 'faith' in the primacy of Euclid's geometry?

There are several answers to these questions here. As we are interested in the emergence of structural thinking in geometry, one is of primary importance. This is that that Saccheri, unlike later geometers, had not yet begun to question the nature and role of his own spatial intuition as a guarantor of the validity of geometry, nor of geometry as a unified science of physical space. In other words, Saccheri, like his predecessors Euclid, Proclus, Khayyam, and others, never seriously doubted the assumption that Euclidean geometry contained a true (if idealized) description of the laws governing the physical space of our world, nor did he question the assumption that the certainty of its axioms was grounded in our immediate intuitive grasp of the properties of that space. Taking Euclidean geometry as the science of space, it made little sense to inquire after 'alternative' geometries – for to be an alternative was straightforwardly to be false. Saccheri, of course, was not alone in these assumptions. Indeed, even in Gauss, one of the earliest mathematicians to take seriously the possibility of an alternative to Euclidean geometry, we find reverence for the intuitive basis of geometry. He writes in an 1816 review of a work on parallels, for instance, that:

> A great part of the text [of Schwab] turns on the contention against Kant that the certainty of geometry is not based on intuition but on definitions and on the *principium identitatis* and the *principium contradictionis*. Kant certainly did not wish to deny that use is constantly made in geometry of these logical aids to the presentation and linking of truths: but anybody who is acquainted with the essence of geometry knows that they are able to accomplish nothing by themselves, and that they put forth only sterile blossoms unless the fertilizing living intuition of the object itself prevails everywhere. Schwab's contradiction seems moreover to rest in part only on a misunderstanding: at any rate, it seems to us that in the sixteenth paragraph of his work (which uses precisely the faculty of intuition from beginning to end, and which claims at the end to prove that "[the postulates of Euclid can be resolved into more general ones, based not on sensation and *intuition*, but on *understanding*]") Schwab must have imagined something else than did the Königsberg philosopher in this terminology for two different branches of the faculty of knowledge.[40]

The view that the certainty of basic geometrical claims was based on intuition (in contradistinction to the apparently non-intuitive character of, e.g., arithmetical claims) was popular throughout the seventeenth, eighteenth, and nineteenth centuries. Its widespread popularity, however, does not mean that it had a clear or even a consistent meaning. Given that (most) mathematicians are not directly interested in the epistemological aspects of mathematics, it is perhaps not surprising that we find very few direct explanations of the notion of intuition in mathematical thinking. Nevertheless, it is mentioned almost everywhere in the 18$^{th}$ and 19$^{th}$ centuries as an essential part of mathematics.

The view that intuition (whatever it might be) was a guarantor of the truth and necessity of Euclidean geometry was not spelled out in detail by the mathematicians who held the view. In Saccheri, we find only the remark that the acute-angled hypothesis is inimical to the nature of the Euclidean straight line. The acute-angled hypothesis was untenable, therefore, only because it opposed Euclidean geometry and Euclidean geometry was the sole science of space. If Euclidean geometry is the only game in town, we cannot recognize the acute-angled hypothesis as true because it contradicts what we already know to be true. But what exactly was the basis of the truth of Euclid's geometry? Saccheri, like most other mathematicians, did not take up this

---

[40] [Gauss 1816, in Ewald 1999, 299-300]; Gauss's conception here of the sterility of any geometrical reasoning not fertilized by intuition clearly echoes Kant's remark that "mathematics must first present all its concepts in intuition, and pure mathematics in pure intuition; that is, it must construct them. If it proceeded in any other way, it would be impossible to take a single step; for mathematics proceeds, not analytically by dissection of concepts, but synthetically, and if pure intuition be wanting there is nothing in which the matter for synthetical judgments *a priori* can be given" [Kant 1950, 30].

question, but an answer was given in much greater philosophical detail in the work of Immanuel Kant (1724-1804). Subsequently, Kant's philosophy became, with varying degrees of explicitness, the *de facto* conceptual framework for many mathematicians, philosophers, and laymen interested in the nature of geometry and geometrical reasoning. Even despite Kant's own lack of mathematical expertise, his account of the intuitive basis of geometry seemed to capture the prevailing view of geometry as a spatial science, and was, accordingly, widely adopted and influential in variously altered forms well into the twentieth century. For instance, we find Paul Bernays, as late as 1930—well after the development of a seemingly anti-Kantian, non-intuitive approach to geometry—writing that:

> We can separate Kant's basic idea that mathematical knowledge, and in general the *successful application of logical inference, rests on intuitive evidence*, from the particular formulation that Kant gave to this idea in his theory of space and time. In this way we simultaneously gain the possibility of doing justice to the very elementary character of mathematical evidence and to the level of abstraction of the mathematical attitude.[41]

Thus Kant's basic claim—that the intuition of space lies at the basis of geometric knowledge and reasoning—was taken up in forms which Kant himself could not possibly have conceived. As we'll see below in more detail, the Kantian view of geometry was surprisingly flexible in the face of the radically new results and approaches to geometry developed throughout the nineteenth century. Before we turn to these results, however, it will be useful to explore Kant's views on geometry, especially given their influence over the last two centuries or so.

## 2.1 Kant's philosophy of geometry

Despite a brief early interest in the possibility of higher-dimensional geometries,[42] Kant's most important contribution to the philosophy of geometry was in the systematic philosophical form he lent to the traditional belief in the certainty of Euclidean geometry. He helped to cement the traditional view in two ways. First, he offered a somewhat plausible (if complex) description of the seemingly contradictory combination of intuition and apriority involved in geometric certainty; and, second, he tacitly worked within the universe of Euclidean geometry and its 'space'—thus connecting his novel view of the nature of physical space with the familiar particulars of three-dimensional, Euclidean space. These views are found at various points throughout his work, though perhaps most prominently in the *Critique of Pure Reason*, where he writes that

> Geometry is a science which determines the properties of space synthetically and yet *a priori*. What, then, must be our representation of space, in order that such knowledge of it may be possible? It must in its origin be intuition; for from a mere concept no propositions can be obtained which go beyond the concept—as happens in geometry. [...] Further, this intuition must be *a priori*, that is, it must be found in us prior to any perception of an object

[41] [Bernays 1930-1931, in Mancosu 1998, 243]. These two approaches to geometry (the intuitive and the logical) are even more explicitly outlined and separated in [Hilbert and Cohn-Vossen 1999, iii-v], with the focus of the work being the development of the intuitive approach.

[42] Cf. [Kant 1747, in Ewald 1999, 133-134].

> and must therefore be pure, not empirical, intuition. For geometrical propositions are one and all apodeictic, that is, are bound up with the consciousness of their necessity; for instance, that space has only three dimensions. Such propositions cannot be empirical or, in other words, judgments of experience, nor can they be derived from any such judgments [...]. How, then, can there exist in the mind an outer intuition which precedes the objects themselves, and in which the concept of these objects can be determined *a priori*? Manifestly, not otherwise than in so far as the intuition has its seat in the subject only, as the formal character of the subject, in virtue of which in being affected by objects, it obtains *immediate representation*, that is, *intuition*, of them; and only in so far, therefore, as it is merely the form of outer *sense* in general. Our explanation is thus the only explanation that makes intelligible the *possibility* of geometry, as a body of *a priori* synthetic knowledge.[43]

Here Kant makes at least two important claims regarding the nature of geometry. First, he characterizes it as a science of space. And, second, he suggests that this science is both synthetic and *a priori*—i.e., that its truths can be known prior to all experience *and* that knowledge of these truths requires some form of intuitive input. This remark would have appeared contradictory to a philosophical public for whom intuition was a type of experience, which meant that it was, by definition, already *a posteriori*. To avoid this apparent contradiction, Kant introduces the concept of a *pure* intuition, which would become one of the cornerstones of his mature critical philosophy. There are two such pure intuitions (space and time) which are, for Kant, the epistemological sources of geometry and arithmetic, respectively. He divides up intuitions and concepts as follows:

> Both intuition and concepts are either pure or empirical. They are *empirical* if they contain sensation (sensation presupposes the actual presence of the object); they are *pure* if no sensation is mixed in with the presentation. Sensation may be called the matter of sensible cognition. Hence pure intuition contains only the form under which something is intuited, and a pure concept contains solely the form of the thought of an object as such. Only pure intuitions or concepts are possible a priori; empirical ones are possible only a posteriori.[44]

As one of Kant's main interests is to explain the certainty of geometrical reasoning—by which is meant, tacitly at least, the rigorously developed axiomatic presentation we find in Euclid—he begins here by contrasting the source of geometrical knowledge with the error-prone sources of knowledge within the 'empirical' sciences. These other sciences, along with other unscientific areas of human endeavour, are rooted in experience and the empirical—they deal with the specific matter of cognitions. On Kant's view, the so-called matter of these cognitions, what he called 'sensations', can only be delivered through our contact with the physical world. For Kant, this places the truths related to these areas of cognition solely within the realm of the *a posteriori*. The sciences of pure intuition (e.g., geometry), by contrast, deal exclusively with the form of cognition, and this form is given prior to all possible experience. Indeed, Kant takes great pains to argue that there must be such a form, prior to experience, for experience to be possible in the first place. So, geometrical intuition differs from empirical intuition in that it does

[43] [Kant 1787, B40-41].
[44] [Kant 1787, A50-51/B74-75].

not come into contact with the specific matter or content of sensations, but concerns only their form.

Another distinction Kant is interested in making is that between pure intuitions and pure concepts. This distinction allowed him to distinguish analytic *a priori* knowledge from the peculiar synthetic *a priori* knowledge he wished to attribute to geometry and arithmetic. For Kant, if geometry were to be construed as having its basis in a pure *concept* of space, it would be a purely analytic discipline, in which no new truths other than those already contained in its basic concept(s) could be found. A purely conceptual geometry would be very similar, for Kant, to the discipline of logic. If it was in fact the case that geometry was merely a formal discipline, this would not be a problem. However, geometry is construed by Kant as a science of space, which has experiential import. By introducing the new category of pure intuition, Kant attempts to break the close connection he and others had previously thought to exist between analyticity and apriority, and thus to allow geometry formal necessity as well as the capacity to help explain the world of spatial experience.

With these more general distinctions in mind, let us now turn to Kant's remarks on the specific qualities of our pure intuition of space. Early on in the Transcendental Aesthetic, Kant writes that

> Space is not an empirical concept that has been abstracted from outer experiences. For the presentation of space must already lie at the basis in order for certain sensations to be referred to something outside me (i.e., referred to something in a location of space other than the location in which I am). And it must similarly already lie at the basis in order for me to be able to present [the objects of] these sensations as outside and *alongside* one another, and hence to present them not only as different but as being in different locations. Accordingly the presentation of space cannot be one that we take from the relations of outer appearance by means of experience; rather, only through the presentation of space is that outer experience possible in the first place.[45]

Here, Kant suggests that in order for us to present the objects of experience as distinct from ourselves (or from each other) we must already possess some notion of space within which these differences can be located and made apparent. Thus, this pure intuition is prior to and engenders the possibility of our outer experience. Following this line of thought, Kant further characterizes space as

> a necessary a priori presentation that underlies all outer intuitions. We can never have a presentation of there being no space, even though we are quite able to think of there being no objects encountered in it. Hence space must be regarded as the condition for the possibility of appearances, and not as a determination dependent on them. Space is an a priori presentation that necessarily underlies outer appearances.[46]

Or, stated more clearly with respect to the notion of intuition, he writes that "Space is nothing but the mere form of all appearances of outer senses; i.e., it is the subjective condition of

[45] [Kant 1787, A23/B38].
[46] [Kant 1787, A24/B38-39].

sensibility under which alone outer intuition is possible for us"; space is thus conceived as "a pure intuition in which all objects must be determined".[47] In sum, then, for Kant space is a pure intuition which determines the form of all outer appearance. Because it determines the form of appearances without itself being determined by a further appearance or experience, this intuition is both *a priori* and necessary for any possible experience.

Since, for Kant and those following the Euclidean tradition, geometry was *the* science of space, all of these claims about the nature of space were intended to support the view that the existing body of geometrical truths was certain, necessary, and knowable *a priori*. Kant writes that

> On this a priori necessity [i.e., that of the pure intuition of space] rests the apodeictic certainty of all geometric principles and the possibility of geometry's constructions. For if this presentation of space were a concept acquired a posteriori, drawn from general outer experience, then the first principles for determining [things] in mathematics would be nothing but perceptions. Hence they would have all the contingency that perception has; and it would then precisely not be necessary for there to be only one straight line between two points, but this would be something that experience always teaches us. By the same token, what we take from experience has only comparative universality, viz., through induction. Hence all we could say is: as far as we have been able to tell until now, no space has been found that has more than three dimensions.[48]

So, the crucial source of our geometrical knowledge is the pure intuition of space. As Kant's explanations here are intended to show, the concept of a pure intuition is meant to safeguard the necessity of mathematical (i.e., arithmetical and geometrical) reasoning while at the same time making sense of the obviously synthetic character of specifically geometric reasoning. By countenancing an epistemological source other than pure reason and sensory input, Kant allows for what he had formerly taken to be a contradictory combination of synthetic knowledge with

---

[47] [Kant 1787, A26/B42].

[48] [Kant 1787, A25]. At this point in the second edition of the *Critique of Pure Reason*, Kant slightly amends this paragraph as follows: "By the same token, no geometric principles—e.g., the principle that in a triangle two sides together are greater than the third—are ever derived from universal concepts of *line* and *triangle*; rather, they are all derived from intuition, and are derived from it moreover a priori with apodeictic certainty" [Kant 1787, B39]. It is interesting to contrast the first edition quotation with Kant's 1747 remarks, in which he appears to countenance precisely the empirical sort of geometry that he rejects here. One of Kant's translators, Werner S. Pluhar, makes the following remark: "But, as Kant has indicated, from this original intuition of space [i.e., the pure intuition of space] concepts can be formed, including such concepts as those of empirical space, relative space, Euclidean space, mathematical space" [Kant 1787, B40, translator's footnote 48]. While there is some truth to Pluhar's remark here—we *can* form less general concepts of space from the pure intuition of space, concepts of space limited by, e.g., epistemological, mathematical, or physical criteria—it seems partially misleading with respect to Kant's own views to suggest that we might develop what appear to be more general conceptions of space, like those we find in the theory of relativity or in other non-Euclidean domains of mathematics. The suggestion is misleading because it seems that, for Kant's critical philosophy at least, our pure intuition of space was necessarily three-dimensional and (very likely) 'Euclidean' in other details as well. Unfortunately, a detailed discussion of the specific qualities of the pure intuition of space is not forthcoming in Kant, and it is not clear if the object of this intuition is identical with Euclidean space—though Kant's offhand remarks regarding certain theorems of Euclidean geometry incline one to think as much. This view is perhaps reinforced by Kant's claim that "We have no one who has exceeded Aristotle or enlarged his pure logic (which is in itself fundamentally impossible) just as no mathematician has exceeded Euclid" [Kant 1992, 438].

the *a priori*.[49] In addition to anchoring the putative necessity of geometrical truths, Kant also helped to reaffirm the privileged position of Euclid's geometry as the true description of the characteristics of (our pure intuition of) space. Kant, like most of his contemporaries, was of the opinion that the true science of space had already been given, by and large, in Euclid's *Elements*. Though he was likely aware of the nagging problem of the parallel postulate,[50] this did not hinder his appreciation of Euclidean geometry a jot.[51] He insists, as we see above, that it is not mere chance or empirical fact that there can be only one straight line through any two points, or that space is three-dimensional. Instead, he provides us with a philosophical basis for understanding how it is that these geometric truths, already present in Euclid, are necessary and a *priori* without being deducible in a purely formal manner.

Thus, Kant provided an appealing attempt to explain the certainty of geometry by rooting it in pure intuition. This move accomplished two things. First, it explained geometrical reasoning's separation from the error-ridden arena of everyday experience and, second, it unveiled a realm of certainty distinct from the (apparently empty) realm of analytic truths. Thus, geometry was at once creative and assured; a neat epistemological trick. In addition to this explanation of geometry's certitude, Kant also reinforced the view that this certitude was the exclusive property of the Euclidean view of space. For Kant, our basic intuition of space could only be understood in one way, and that way was paved by Euclid. (As we will see, however, the very general way in which Kant characterized his views of geometry—i.e., with very little direct reference to the specific claims of Euclidean geometry—provided a certain measure of resilience for those views in the face of seemingly contradictory evidence provided by non-Euclidean geometries). Kant's views on geometry gave many geometers (e.g., Gauss) an explicit philosophical basis for their own views. This gave the philosophy of geometry a common language to work with. This resulted, on the one hand, in a firmer entrenchment of the classical Euclidean position (as we can see in the remark from Gauss quoted at the beginning of §2 of this chapter). On the other hand, it also gave those unsatisfied with Euclidean geometry some specific claims to attack. If geometry

---

[49] While it is commonly held (indeed it was held by Kant himself) that he was the discoverer of the synthetic a priori character of mathematical judgments (see [Kant 1783, §2]), [Rusnock 2013] prevents a convincing argument that he was preceded in this discovery by (at least) John Locke, as evidenced by the following passage from Locke's *Essay Concerning Human Understanding*: "We can know then the Truth of two sorts of Propositions with perfect *certainty*; the one is, of those trifling Propositions, which have a certainty in them, but 'tis but a *verbal Certainty*, but not instructive. And, secondly, we can know the Truth, and so may be *certain* in Propositions, which affirm something of another, which is a necessary consequence of its precise complex *Idea*, but not contained in it. As that the *external Angle of all Triangles is bigger than either of the two opposite internal Angles*; which relation of the outward Angle, to either of the opposite internal Angles, making no part of the comple *Idea*, signified by the Name *Triangle*, this is a real Truth, and conveys with it instructive, *real Knowledge*" [Locke 1690, Book IV, Chapter viii, §8].

[50] Kant corresponded with Lambert, and indeed called him "the greatest genius in Germany" [Kant to Lambert December 31, 1764; Kant 1999, 81]. It seems likely that Kant would have been familiar with Lambert's 1766 work on parallels, at least in passing. Given the prominence of the problem of parallels for Euclidean geometry at the time, it seems unlikely that Kant would have been entirely ignorant of the existence of the problem, given his passing familiarity with Euclid and his interest in the nature of geometry and its epistemology. These remain, however, speculative remarks, as, to my knowledge, Kant provides us with no material on the basis of which we can make more definitive claims about his views.

[51] Though Kant only rarely mentions Euclid or his *Elements* by name, he recapitulates Euclidean formulae and proofs at several points. See, for instance, his rendition of Proposition 32 of Book I of the *Elements* at [Kant 1787, A716-717/B744-745].

is the science of space, and we know space through pure intuition, perhaps this intuition is not Euclidean, or perhaps this intuition does not include the parallel postulate as an essential part. Kant's philosophy of geometry both solidified the traditional conception of geometry and pinned it in place.

### 2.2 The emergence of non-Euclidean geometry: Gauss, Bolyai, and Lobachevsky[52]

While geometers from Khayyam to Saccheri were unable or unwilling to countenance the so-called 'acute hypothesis' as a viable alternative to Euclid's geometry, a century later we find no less than three brave souls willing to do so. The men in question were Carl Friedrich Gauss, János Bolyai, and Nikolai Lobachevsky, each of whom arrived at relatively similar conceptions regarding hyperbolic geometry independently of the others. Gauss was the first to countenance non-Euclidean geometry as a worthwhile scientific endeavour, but, due to his notorious reticence to publish, he was also the last of the three to have his findings published.[53] Lobachevsky was the earliest of the three to have published, though his work had very little influence at the time given his isolated position at the University of Kazan, far from the centers of nineteenth century mathematical activity. Bolyai's work appeared in a rather timely manner, but, like that of Lobachevsky, it did not have the immediate impact one would have expected, a circumstance likely due to the relative obscurity of the author, the fact that his results were published in an appendix to a work of his father's, and the still-formidable status of Euclidean geometry with its alleged bedrock of intuitive certainty recently reinforced by Kant. In what follows, I will trace some elements of the early development of non-Euclidean geometry, beginning with one of its earliest forms in the notebooks and correspondence of Gauss.

#### 2.2.1 Carl Friedrich Gauss

It is widely supposed that Gauss was the first to seriously consider non-Euclidean geometry.[54] Due, however, to his unwillingness to publish and his fear of the "clamour of the Boeotians"[55] over the suggestion that we ought to take seriously any geometry other than Euclid's, almost nothing of his work on the subject was known publicly until after his death in 1855. Collected, this material is rather thin, comprising scattered remarks in letters, a few notebook sketches, and two papers.[56] From the letters, however, we find suggestions from Gauss that he had expended a

---

[52] In this sketch of a rather complex conceptual history I must remain somewhat selective. After Saccheri and prior to the publication of Lobachevsky's results, work of undoubted importance was accomplished in the theory of parallels by several men. Of particular note are works by Lambert, by the German jurist Ferdinand Karl Schweikart (1780-1859), and two publications by Schweikart's nephew, Franz Adolf Taurinus (1794-1874).

[53] Gauss is known to have followed the maxim "*Pauca sed matura*" (Few, but ripe), which was exemplified in his attitude toward publishing; despite the lateness with which his thoughts on non-Euclidean geometry appeared, they nevertheless became rather more influential than those of either Bolyai or Lobachevsky, due in large part to his reputation.

[54] Cf. [Kline 1972, 877-878], [Ewald 1999, 297], [Gray 1979, 76], and [Dunnington 1955, 174-190].

[55] Cf. Gauss' letter to Friedrich Bessel (1784-1846) of the 27th of January, 1829: "But I shall not for a long time work up for public consumption my very *extensive* investigations into these matters [i.e., non-Euclidean geometry], and this will possibly not occur in my lifetime, since I fear the clamour of the Boeotians if I were to express my opinion *completely*" [Gauss 1863-1929, Band VIII, 200; translated in Ewald 1999, 301].

[56] His correspondence, beginning at the close of the eighteenth century, with Farkas Bolyai, Gerling, Taurinus, and Bessel, is of particular note in connection with his researches into non-Euclidean geometry. The papers in question

considerable amount of mental energy upon the subject of non-Euclidean geometry, even if he had not deigned to commit much to paper in any definitive form.

Gauss was a long-time correspondent with the Hungarian mathematician Farkas Bolyai, the father of János Bolyai. After years of toil, the latter had finally produced a detailed work on non-Euclidean geometry. This work, *The Science Absolute of Space: Independent of the Truth or Falsity of Euclid's Axiom XI (which can never be decided a priori)*, was published in 1831 as an appendix to a work of his father. Subsequently, his father forwarded a copy to his friend Gauss in 1832.[57] Gauss' reply, from the 6th of March 1832, was as follows:

> If I commenced by saying that I am unable to praise this work (by Janos), you would certainly be surprised for a moment. But I cannot say otherwise. To praise it would be to praise myself. Indeed the whole contents of the work, the path taken by your son, the results to which he is led, coincide almost entirely with my meditations, which have occupied my mind partly for the last thirty or thirty-five years. So I remain quite stupefied. So far as my own work is concerned, of which up till now I have put little on paper, my attention was not to let it be published during my lifetime. Indeed the majority of people have not clear ideas upon the questions of which we are speaking, and I have found very few people who could regard with any special interest what I communicated to them on this subject. To be able to take such an interest it is first of all necessary to have devoted careful thought to the real nature of what is wanted and upon this matter almost all are most uncertain. On the other hand it was my idea to write down all this later so that at least it should not perish with me. It is therefore a pleasant surprise for me that I am spared this trouble, and I am very glad that it is just the son of my old friend, who takes the precedence of me in such a remarkable manner.[58]

As is not terribly surprising, János was rather irked by Gauss' reply; Gray writes that he "never really forgave the Prince of Geometers. He was incredulous that Gauss should have made his discoveries before him, and became so convinced of plagiarism that he never published again".[59] We are perhaps in a better position than Bolyai to estimate Gauss' intellectual honesty on the matter.

Whether or not the specific findings later contained in Bolyai's work were known to Gauss, it does seem clear that he had begun to take seriously the problems with the traditional conception of Euclidean geometry as the only possible science of space. From scattered remarks in his correspondence it is evident that Gauss was dissatisfied with the current understanding of geometry, particularly with the belief that the theorems of Euclidean geometry were straightforward truths about the physical world.[60] Moreover, he was cognizant that whatever new

---

are [Gauss 1816] and [Gauss 1822], both of which were published locally in the *Göttingensche gelehrten Anzeige*. Both the notebook material and the relevant correspondence can be found in [Gauss 1863-1929, Band VIII, 157-268].

[57] Farkas Bolyai had actually sent Gauss a copy of the work in 1831, but due in part to the instability caused by the cholera pandemic then making its way across Europe, the book never arrived.

[58] [Gauss 1863-1929, Band VIII, 221; translated in Gray 1989, 97].

[59] [Gray 1989, 97-98].

[60] See, e.g., [Gauss 1863-1929, Band VIII, 159, 200-201; translated in Ewald 1999, 299, 301].

approach was developed, it would likely alter radically our conception of geometry as an *a priori* science grounded in pure intuition, not to mention our understanding of the connection between geometry and the physical world. His distance from accepted tradition is further evidenced in an 1817 letter to the German astronomer Heinrich Wilhelm Matthäus Olbers (1758-1840), in which Gauss writes:

> I come more and more to the view that the necessity of our geometry cannot be proved... Perhaps we shall come to another insight in another life into the nature of space, which is unattainable for us now. But until then one must not rank Geometry with Arithmetic, which is truly *a priori*, but with Mechanics.[61]

Here we see the dual effects of Kant's philosophy of geometry. Gauss speaks, like Kant, in the language of the a priori and the a posteriori but he also comes to doubt that traditional geometry is merely an elucidation of the a priori, pure intuition that Kant had suggested. Given the problems in establishing the parallel postulate, Gauss, aware of Kant's philosophy but considerably more knowledgeable about geometry than Kant, is ready to (privately) consider that the application of Euclidean geometry to the physical world might be an a posteriori affair after all. There are inklings of the empirical grounding of geometrical axioms which we will later find in the work of Gauss' student Riemann. But, from Gauss himself, we do not find much more in his non-Euclidean works than isolated theorems and hints toward a new vision of the subject; for more substantial results, we must turn to the work of the two men whose names have since become most closely associated with the birth of non-Euclidean geometry, namely, Bolyai and Lobachevsky.

### 2.2.2 Bolyai and Lobachevsky

Both Bolyai and Lobachevsky were isolated from the main mathematical centers of their time, with the result that they and their works were little known until the surge of interest in non-Euclidean geometries prompted by the publishing of Gauss' work after his death. Indeed, the two men remained isolated even from each other for many years; Gray writes that

> The work of Bolyai and Lobachevsky is astonishingly similar, and yet each remained in ignorance of the very existence of the other until some years after their work was published. The Hungarian was aware through his father of the Western European work on non-Euclidean geometry, although he appears to have largely gone his own way. The Russian kept apart from the controversy, and his earliest publication on the subject, in the *Kazan Messenger* for 1829, passed completely unnoticed.[62]

Let us turn now to a discussion of some elements of the conceptual content of their work and the ways in which that work diverges from the traditional understanding of geometry as grounded in a basic, Euclidean intuition of physical space. Here we will begin to see the philosophical vacuum that begins to open up in the wake of the mathematical dismantling of the primacy of

---

[61] [Gauss 1817, translated in Gray 1989, 76; also in Gauss 1863-1929, Band VIII, 177].
[62] [Gray 1979, 96].

Euclidean geometry. It is in this vacuum that the structure-focussed approach to mathematics begins to come into focus.

The very title of Bolyai's work, *The Science Absolute of Space: Independent of the Truth or Falsity of Euclid's Axiom XI*, indicates his belief that he had penetrated more deeply into the nature of space than had Euclid or his followers. His own early frustrations— as late as 1821 he was convinced "that the parallel postulate must hold"[63]— certainly encouraged his resolve. In a remarkable letter of 1823 to his father, he writes:

> I have now resolved to publish a work on parallels ... I have not yet completed the work, but the road I have followed has made it almost certain that the goal will be attained, if that is at all possible: the goal is not yet reached, but I have made such wonderful discoveries that I have been almost overwhelmed by them, and it would be the cause of continual regret if they were lost. When you see them, you too will recognize them. In the meantime I can say only this: *I have created a new universe from nothing*. All that I have sent you till now is but a house of cards compared to a tower. I am as fully persuaded that will bring me honour, as if I had already completed the discovery.[64]

His researches were not published until 1831, and even then "passed straight into obscurity,"[65] despite this effusive self-praise. Only much later was his approach to geometry noticed and properly evaluated by others.

One of the key conceptual innovations that Bolyai introduced was his clear use of the notion of 'absolute' geometry (often also called neutral geometry, or 'pangeometry' by Lobachevsky). He used this term to refer to the collection of geometrical theorems which remain true independently of the parallel postulate, i.e., those which can be proved from Euclid's other axioms alone. One of the remarkable achievements of his work was the tacit ascription of independent importance to the absolute or neutral perspective from which we can make decisions about the status of the parallel postulate. This perspective had been implicitly employed as early as Khayyam's discussion of the three possible cases which arise when we suspend judgment regarding the parallel postulate, but neither Khayyam nor Saccheri had dwelt overmuch on the nature of this neutral geometrical perspective itself. What is novel in Bolyai is not solely his discussion of *theorems* from within this perspective, but also his analysis of which of these theorems remained true when the parallel postulate was denied (or, correspondingly, accepted). Thus, he makes a distinction between theorems which are 'absolutely' true—i.e., those which are true whether we accept, reject, or suspend judgment upon the parallel postulate—and those which are true only given the further assumption/denial of the parallel postulate. This is a remarkable change in attitude towards the 'acute angled hypothesis' from that of Saccheri, who, recall, thought of this hypothesis as inimical to the nature of the Euclidean straight line and, therefore, as absolutely false. From within this neutral perspective, Bolyai proves a number of theorems. He also goes on to work within a geometry in which the parallel postulate is denied, namely a hyperbolic geometry developed from the assumption that the acute angled case of Saccheri is true.

---

[63] [Gray 1979, 97].
[64] [J. Bolyai to F. Bolyai, 3 November 1832; quoted in Gray 1979, 97].
[65] [Gray 1979, 98].

As Gray notes, Bolyai and Lobachevsky produced works which were ‘astonishingly similar’. Several of the theorems proved by Bolyai are proved analogously by Lobachevsky. Lobachevsky’s ‘imaginary geometry’ was, like Bolyai’s, based upon the assumption that Saccheri’s hypothesis of the acute angle holds. One of the notable differences between their works is that Lobachevsky did not limit his geometry to representations in the plane. He offers, for example, the figure of the horosphere (a sphere of infinite radius with constant negative curvature) as a surface upon which a non-Euclidean geometry holds when its basic terms are given appropriate interpretations. Such spherical ‘models’ of hyperbolic and other geometries would later play a role in making non-Euclidean geometries acceptable and scientifically useful.

Lobachevsky, perhaps even more than Bolyai, was also quite well aware of the implications which his pan-geometrical perspective had for the traditional, Kantian underpinnings of Euclidean geometry. He writes:

> The fruitlessness of the attempts made, since Euclid’s time, for the space of 2000 years, aroused in me the suspicion that the truth, which it was desired to prove, was not contained in the data themselves; that to establish it the aid of experiment would be needed, for example of astronomical observations, as in the case of other laws of nature. When I had finally convinced myself of the justice of my conjecture and believed that I had completely solved this difficult question, I wrote, in 1826, a memoir on this subject.[66]

Here we see that the continual failure of Lobachevsky’s early (1815-1817) attempts to prove the parallel postulate brought him to the conclusion which Gauss had suggested already in 1799. Namely, that the basis of the truth of Euclidean geometry was not to be found in our *a priori* intuition of space but, rather, upon empirical considerations. Two thousand years of history had shown that establishing the truth of at least some aspects of geometry *did* require what Kant would call the ‘matter’ of sensation (i.e., astronomical observation). Thus, Euclidean geometry could not be considered an *a priori* science. Accordingly, Lobachevsky’s work on Euclidean geometry “compelled [it] to take its place again among the experimental sciences”.[67] With his work on pangeometry, however, Lobachevsky yet secured the *a priori* status of something recognizably geometric: “... Pangeometry becomes a branch of analysis, including and extending the analytical methods of ordinary geometry”.[68] Lobachevsky’s pangeometric approach detached an important portion of Euclidean geometry from the intuitive underpinnings which had previously linked geometry to physical space. This conceptual move partially enabled the shift away from the conception of geometry as the science of physical space towards a more abstract understanding of geometry as the science of a number of possible spaces. From the pangeometric perspective, as from those of Gauss’s astral geometry and Bolyai’s absolute geometry, Euclidean geometry is treated as a special case whose applicability to the physical world cannot be proven *a priori*, but must instead be shown by empirical means. The peculiarity of the parallel postulate in relation to Euclid’s other postulates finally makes sense. Unlike the other postulates (whose

---

66 [Lobachevsky, quoted in Bonola 1955, 92]; the memoir referred to is his *Exposition succincte des principes de Géométrie*.

67 [Bonola 1955, 93].

68 [Lobachevsky, quoted in Bonola 1955, 94].

apriority and certainty were not yet to be challenged), the parallel postulate's truth was connected to some peculiarity about our physical world, a peculiarity which could not be derived from our intuition of space, or from the nature of reason, but instead had to be measured. This initial step, shared in common by Gauss, Bolyai, and Lobachevsky, allowed the dismantling of the Kantian dogma of the intuitive basis of geometrical necessity to begin. It also planted the seeds for the emergence of geometrical spaces—plural—conceived abstractly without direct reference to any intuitive or physical manifestations.

## 2.3 Challenges to Kant

With their work, Bolyai and Lobachevsky performed a key conceptual manoeuvre which would eventually help enable a significant alteration (one might say revolution) in the way that geometry was perceived. Their countenancing of alternative geometries was a necessary first step in the dismantling of the traditional picture, but it was not sufficient.

It was a necessary step because without presenting a coherent and reasonably consistent alternative to Euclid's geometry, there could be no question of abandoning the view of geometry as the science of space. If geometry is about space, and there is one true description of space, then any putative alternatives to this description must necessarily be false. But, if it can be shown that a non-Euclidean form of geometry is a viable description of space, then the question must be asked: how can we know *which* alternative is the correct one? Prior to the work of Gauss, Bolyai, and Lobachevsky, this question seemed literally nonsensical. But this step was not enough to usher in a completely new understanding of the nature of geometry overnight.

Many mathematicians remained certain of geometrical truths, and this certainty seemed to require explanation, given the apparent similarity between geometry as a concrete science (the science of space) and the error-prone empirical sciences. Kant's views seemed to offer an appealing path towards explaining this apparently contradictory combination of certainty with the concrete. Though it is a commonplace to insist that the development of consistent non-Euclidean geometries overturned the certainty with which the Kantian view of geometry was held, this is a somewhat misleading claim. The claim is misleading insofar as it fails to take into account the remarkable resilience which those views possessed in the face of seemingly recalcitrant evidence. And, even more, such a claim can seem to anachronistically mischaracterize the nature of the relevant counter-evidence as understood by proponents of the Kantian view.

As we've seen, prior to the emergence of non-Euclidean geometries, Kant had attempted to ground the apriority of geometry (i.e., Euclidean geometry) in its *intuitive* necessity. The basis of geometry was our pure intuition of space, and, I have argued, this intuition itself was at least tacitly assumed to be Euclidean in nature. For Kant and the legion of geometers who took his view to be a more or less accurate description of their pre-theoretic conception of geometry, our pure intuition of space presented (with certainty and necessity) a space which was three-dimensional, isotropic (i.e., whose geometrical properties and relations are invariant with respect to direction, location, or size), which obeyed the parallel postulate, etc. In other words, the form of our pure intuition of space was taken to be Euclidean. Thus, while the *logical* possibility of non-Euclidean geometries would likely have been admitted by Kant, such geometries would nevertheless have been impossible for us to intuit (as the pure intuition of space already contained specific geometric content) and could therefore not be true in the sense that Euclidean geometry was understood to be true. And, given Kant's popular dictum that thoughts without (intuitive) content are empty,[69] it seems likely that he would have seen non-Euclidean geometries as interesting but scientifically vacuous constructions—merely 'logical' constructions which could not be applied to the realm of experience insofar as that realm is always structured and made possible by our pure, Euclidean intuition of space. Thus, non-Euclidean geometries could not be considered sciences at all, at least not in the sense of a theory applicable to some empirical subject matter.[70] Far from sciences, these descriptions of alternative 'spaces' would be better understood as arbitrary logical games, with no application to reality. Indeed, something like this seems to have been Frege's view, inspired by Kant, as late as the turn of the century:

> The question at the present time is whether Euclidean or non-Euclidean geometry should be struck off the role of the sciences and made to line up as a museum piece alongside alchemy and astrology. If one is content to have only phantoms hovering around one, there is no need to take the matter so seriously; but in science we are subject to the necessity of seeking after truth. There it is a case of in or out! Well, is it Euclidean or non-Euclidean geometry that should get the sack? That is the question. Do we dare treat Euclid's elements, which have exercised unquestioned sway for 2000 years, as we have treated astrology? It is only if we do not dare to do this that we can put Euclid's axioms forward as propositions that are neither false nor doubtful. In that case non-Euclidean geometry will have to be counted amongst the pseudo-sciences, to the study of which we still attach some slight importance, but only as historical curiosities.[71]

Frege here puts forward the view that one must inevitably choose between Euclidean and non-Euclidean geometries—with the heavy bias being in favour of Euclid. This view was quite common during the early years of non-Euclidean geometries and was based upon two key presuppositions. The first is the (Aristotelian) view of distinct sciences differentiated according to classes of truths related to particular domains of entities. Geometry, on this view, was the science which studied the class of truths related to spatial objects, whose characteristics are idealized in points, lines, planes, etc. And, second, it is based upon the view that Euclidean and

---

[69] The oft-quoted remark is to be found early on in the *Critique of Pure Reason*, in the introductory preamble to the transcendental logic: "Thoughts without content are empty; intuitions without concepts are blind" [Kant 1787, A51/B75].

[70] Kant writes in the *Prolegomena*, for instance, that "Pure mathematics, and especially pure geometry, can have objective reality only on condition that they refer merely to objects of sense" [Kant 1950, 34]. Geometry's way of fulfilling this requirement is by referring only to the necessary spatial form of the objects of sense. Thus, a geometry constructed upon an alternative concept of space (i.e., one which differs from our pure intuition of space) can never connect with objective reality insofar as it fails to refer in any way to the objects of sense, which are determined exclusively by way of our pure intuition of space. So, if our pure intuition of space is Euclidean, then any alternative conception of space which hopes to achieve objective reality mustn't contradict any of the postulates of Euclidean geometry. On such a view, alternate geometries which endeavour to attain the status of sciences can only do so on the condition that they are special cases of Euclidean geometry. Given these views, we might consider Kant a kind of proto-structuralist about geometry, but I think this would be to ignore the fact that for Kant such 'structures' would be empty and meaningless (something which all but the most extreme mathematical structuralists would find deeply troubling).

[71] [Frege c.1900, 169].

non-Euclidean theories were both geometries – thus, that they were competing descriptions of the same domain of objects. So, accordingly, if a science is a (true) description of some collection of objects, and if there are two sciences which purport to describe the same aspects of the same collection of objects, it seems to be the case that either the two sciences say the same things in different ways (in which case one will do), or it will be the case that one of the sciences is simply false—as Frege implies here. Alberto Coffa describes the early prevalence of this supposed competition between Euclidean and non-Euclidean geometries as follows:

> When non-Euclidean geometries first appeared on the scene, they were naturally understood as rivals of Euclid, and their primitives were therefore taken to designate the same things as Euclid's, whatever they might be. Presumably, since both Euclid's and Lobachevskii's were *geometries,* they had to be interpreted as attempts to identify the class of truths concerning the very same domain of entities. Their points, straight lines, and so on had to be the same in both theories.[72]

Given the long tradition and apparent intuitive plausibility supporting Euclidean geometry, it is not surprising, then, that early work on the legitimation of non-Euclidean geometry focussed on connecting it in some way to 'real' geometry, i.e., Euclidean geometry. One of the best known, and most influential, examples of such an attempt was Eugenio Beltrami's (1835-1900) model of hyperbolic geometry constructed within a fragment of Euclidean space. Somewhat surprisingly, such attempts to connect non-Euclidean geometry to the legitimacy of Euclidean geometry were instrumental in the (eventual) dismantling of the Kantian picture. Beltrami's model is developed in two separate papers of 1868. These papers were subsequently generalized by Felix Klein (1849-1925) in 1871. Thus, the resulting model of hyperbolic geometry is usually called the Beltrami-Klein model.[73] Given that the emergence of this model played an important role in challenging the pervasive Kantian view of geometry, it will be helpful to discuss some of the key conceptual moves employed in its construction.

### 2.3.1 The Beltrami-Klein model of hyperbolic geometry

The Beltrami-Klein model begins within a two-dimensional Euclidean space, upon which we construct an open disc (i.e., all the points contained within the radius of a circle $\gamma$, excepting the points which lie on the circumference of the circle). Now, we interpret the 'lines' of Lobachevsky's hyperbolic geometry as the open chords of this open disc (i.e., the chords of $\gamma$ excluding the endpoints which lie upon the circumference of $\gamma$). We then consider two 'lines' (i.e., open chords) $a$ and $b$ to be parallel if they have no point within $\gamma$ in common. If the Euclidean lines which extend the open chords intersect outside $\gamma$, the hyperbolic 'lines' are nevertheless called parallel as long as they have no point *within* $\gamma$ in common (see Figure 3). Here the open disc represents the hyperbolic plane, the open chords of $\gamma$ represent hyperbolic lines, and the relation holding between two lines which do not have a point within the hyperbolic plane in common replaces the Euclidean parallel relation. Other basic Euclidean concepts are translated with varying degrees of familiarity: a 'point' is a Euclidean point within the open disc; such a point 'lies on' a hyperbolic line in just the same way it would lie on a Euclidean line; a

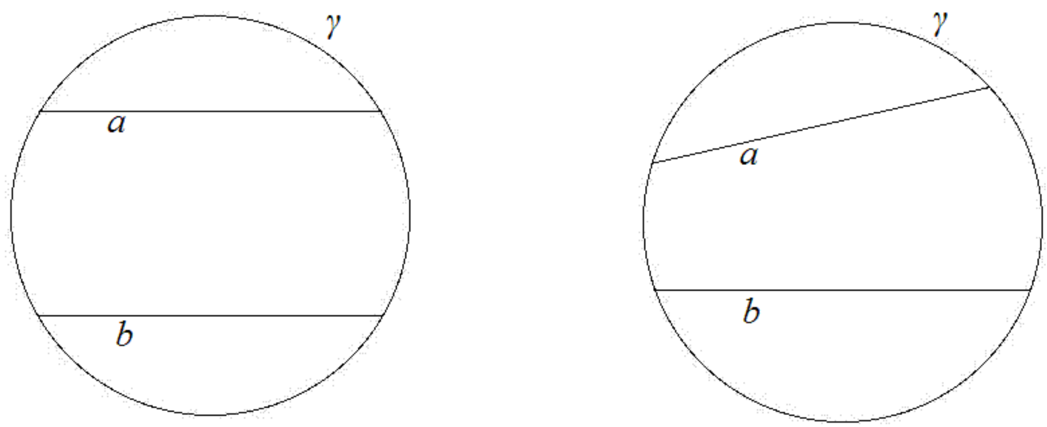
---

[72] [Coffa 1986, 7].

[73] The references are to [Beltrami 1868a], [Beltrami 1868b], and [Klein 1871], respectively.

hyperbolic point lies 'between' two others in the familiar way, etc. The key difference here is that the points and lines of the hyperbolic plane all lie within the open disc. (Hyperbolic congruence and perpendicularity are not as easily defined, as they require some tinkering with rather tricky metrical concepts. I will delay discussion of these difficulties until the following subsection.)

**Figure 3: Parallel lines within the Beltrami-Klein model**




Essentially, then, this method of modelling shows that we can realize Lobachevsky's hyperbolic geometry inside the space of traditional Euclidean geometry. With this manoeuvre, Beltrami illustrated not only that non-Euclidean geometry was consistent,[74] but, also, that it was strictly reducible to Euclidean geometry—an apparent vindication of the Kantian view that *all* geometry, properly speaking, was Euclidean in the final analysis. So, while Beltrami had provided a model of non-Euclidean geometry, he also reinforced the Kantian view insofar as his model was constructed *within* Euclidean geometry. Somewhat oddly, then, the first real vindication of non-Euclidean geometry as more than a tangle of counter-intuitive propositions at the same time bolstered the confidence of those who believed geometry was ultimately grounded in our Euclidean intuition of space. Coffa describes the situation as follows:

> Despite appearances, the doctrines put forth could have brought nothing but comfort to Kantian souls. For Beltrami's ultimate goal was not so much to interpret as to *reduce* hyperbolic geometry to Euclidean geometry and to argue that there was no more geometric sense in the former than it could derive from the latter.[75]

And, further, that

[74] Consistent, of course, only relative to the consistency of Euclidean geometry – which was not in question at the time given the widely held belief in the supposed epistemological guarantee of spatial intuition (and the assumption that the physical world itself is consistent).
[75] [Coffa 1991, 48].

> Far from posing any threats to Kant's philosophy, Beltrami's work was consistent with and possibly even grounded upon it. Kant had never doubted the logical consistency of non-Euclidean geometries. He would surely have said of hyperbolic geometry that it is impossible but not *logically* impossible (since its "negation," Euclidean geometry, is not logically necessary but only intuitionally necessary). So the fact that there is an interpretation of hyperbolic geometry is hardly surprising, nor is it surprising that the interpretation has to be given in terms of Euclidean intuitable notions. Nor is it surprising that wherever that reduction to Euclidean intuition fails, we must abandon the project of giving an interpretation to Lobachevski's theory. One could hardly find a more appealing package of good news for Kantians in a geometric monograph.[76]

To the Kantian geometer of the time, then, any scientific importance we might wish to lend to new forms of geometry was determined entirely by the possibility of connecting those geometries to Euclid's intuitively guaranteed geometry. Once such a connection was made evident—as in Beltrami's case—Kantian scruples were satisfied and normal geometric business could continue, for a while at least.[77]

But, despite the increased assurance lent to the Kantian understanding of geometry by Beltrami's model, subsequent work by Felix Klein on hyperbolic models, influenced heavily by the British mathematician Arthur Cayley (1821-1895), was to prove instrumental in dislodging the Kantian view.

### 2.3.2 Cayley's projective concept of distance and Klein's generalizations

Let's begin our discussion of Klein and Cayley by first returning to the discussion of the interpretation of hyperbolic congruence and perpendicularity within the Beltrami-Klein model. Now, one might begin by attempting to treat congruence for angles in the normal manner, i.e., by introducing some metric for angles and then calling two angles congruent just in case they possess the same number of degrees according to this metric. The problem with this approach in our case is that we cannot employ a Euclidean metric. If we employ a usual Euclidean measure of length, for instance, then all our hyperbolic lines will have to be considered finite. This is so because once we import a Euclidean metric, the circumference of the circle within which the hyperbolic plane is embedded becomes a measurable part of the hyperbolic universe, whereas in the model it is importantly outside that universe.[78] It seems that the use of any standard Euclidean metric will result in the collapsing of the horizon of the non-Euclidean world into that world, thus eliminating the very distinction required for the model to function in the first place. If we cannot employ any of the usual Euclidean approaches toward measurement, how then do we discuss congruence and perpendicularity within the hyperbolic plane?

---

[76] [Coffa 1991, 49].

[77] This reading of the import of Beltrami's work should not suggest that Beltrami saw himself as proving the relative consistency of hyperbolic geometry. This is chiefly because the notion of a relative consistency proof was not at all clearly formulated, by Beltrami or anyone else, in 1868. If it is misleading, then, to suggest that Beltrami *intended* to prove the consistency of hyperbolic geometry relative to Euclidean geometry, it is nevertheless true to state that his work did not immediately call into question the primacy or truth of Euclidean geometry. See [Gray 2007, 219-220] for more detailed remarks to this effect.

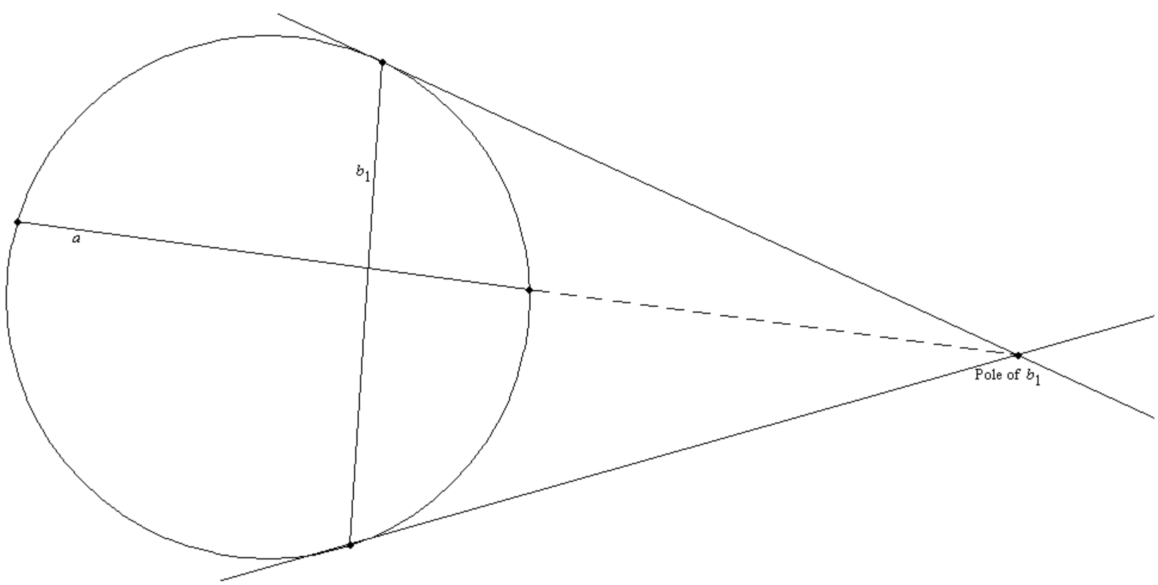
[78] See [Kvasz 1998, 155f.] for a useful discussion of this problem.

The standard approach, developed by Beltrami and further specified by Klein, runs roughly as follows. We say two hyperbolic lines (i.e. open chords of the circle $\gamma$) are ‘perpendicular’ in the following two cases:

1. If one of the two Euclidean chords $a_1$ and $b_1$ is a diameter of the circle $\gamma$, then the hyperbolic lines $a$ and $b$ (i.e., $a_1$ and $b_1$ excluding their endpoints) are said to be hyperbolically perpendicular iff they are perpendicular in the usual Euclidean sense.

2. If neither $a_1$ nor $b_1$ is a diameter of the circle $\gamma$, then $a$ and $b$ are perpendicular iff the Euclidean line which extends $a$ passes through the pole of $b$, where by ‘pole of $b$’ is meant the (Euclidean) point of intersection of the tangents of the radius of $\gamma$ with the endpoints of $b_1$ (Figure 4).

**Figure 4: Hyperbolic perpendicularity in the Beltrami-Klein model**




Now, what is interesting here is that the definition of hyperbolic perpendicularity requires that we extend hyperbolic lines to points at infinity, i.e., ‘Euclidean’ points which do not lie within the hyperbolic plane itself. The use of points at infinity is an important technique of projective geometry, which Cayley had examined and generalized in the 1859 paper from which Klein drew. In contrast to Beltrami’s efforts to legitimate non-Euclidean geometry by embedding it within intuitively acceptable Euclidean geometry, one of Cayley’s main goals was to eliminate the elevated position occupied by Euclidean geometry in favour of a more general standpoint. He famously found such a standpoint within projective geometry, to which we now turn.

The paper which was to influence Klein’s approach to geometry was Cayley’s “Sixth Memoir Upon Quantics”. In that paper, Cayley argues that the projective properties of geometrical figures are their most fundamental, and, further, that the whole of geometry can be reconstructed in exclusively projective terms. Projective geometry was a burgeoning field of study within Cayley's time, but it was still generally treated as a sub-field of properly Euclidean geometry. Historically, it had been developed by Gérard Desargues (1591-1661) and others in order to help explain very concrete—hence *Euclidean*—problems of perspective in Renaissance painting. In the early nineteenth century there was something of a projective renaissance begun by the French mathematician Jean-Victor Poncelet (1788-1867), which gathered steam throughout Germany, France, and Italy.[79]

While Beltrami himself had used projective techniques in the construction of his model of non-Euclidean geometry, his focus was never explicitly upon the medium of projection, or on the nature of the projective plane itself. As we saw above, the Beltrami-Klein model employs certain points outside the circle $\gamma$ in order to define the concept of perpendicularity. These points at infinity are not part of the hyperbolic universe, but are nevertheless still situated within the Euclidean plane into which the hyperbolic plane is projected. The relationship between these points and the lines of the hyperbolic plane is not describable from within the hyperbolic universe itself, but requires a more general perspective. For Beltrami, this perspective was satisfyingly that of traditional, Euclidean geometry. Cayley, on the contrary, developed a point of view from which Euclidean geometry itself appeared to be merely a special case, a specific *type* of metric geometry. Conceptually, the role played by Cayley's projective geometry is quite similar to that of 'neutral geometry' in the works of Bolyai and, more powerfully, Lobachevsky.

One of the key results of Cayley's 1859 paper is his celebrated treatment of the distance function. This treatment of the concept of distance involves characterizing it solely in projective terms (he writes that "a chief object of [this 1859 paper] is the establishment, upon purely descriptive principles, of the notion of distance"[80]); effectively, Cayley elides all the peculiarities of the Euclidean plane by mapping it onto the projective plane and back onto the Euclidean plane once more. He introduces the distance function relative to a particular figure he called the Absolute. The Absolute functioned as a yardstick against which distance could be measured irrespective of the specificities of a given space. More precisely, the process of projecting from a given metric plane into the projective plane and back again to the metric plane allowed Cayley to define the Absolute (and hence distance) by appealing only to projective properties—its non-projective properties were essentially erased by the successive projective transformations. Thus, Cayley provided a method for characterizing distance as a purely projective concept. By doing so, he successfully illustrated that we can treat Euclidean geometry as a particular form of the more general projective geometry, and thus we have his well-known remark that "Metrical geometry is thus a part of descriptive geometry, and descriptive geometry is *all* geometry".[81]

---

[79] Poncelet's two-volume work, *Traité des proprieties projectives des figures* ([Poncelet 1822]) was the instigator of this trend; throughout the middle third of the nineteenth century, algebraic techniques were fruitfully applied to projective geometry by, e.g., Christian von Staudt (1798-1867) and his school in Germany, and by the Italian geometer Luigi Cremona (1830-1903). The algebraic approach to projective geometry was instrumental in the establishment of the school of Alfred Clebsch (1833-1872) at Göttingen, which would prove to be so influential in the development of geometry towards its more recognizably modern forms in the work of Riemann, Klein, Hilbert, and others. (On the German side of these developments, see [Ziegler 1985]).

[80] [Cayley 1859, 61].

[81] [Cayley 1859, 90]; Cayley uses the term 'descriptive geometry' for what we call projective geometry.

Cayley restricted his investigation to the properties of a particular type of Absolute (he employed a specific degenerate conic section), which limited the applicability of his work considerably. Even still, his work allowed him to establish the important result that Euclidean metric geometry can be included within projective geometry. Though this was already quite a general standpoint, it would become further generalized by Felix Klein, beginning in the 1870s with his famous *Erlanger Programm*. One of the central ideas of Klein's papers of the early 1870s was to vary the form of the Absolute itself in order to investigate other possible metric structures, with other sorts of distance functions, all still characterized solely in terms of their projective properties. In effect, Klein combined Cayley's projective insight into the metric structure of Euclidean geometry with an appreciation for the variety of metric geometries made evident in the works of Lobachevsky and Bolyai.[82] Ultimately, this led Klein to classify different metric geometries according to the existence of isomorphisms between those geometries and particular projective transformation groups. Later, with the help of group-theoretic work by his good friend Sophus Lie (and perhaps also an increasing appreciation of the geometric insights of Riemann), Klein extended this program to include transformation groups more generally.

From the point of view of someone like Gauss, it would likely have seemed impossible for Klein's view of a group-theoretically classified system of interrelated geometries to have arisen over the brief span of a few decades. And, indeed, it was only because of a remarkable series of profound conceptual mutations combined with seemingly minor shifts of technical emphasis that such alteration was at all possible. From Beltrami's limited projection of the hyperbolic plane into the Euclidean plane we arrive (with Klein) at a radically more general understanding of the logical relations amongst a wide variety of geometries. For Beltrami, hyperbolic geometry was contained within true, Euclidean geometry; for Cayley, Euclidean geometry was contained within true, projective geometry; for Klein, these isolated logical relations between two geometries were replaced altogether by a higher standpoint, from which many different geometries stand in many different logical relations according to a group-theoretic typology.[83] With the development of the full-blown *Erlanger Programm* we have a clear example of the

---

[82] Klein was well aware that the influence of Cayley on his work was decisive. He wrote, well after the development of his *Programm*, that "This whole manner of viewing the subject [i.e., the viewpoint of the *Erlanger Programm*] was given an important turn by the great English geometer A. Cayley in 1859. Whereas, up to this time, it had seemed that affine and projective geometry were poorer sections of metric geometry, Cayley made it possible, on the contrary, *to look upon affine geometry as well as metric geometry as special cases of projective geometry*" [Klein 1939, 134]. For a clear explanation of the technical details of Klein's generalization of Cayley's approach to metric geometry [Gray 2007, 225-226] is helpful.

[83] It should be noted here that, despite the central role Klein's *Erlanger* typology had in re-orienting understanding of the nature of geometry, Klein himself still resisted the fully abstract conception of geometries developed purely from postulate sets which is characteristic of the most modern approaches. In contrast to Riemann (about whom, more below), Klein's typology was meant to organize—rather than produce—geometrical knowledge. Because Klein was most interested in treating projective geometry as the ultimate foundation of *all* geometry, he was not inclined to include in his typology geometries which could not be easily characterized in projective terms (e.g., elliptic and various differential geometries). After meeting the British mathematician William Clifford (1845-1879)—who translated Riemann's *Habilitationsvortrag* and was very impressed by the possibility of geometries of discontinuous spaces—Klein became more tolerant of a wider variety of geometries, though he was still convinced of the fundamental place occupied by the perspective of projective geometry. (For some insightful remarks on Klein's position and its development, see [Corry 2004a, 33f]).

increasing focus on structure as a central, organizing principle within geometry. It is important to note, though, that the initial developments which led to this shift (e.g., Cayley's and Beltrami's respective works) were not in the least self-conscious attempts to introduce structural thinking into geometry. Beltrami was concerned to illustrate the legitimacy of non-Euclidean geometry by connecting it to the intuitively acceptable objects of Euclidean geometry, while Cayley sought a deeper understanding of the science of space by embedding the objects of Euclidean geometry in the more fundamental properties of projective space. And, while Klein helped shift the focus away from the competition for the rights to our intuition of space and toward the more interesting (and, eventually, far more influential) problem of determining the relationships between a variety of possible geometries, he too was still concerned that the source of our knowledge of these geometries should be found in intuition. So, while each of the moments addressed in the historical story I give above is crucial to the development of a structural, non-Aristotelian conception of geometric science, none of them, not even the highly abstract system of Klein, can be said to fully or self-consciously embody such a conception.

2.3.3 Physics and geometry

In addition to the challenges posed to Kant's philosophy of geometry by mathematicians, there was significant trouble raised on the physical front as well. In order to see this, let's take a (very brief) detour through the pre-Gaussian understanding of the relation between geometry and physics.

For many physicists and mathematicians prior to the acceptance of non-Euclidean geometry, the mathematical study of geometry was understood to be separate from physics, though whether this separation was a matter of logical priority or a methodology varied from thinker to thinker. Thus for instance we find Isaac Newton (1643-1727) claiming that, while geometry is, in a sense, a part of (universal) mechanics, it was nevertheless importantly distinguished by its peculiarly perfect method. As he notes in the preface to his *Principia*,

> To practical mechanics all the manual arts belong, from which mechanics took its name. But as artificers do not work with perfect accuracy, it comes to pass that mechanics is so distinguished from geometry, that what is called perfectly accurate is called geometrical, what is less so is called mechanical.[84]

So, while Newton's geometry was a part of mechanics insofar as it pertained to the study of the mechanical principles which animated the physical world, it was nevertheless that specific part of mechanics which abstracted from the particularity and inaccuracy of our actual measurements of the physical world to consider the exact principles, proportions, and demonstrations which made them possible in the first place. As Guicciardini puts it,

> Rather than excluding mechanics from the realm of geometrical exactness, Newton proposed to subsume geometry under mechanics. Geometry [for Newton] is founded upon

[84] [Newton 1845, lxvii].

> mechanics, since the description (or the construction) of geometrical objects appertains to mechanics.[85]

The important difference between Newton's view and the more straightforward (and later) view that physical geometry is subject to the dictates of physics, is that Newton held that the object of geometry was the perfectly precise realm of absolute space—a space whose perfection is foreign to the imprecision and relativity of the space we actually experience. As he puts it,

> Absolute space, in its own nature, without regard to anything external, remains always similar and immovable. Relative space is some movable dimension or measure of the absolute spaces; which our senses determine by its position to bodies; and which is vulgarly taken for immovable space.[86]

The imperfect measurement and theorization of this absolute space was the domain of practical mechanics, while the objects of geometry were generated by an appreciation of the perfect or idealized motions which lay beneath our imperfect measurements and observations. In this sense, then, geometry at least partially shared its domain with practical mechanics, and both of them were subsumed under the general heading of rational or universal mechanics. Despite the fact that the perfect objects of geometry are, in this loose sense, generated by our imperfect observations of the physical world, they remain methodologically separate from them.

Now, while Newton's interest is not otherwise in the elaboration of a complete philosophy of geometry, his general views dovetail (in some respects) with the later Kantian philosophy of geometry. Newton suggests that geometry is a component of rational (or universal) mechanics, and that its objects are generated by our observations of motion. This is not to say that they are generated by any particular motions or observations, but by the perfect motions which are possible within absolute physical space. An obvious and important distinction between Kant's account of space and Newton's understanding of physical space is Newton's eagerness to attribute the perfection and endurance of physical space to God's good graces.[87] Kant, by contrast, argues that the nature of physical space is determined by our pure intuition of space. Of course, on Kant's view, for all we know, this pure intuition itself is generated by some noumenal God. For our purposes here, however, this difference in the origins and sustenance of space is not terribly important. What is important is the notion that, for both men (and many others), geometry is, in some respects, prior to any specific physical investigation. We may personally arrive at the investigation of (absolute) space through our empirical considerations. But, as both Newton and Kant hold, the objects of geometry are not themselves solely empirical in character.

This view of geometry as importantly separate (in method or content) from physics endured in various forms for quite some time. It was not really until after the advent of non-Euclidean geometries that the question of the *empirical* investigation of geometry became a live possibility. It was only really when the idea of geometry as connected to empirical investigation became

[85] [Guicciardini 2009, 297].
[86] [Newton 1854, 77].
[87] Though this isn't directly stated in this manner in the *Principia*, this view is made implicitly clear throughout, and particularly in the *General Scholium*.

possible—as in the (probably apocryphal and very early) tale of Gauss' mountaineering expeditions aimed at establishing the curvature of space[88]—that the connection between Kant's pure intuition and space came to be challenged by in earnest by physicists.

Prior to this conceptual move (or 'discovery,' if we prefer), on Kant's account, our conception of space was understood to be entirely synthetic a priori and, therefore, not in need of any external input from physics. Newton's view is not entirely dissimilar: we take the a posteriori deliverances of measurement and empirical physics, and from them we gain access to the geometrically perfect realm of absolute space.

Afterwards, the 'pure' intuitive grasp of space no longer seems a reasonable option. The properties of space are no longer given to us but instead become a matter for investigation. What this meant for physicists, in particular, was that space might end up being *nothing* like Euclid, Kant, and Newton imagined. And, given the development of strange new geometries by Riemann and others, it might turn out to be stranger than anyone could possibly have imagined. Ultimately, the shift towards a conception of the geometry of physical space as a matter for *physical* investigation prompted both physicians and mathematicians to develop interesting new geometries and interesting ways of testing their applicability to physical space. In the end, then, Riemann's weird new geometries—which, to Kant and those of like mind might have seemed mere flights of fancy or amusing but scientifically inert games—ended up providing mathematical descriptions of physical space in the works of men like Einstein and Minkowski. So it is that we find Einstein remarking in 1921 that

> It is clear that the system of concepts of axiomatic geometry alone cannot make any assertions as to the behaviour of real objects of this kind, which we will call practically-rigid bodies. To be able to make such assertions, geometry must be stripped of its merely logical-formal character by the coordination of real objects of experience with the empty conceptual schemata of axiomatic geometry. To accomplish this, we need only add the proposition: solid bodies are related, with respect to their possible disposition, as are bodies in Euclidean geometry of three dimensions. Then the propositions of Euclid contain affirmations as to the behaviour of practically-rigid bodies. Geometry thus completed is evidently a natural science; we may in fact regard it as the most ancient branch of physics. Its affirmations rest essentially on indication from experience, but not on logical inferences only. We will call this completed geometry 'practical geometry,' and shall distinguish it in what follows from 'purely axiomatic geometry.' The question whether the practical geometry of the universe is Euclidean or not has a clear meaning, and its answer can only be furnished by experience. […] I attach special importance to the view of geometry, which I have just set forth, because without it I should have been unable to formulate the theory of relativity.[89]

[88] As the tale goes, Gauss measured the angles of the triangle formed between three mountain peaks in order to determine the curvature of physical space. While it is certain that Gauss did engage in just such a survey, there is little evidence to support the idea that his aim was such a geometrical test. More serious tests of the geometry of physical space would have to wait until considerably later.

[89] [Einstein 2002, 387-388].

As Einstein makes clear here, a prerequisite of his own fundamental advances in physical theory was the decoupling of geometry from the moorings of Kantian pure intuition. After severing the connection between the a priori certainty of Euclidean geometry and the physical investigation of space, physicists like Einstein were free to apply (or attempt to apply) a variety of axiomatic geometries to a particular physical problem. And, from the mathematical point of view, geometers like Minkowski had a new physical incentive to develop entirely new forms of geometry. The result, particularly throughout the first quarter of the twentieth-century was an incredibly fruitful series of collaborations between physicists and mathematicians.[90]

## 2.4 Some philosophical lessons

The overall aim of this chapter has been to illustrate the content of the traditional Euclidean picture of geometry, and the conceptual steps through which this view of geometry as a science of idealized physical objects was eventually overturned in favour of a picture of geometry as a science interested in a variety of abstract structures whose relations to physical space were not of particular mathematical interest. Though we have yet to explore this second picture, with Felix Klein's work, we have arrived at an important threshold between these two conceptions of geometry.

I began by describing several aspects of Euclid's *Elements*. This description is intended to illustrate that his work, and the tradition stemming from his work, exemplified a conception of science which was importantly distinct from the structure-focussed conception of mathematics which emerges just after our historical survey with the work of Hilbert and others. Rather than the study of formal structures, traditional Euclidean geometry was understood by many of its practitioners as a concrete science distinguished by the sort of *objects* it studied—namely spatial objects. In Euclid, we find clear evidence of this object-oriented view: the *Elements* begins with a series of definitions, which are intended to familiarize the reader with the particular objects about which geometers reason (e.g., points, lines, circles, etc.). It then proceeds to the presentation of a number of basic truths or axioms about these objects, truths which must be accepted before geometrical argumentation begins.

My own exposition focussed upon one of these axioms, namely, the parallel postulate. This postulate was chosen as a historical focal point largely because the attempts to prove it gave rise to a number of developments that proved central to the emergence of a structural view of geometry. In hindsight, it is certainly true that history could have taken quite a different route, and geometers might have focused more of their energies on proving, e.g., the Archimedean axiom, or any other of Euclid's axioms for that matter. Had Euclid phrased his axiom of parallels

---

[90] This sort of collaboration is perhaps most evident in the interactions between Minkowski and Einstein, about which Einstein writes the following: "The generalization of Relativity theory has been made much easier through the form given to the special Relativity theory by Minkowski, which mathematician was the first to recognize clearly the formal equivalence of the space-like and time-like co-ordinates, and who made use of it in the building up of the theory. The mathematical apparatus useful for the general relativity theory, lay already complete in the 'Absolute Differential Calculus,' which were based on the researches of Gauss, Riemann and Christoffel on the non-Euclidean manifold, and which have been shaped into a system by Ricci and Levi-civita, and already applied to the problems of theoretical physics" [Einstein 1916, 89]

more simply (imitating Wallis' formulation, perhaps) the history of geometry would surely have been quite different. Despite the fact that the centrality of the parallel postulate for the history of geometry is a mere historical accident, however, it is worthwhile, from a philosophical standpoint, to study attempts at proving the postulate insofar as such a study reveals a number of deep and important conceptual shifts in the understanding of geometry as a science. Through this history, we can witness the slow and largely accidental shift away from the traditional understanding of geometry as the science characterized by its study of the genus of spatial objects towards the modern understanding of geometry as a science which studies a variety of abstract structures and relations between those structures.

From the array of attempts to prove the postulate, I have focused my attention on a mere handful. We began with those of Ptolemy and Proclus for two reasons. First, because they are some of the earliest attempts of which we know. And, second, because they are exemplary of a common sort of 'proof' of the parallel postulate, the sort which appeals to a disguised equivalent of that very postulate. The unwitting character of these appeals is in direct contrast to attempts at simplification which show that the postulate is equivalent to some simpler and intuitively more appealing statement. Wallis' proof that the parallel postulate is equivalent to the claim that space is everywhere homogenous, or that there exist similar triangles which differ in size, is an example of the latter sort of approach.

From Ptolemy and Proclus I turn to the work of Omar Khayyam. For my purposes, the chief conceptual importance of Khayyam's work lies in his brief exploration of the logical possibility of geometries in which the parallel postulate does not hold. Though Khayyam is quick to dismiss these alternative non-Euclidean geometries, his suggestion was later exploited in much more detail by Girolamo Saccheri. With Saccheri's work, we have an example of the detailed exploration of logically possible geometries which do not obey Euclid's parallel postulate. One of these possibilities, based upon an 'acute-angled' hypothesis, is explored in particular detail. Despite the effort which Saccheri devotes to the development of this non-Euclidean geometry (as we would understand it today), he abruptly rejects the coherence of this alternative in favour of traditional Euclidean geometry. To a modern eye, this rejection is rather surprising—it seems as if he simply decides to stop arguing. This is because we are intimately familiar with the structural approach to geometry which involves dealing with multiple geometries and multiple spaces as a matter of course. So why does Saccheri not take the opportunity to describe just this sort of alternative geometry?

My claim was that an implicit cluster of assumptions about the nature of geometry as a science guided Saccheri away from the implications of his own detailed work and toward a more traditional understanding of the status of logically possible alternative geometries. I discuss the philosophy of geometry of Immanuel Kant as a particularly influential codification of many of these implicit assumptions about geometry. Of central importance among these assumptions are the claims that geometry is *the* science of space, and that geometrical truths are certain and necessary because they are delivered to us by a special faculty of intuition somehow exempt from empirical doubt. Given the widespread acceptance of these sorts of claims, then, Saccheri's refusal to countenance the coherence of a geometry alternative to Euclid's is not terribly surprising. In order to do so, Saccheri would have had to reject or seriously alter the view that geometry as he understood it was the science of space: for if there is only one science of space, and there are conflicting claims about the nature of that space, one of those claims must necessarily be rejected. Given the traditional weight lent to the theorems of Euclidean geometry, it is not at all surprising that Saccheri would have chosen Euclid's geometry over his own upstart alternative (a sentiment which was echoed much later by Frege). Even if he could have rejected the view that there is only one science of space, Saccheri would still have had to face the epistemological problem implicit in the view that the truths of geometry were certain, and that this certainty was grounded somehow in our very apprehension of the nature of space. No small challenge, then.

It would take another hundred years or so before we find mathematicians capable of overturning or questioning these implicit assumptions. Gauss, Bolyai, and Lobachevsky all began to call into question the view that Euclidean geometry was the true description of physical space, and in so doing they slowly opened the door for the view that the applicability of the theorems of geometry to physical space was itself an empirical question which could not be decided on purely *a priori* grounds. This move required a huge conceptual shift away from the understanding of geometry as the science of space. From the work of these three men, a view developed within which it was possible to explicitly investigate questions pertaining to particular sorts of spaces (e.g., Euclidean or hyperbolic space). The development of a 'neutral' geometric perspective from which claims about the structure of a particular *kind* of space could be decided was a crucial element in the larger story of the emergence of a structure-focused understanding of geometry.

With Beltrami's work on modelling non-Euclidean geometry within the Euclidean plane, we have a much more explicit instance of the tendency to discuss the properties of particular kinds of geometrical structures. Beltrami's work is also important from a philosophical standpoint for the fact that it was one of the earliest instances in which relations between what we might anachronistically call 'geometrical structures' were discussed with technical sophistication. Beltrami's illustration that we can model hyperbolic geometry by embedding hyperbolic space within a fragment of Euclidean space is highly suggestive of a more general perspective from which both Euclidean and hyperbolic 'space' appear to be special cases. Despite Beltrami's own understanding of his model as an illustration of the conceptual priority of Euclidean geometry, his work had the effect of drawing geometric attention to relations between different spaces, or different logically possible spatial structures.

In the hands of Felix Klein, the gestures towards this more general perspective—within which different spaces and relations between those spaces are examined—became far more sophisticated. Relying upon earlier work of Cayley in projective geometry, Klein extended Beltrami's work by characterizing a whole host of different geometrical structures (different 'spaces') in terms of a limited group of projective properties.

All of this helps to illustrate the slow unravelling of an object-oriented approach in favour of a more structure-focussed view within geometry. Early attempts to prove the parallel postulate within this traditional object-oriented understanding of geometry were largely unsuccessful. When we arrive at the work of Girolamo Saccheri, we have the first real hints that progress in the theory of parallels might require overturning one or many of the implicit assumptions which undergirded the traditional view of geometry. With Gauss, Lobachevsky, and Bolyai, these hints are given a more tangible form and the assumption of the applicability of Euclidean geometry to physical space is made into an empirical matter. From this questioning of the *a priori* certainty of the Euclidean nature of physical space, the traditional understanding of geometrical research as the exploration of the properties of idealized physical space was slowly supplanted by a view of geometry as the logical development of various abstract structures. With the work of Cayley, Beltrami, and, eventually, Klein, this structure-focussed view became increasingly more explicit.

## 2.5 Bernhard Riemann's *n*-dimensional manifolds

While there was, as we've seen, a lengthy history leading to the eventual development and legitimation of the specific non-Euclidean geometries created by Bolyai and Lobachevsky, and to the subsequent establishment of Klein's *Erlanger Programm* as a partial result of attempts to model those geometries, it is difficult to avoid treating Bernhard Riemann's 1854 *Habilitationsvortrag*, "On the Hypotheses Which Lie at the Basis of Geometry," as an example of an almost entirely isolated development.[91] In that short lecture, prepared at the behest of his doctoral supervisor Gauss, Riemann produced a radical re-imagining of the nature of geometrical science and its relationship to physical space. We've seen how difficult it was to achieve even minor headway in the advancement of non-Euclidean geometries even up to Gauss's death in 1855. Given these difficulties, the dramatic progress represented by Riemann's 1854 paper is all the more surprising. The first few paragraphs, which set the tone for the rest of the work, are worth quoting at length here:

> It is known that geometry assumes, as things given, both the notion of space and the first principles of constructions in space. She gives definitions of them which are merely nominal, while the true determinations appear in the form of axioms. The relation of these assumptions remains consequently in darkness; we neither perceive whether and how far their connection is necessary, nor, *a priori*, whether it is possible. [...] From Euclid to Legendre (to name the most famous of modern reforming geometers) this darkness was cleared up neither by mathematicians nor by such philosophers as doubtless concerned themselves with it. The reason of this is doubtless that the general notion of multiply extended magnitudes [*ausgedehnten Grösse*] (in which space-magnitudes are included) remained entirely unworked. I have in the first place, therefore, set myself the task of

[91] Riemann himself refers to some remarks made by Gauss in a paper on biquadratic residues (i.e., [Gauss 1831]) and the works of the philosopher Johann Friedrich Herbart (1776-1841) as his main influences in this respect, though he admits their influence was rather limited. Gauss's paper presents a rather general approach to the theory of numbers in which relations across various number domains (natural, real, complex, etc.) are treated via the concept of a manifold. The generality implied by the concept of a number manifold is likely the particular element in Gauss's paper which partially inspired Riemann's own generalizing manoeuvres in geometry, though there is little there to suggest the specific application to 'space' that we find in Riemann. The philosophical views of Herbart—who worked at Göttingen from 1833 until his death in 1841—significantly influenced Riemann's philosophical view of the world. Particularly relevant for us is Herbart's view that the so-called 'pure' intuitions of Kant (i.e., space and time) are not 'pure' in Kant's sense but are, in fact, based upon the mind's own attempt to understand causal experiences. Thus, they are not pre-experiential, *a priori* sources of knowledge, but the result of our attempts to make sense of experience in a coherent fashion. Riemann, as we will see, significantly developed Herbart's philosophical point within a more strictly mathematical setting. (For some of Herbart's views which influenced Riemann, see [Herbart, 1824-1825]—a work that deals with the problem of Kant's pure intuition among other things, and a work which Riemann is known to have read. For a careful analysis of the specific influences of Herbart on Riemann's work, see [Scholz 1982]).

> constructing the notion of a multiply extended magnitude out of general notions of magnitude. It will follow from this that a multiply extended magnitude is capable of different measure-relations, and consequently that space is only a particular case of a triply extended magnitude. But hence flows as a necessary consequence that the propositions of geometry cannot be derived from general notions of magnitude, but that the properties which distinguish space from other conceivable triply extended magnitudes are only to be deduced from experience. Thus arises the problem, to discover the simplest matters of fact from which the measure-relations of space may be determined; a problem which from the nature of the case is not completely determinate, since there may be several systems of matters of fact which suffice to determine the measure-relations of space—the most important system for our present purpose being that which Euclid has laid down as a foundation. These matters of fact are—like all matters of fact—not necessary, but only of empirical certainty; they are hypotheses. We may therefore investigate their probability, which within the limits of observation is of course very great, and inquire about the justice of their extension beyond the limits of observation, on the side of both of the infinitely great and of the infinitely small.[92]

In this relatively brief excerpt we have a startlingly clear formulation of what would become the basis of the 'modern' conception of geometry. Recalling Saccheri's exasperation and Farkas Bolyai's desolation, it is mildly scandalous to witness the casual air with which Riemann overturns the centuries-old understanding of geometrical science. Instead of the study of the idealized elements of an intuitively given, three-dimensional, Euclidean space, Riemann suggests that we treat geometry as a subfield of the study of multiply extended magnitudes (or, more generally, *n*-dimensional manifolds [*Mannigfaltigkeiten*]), some of which have properties similar to Euclidean space and some of which are radically different. On this view, the link between intuition and Euclidean space (if, indeed, there is one) becomes rather beside the point. We do not intuit the truth of the axioms of geometry, but develop them as hypotheses which may or may not turn out to be true of empirical reality. The applicability of these axioms to the physical world (i.e., their empirical truth) is a matter for empirical determination and not of purely formal, pre-experiential necessity delivered by Kantian intuition. It is not surprising that Riemann's work was partly inspired by the philosopher Johann Friedrich Herbart's attempts to reform what he perceived to be Kant's errors (or hyperbole). Precisely this sort of philosophical view of the relationship between physical space and geometric theory was lacking in many earlier attempts to explain problems related to Euclid's parallel postulate.

On Riemann's view, the traditional space of geometry (i.e., three-dimensional Euclidean space) becomes one instance of an *n*-dimensional manifold amongst many others. Rather than examining this or that specific manifold, Riemann begins by presenting the concept of an *n*-dimensional manifold in its full generality. In order to fix this concept in its most general form, he proceeds by distinguishing its more specific instantiations. He thus isolates a number of geometric properties by which we might fix particular classes of geometries. Some of the crucial properties are: the number of 'dimensions' of a manifold; whether it is discrete or continuous and, if continuous, whether or not it is isotropic; its metric relations; its curvature. We can see the generality of Riemann's point of view by considering how he comes to understand what had, for

---

[92] [Riemann 1868, translated in Ewald 1999, 652-653].

two millennia, been the traditional content of the whole of geometry. Three-dimensional Euclidean space becomes a specific instance of a three-dimensional, continuous, isotropic manifold of constant curvature, where that curvature is zero, and which possesses a particular distance function. Now, each of these adjectival descriptors of Euclidean space can be varied, producing a distinct class of geometries, each describing particular sorts of spaces. In Riemann's very general view, Bolyai-Lobachevsky hyperbolic space and traditional Euclidean space are not terribly different; indeed they are both examples of continuous three-dimensional manifolds of constant curvature—they differ mainly in the fact that hyperbolic space has a constant negative curvature, while Euclidean space has a constant curvature of 0. Thus, while we might discuss the emergence of non-Euclidean geometry as a troubling development for the intuitively grounded Kantian view, it should nevertheless be apparent that the work of Bolyai and Lobachevsky (though an important first step) was the first step toward a radically different view of geometry which would be made more mathematically explicit in the work of Riemann, and more philosophically explicit in the work of Hilbert a half century later.

With Riemann's work, even more than with the developments which led to Klein's *Erlanger Programm*, a significant blow seemed to have been dealt to the Kantian belief that pure intuition grounds geometrical reasoning, though this was not widely realized until considerably later. Riemann begins his lecture by suggesting that the application of our geometrical axioms to the world is not a matter of necessity but one which is established only by empirical investigation. Instead of intuitively grounded, necessary truths about the structure of (the only possible) space, we arrive at more-or-less applicable, empirically-testable hypotheses whose truth is not at all certain. In addition, some of these hypotheses (e.g., those lying at the basis of higher-dimensional geometries, or geometries radically different from the more familiar Euclidean case) will be difficult if not impossible to 'intuit' at all. In order to salvage the Kantian picture in the face of Riemann's work, it would seem that we require a considerably expanded understanding of the notion of intuition, as well as an alteration in our understanding of the epistemological role which such intuition can play within geometric reasoning. And, beyond even Klein, Riemann's work suggests the systematic development of geometries from various sets of hypotheses and slight alterations in coefficients, leading to a typology of spaces and geometries which is entirely indifferent to the deliverances of intuition, whatever they may be. Riemannian geometers might be interested in the connection between a given geometry and physical space, but they certainly needn't be, and usually will not be.

## 2.6 A conceptual vacuum

The increasingly rich assemblage of new geometrical concepts, as well as the rapid development of highly abstract concepts in other areas of mathematics like group theory, created a situation in which the very nature of geometrical science became unclear. Previously, geometry had been seen as a (relatively) concrete science dealing with the laws governing physical space. Its rather odd status as an *a priori* science that, somehow, also applied to the changing world of appearances had been temporarily propped up by Kant's view of space as a pre-experiential, pure intuition.

Having undermined Kant's position and the privileged position of Euclidean geometry, geometers became more fully capable of questioning what role, if any, intuition played within the new, generalized science of geometry. As we will see, this loosening of the epistemological connection between pure spatial intuition and the logical development of geometry enabled a rather significant alteration in the way geometry itself was conceived. We've already witnessed the shift from a view of Euclidean geometry as the science of space to the view of Euclidean geometry as the description of a very particular sort of three-dimensional manifold. Indeed, the distance between the traditional conception of geometry and the more 'structural' approach is most evident in the alteration of the concept of space itself. For Euclid and those following in his tradition, space was *physical* space: geometry may have studied its properties by way of idealized objects like points and lines, but geometry nevertheless studied the properties of our physical space. With the advent of Klein's *Erlanger Programm*, we have a variety of geometries, which operate on a variety of different spaces. Initially, these spaces are determined by the invariants of particular groups of spatial transformations. In Riemann, we have even more such spaces, which are now determined neither by intuition nor by transformation groups, but simply by sets of hypotheses containing varying numerical coefficients. Geometers are free to construct such spaces at their pleasure, and investigate their properties without ever imagining how or even if they can have any physical application. Thus, geometry moved away from being the science which determines the laws governing physical space to the science which studies the properties of occasionally baffling abstract spaces, conceived as particular sorts of multiply extended magnitudes, and the various relations between them.

With these developments in non-Euclidean, projective, and Riemannian geometry, pure geometrical science became disconnected from the physical world, and it was no longer immediately clear how to understand its results or their purposes. What exactly was geometry? What were the abstract structures which it seemed to be studying? For a dyed in the wool Kantian, they might have been seen as logically interesting fictions. But for most late 19$^{th}$ century mathematicians, they seemed to be something quite different.

With Hilbert's work on the foundations of geometry, a new conception of mathematics as the study of formalized axiomatic systems finally began to emerge. Both Riemann and Helmholtz (who helped to popularize and conceptualize Riemann's work) had, earlier, offered sophisticated explanations of the nature of the new geometry, but it wasn't really until the publication of Hilbert's *Grundlagen* and its axiomatization of Euclidean geometry that the new approach was taken up with gusto by the mathematical community at large. Hilbert's work shored up the burgeoning axiomatic understanding of the nature of geometry by applying it to the core of the old tradition. By showing that Euclidean geometry could be fruitfully axiomatized in the new style, Hilbert at once lent credence to the new approach and vindicated the importance of the traditional heart of the science. In the following chapter I will show how Hilbert's work helped to challenge the prevailing Aristotelian conception of science.

# Chapter Two: Frege and Hilbert on axiomatic systems

## §0. Introduction

In the previous chapter, we examined elements of the historical shift away from a view of geometry as the science of idealized spatial objects. The increasing prominence of the formal structure of and interrelations among geometric theories (as opposed to their content, properly speaking) had a destabilizing effect on the traditional understanding of geometry as the science of space. In place of the traditional view of geometry as a science concerned with the relations amongst ideal objects like points, lines, and planes, we find, toward the end of the nineteenth century, an increasing confusion about the nature of geometry as a science. If geometry was not best understood as the idealized study of physical space, what then was its proper object?

In this chapter, we will examine two distinct philosophical attempts to understand the nature of geometry in light of these developments. We will first turn to David Hilbert's influential views, as developed in his *Grundlagen der Geometrie*. Hilbert's work in geometry focused on the development of the deductive consequences of groups of axioms. His philosophical views of geometry give pride of place to the formal characteristics of these axiomatic systems.

After examining Hilbert's re-imagining of geometry, as well as his philosophical views of the nature of axiomatics, we will turn to the views of Frege. Frege, in some ways, was a geometric traditionalist. Unlike Hilbert, he held on to the view of axioms as meaningful truths about the world, and of geometry as a science with a determinate subject matter. In place of uninterpreted formal systems, Frege viewed Euclidean geometry as the study of a particular collection of objects and concepts. Though sympathetic to the idea of higher-level geometric theories (within which, e.g., Euclidean geometry might fall), Frege was nevertheless harshly critical of what he took to be the logical imperspicuity required for Hilbert-style axiomatics. In this chapter, we will first examine Hilbert and Frege's respective views of geometry and axiomatics, then turn to an analysis of Frege's criticisms of Hilbert.

The chapter has two chief aims. First, it will call into question the relatively common view that Hilbert's prescient account of axiomatics simply wins the day when pit against Frege's old-fashioned, rather curmudgeonly criticisms. This will be accomplished by showing that a) Hilbert's views are not entirely in line with the more contemporary model-theoretic approach to axiomatics and b) that Frege's criticisms of Hilbert are not based on a simple recapitulation of the traditional Euclidean/Kantian view of geometry. Second, the chapter will show that, despite the widespread acceptance of an altered form of Hilbert's approach to axiomatics, Frege poses deep questions about the clarity and meaningfulness of this approach which Hilbert and his defenders leave unanswered. The result of this chapter, when combined with the first, is that the traditional Euclidean approach is shown to be incapable of addressing the reality of modern mathematics, while the conceptual clarification of this new reality offered by Hilbert and others faces significant philosophical difficulties.

As will become clearer in subsequent chapters, Hilbert's views are, with some exceptions, representative of a larger trend in mathematical thought within which abstract mathematical structures are characterized in purely relational terms (i.e., as ontologically un-fixed). By illustrating the force of Frege's criticisms of Hilbert's views, then, I hope also to call into question the viability of this more general method of characterizing mathematical structures in this relational, or free-floating way. We will begin the chapter by differentiating Hilbert's views of geometry and (mathematical) science from the traditional account stemming from Euclid and Aristotle.

## 0.1 The Euclidean tradition and Aristotelian science

As we've seen above, in the works of Euclid one finds two sorts of definition. Definitions of the first type were essentially abbreviations: 'circle' for instance, is defined in terms of the more basic notions of point, line, etc. We could just as well cut out the term 'circle' and replace it everywhere with its definition in terms of points and lines and no essential information would be lost. There were, however, definitions of a different sort, which served as pre-systematic guides to orient the reader towards the most basic objects of geometry (i.e., the points, lines, and planes which form its subject matter). Through these definitions, readers were either oriented towards objects with which they were already familiar or introduced to those objects, with varying degrees of success. Within this tradition, axioms were construed as truths about a determinate subject matter; in this case, the (idealized) elements of physical space. In many cases, too, the truth of these axioms was thought to have been grounded in our basic intuition of space—a source of knowledge which came to be considered importantly distinct from both experience and pure logic. These specific claims about the nature of geometric science were themselves embedded within a broader Aristotelian conception of the nature of science in general.

For Aristotle and those following this tradition, human knowledge was divisible into distinct sciences, differentiated according to the genus of objects which they study.[1] Thus, for instance, we can conceive of biology as a science distinct from geometry insofar as the former studies the genus of living objects, while the latter studies the genus of spatial objects. We can establish certain relations (e.g., relations of inclusion, co-extension, etc.)

---

[1] We find echoes of the traditional Aristotelian concept of sciences (differentiated according to the species of objects which they study) within the works of many of the most important thinkers in the history of Western thought. We find, for instance, Thomas Aquinas (1225-1274) writing in the *Summa Theologica* that "a species is a kind of science" [Aquinas 1957, 2.57.2, objection 1; 1005b] and that "according to the different kinds of knowable matter, there are different habits of scientific knowledge" [Aquinas 1957, 2.57.2, response; 1006a-b]; John Duns Scotus (1265/66-1308), too, refers specifically to the concept of science outlined in the *Posterior Analytics* (at 1.2 79b9-12) early on in the *Ordinatio* (see [Duns Scotus 1950, 141-142]); Leibniz was undoubtedly heavily influenced by both Aristotle and the Schoolmen, and his systematic approach to natural philosophy owes much to the Aristotelian tradition despite his more general understanding of species and essences; Bolzano, a partial follower of Leibniz and Aristotle, calls "an aggregate of truths of a certain kind a *science*" [Bolzano 1837, §1.1] and notes that he "shall call the class of truths which assert something about the constitution of space the science of space (geometry), because these propositions form a separate species of truths" [Bolzano 1837. §1.1]. Bolzano departs from this Aristotelian tradition, however, in that he allows for species of truths which are not determined by classes of objects. As we shall see in what follows, Frege, too, held a rather Aristotelian conception of the nature of science.

amongst the sciences themselves by looking to the relations amongst the classes of objects which they study: thus geometry is a more general science than biology because all living objects are, presumably, also spatial objects, while most spatial objects are not alive. Accordingly, geometrical truths have wider application than those of biology. On this Aristotelian model, some sciences are said to fall within others, in which case the truths of the more general science apply to the objects of the less general, or included, science. In the *Posterior Analytics* Aristotle makes some remarks to this effect which are specifically relevant to the mathematical sciences. He writes:

> [Y]ou cannot prove anything by crossing from another kind—e.g. something geometrical by arithmetic. There are three things involved in demonstrations: one, what is being demonstrated, or the conclusion (this is what holds of some kind in itself); two, the axioms (axioms are the items from which the demonstrations proceed); third, the underlying kind whose attributes—i.e. the items incidental to it in itself—the demonstrations make plain. Now the items from which the demonstrations proceed may be the same; but where the kinds are different, as with arithmetic and geometry, you cannot attach arithmetical demonstrations to what is incidental to magnitudes—unless [the] magnitudes are numbers. [...] Arithmetical demonstrations always contain the kind with which the demonstrations are concerned, and so too do all other demonstrations. Hence the kind must be the same, either *simpliciter* or in some respect, if a demonstration is to cross. [...] For this reason you cannot prove by geometry that there is a single science of contraries, nor even that [the product of two cubes makes] a cube. (Nor can you prove by any other science what pertains to a different science, except when they are so related to one another that the one falls under the other—as e.g. optics is related to geometry and harmonics to arithmetic.) Nor indeed anything that holds of lines not as lines and as depending on the principles proper to them—e.g. whether straight lines are the most beautiful of lines, or whether they are contrarily related to curved lines; for these things hold of lines not in virtue of their proper kind but rather in virtue of something common.[2]

From Aristotle's remarks here we can gather some understanding of the traditional model of science which was, by and large, the tacit background for geometers working with the traditional, Euclidean view of geometry. Sciences are delineated by the genus of objects about which they provide demonstrations. The logical relations amongst distinct sciences are determined by the relations between the classes of objects studied by those sciences; if the classes of objects are entirely disjoint (i.e., if there is no object which is treated in both sciences), then there will be no noteworthy relations amongst the sciences which study those objects, other than mutual exclusion. Various other relations are possible when there is overlap or inclusion amongst the classes of objects studied. Further, the demonstrations of distinct sciences cannot be applied to one another, except under the aforementioned condition, because any given science, qua science, may only contain reasoning pertaining to its proper objects—to 'cross to a different kind' is, in essence, to fail to reason scientifically.

[2] [Aristotle 1994, Book I, Chapter 7, 75a37-75b20; translation slightly altered].

The emergence of consistent non-Euclidean geometries and the development of sophisticated methods of classifying those geometries (e.g., Klein's *Erlanger Programm* and Riemann's differential approach) considerably loosened the link between spatial intuition and geometrical truth. Along with the slackened connection between intuition and geometrical demonstration came a general uncertainty about the *objects* of geometry itself. Traditionally, geometry had been construed as the science of the (idealized) elements of physical space. So, when the intuition which was the putative source of our knowledge of this space and its elements fell into disrepute, and the concept of space mutated into something considerably more general, it should be no surprise that there arose some confusion over the very nature of the science of geometry itself. If geometry could not be characterized as the science of spatial objects, then what sort of a science was it? What genus of objects did it study? What was geometry *about*? As we will see, the new approach to geometry offered by Hilbert and others—an approach which contained an alternative understanding of the nature of definitions and axioms—required a rather different understanding of the nature of (at least 'formal') science than was available within the Aristotelian tradition.

## §1. Hilbert's *Grundlagen*

Formally at least, Hilbert's *Grundlagen der Geometrie* bears some similarity to Euclid's *Elements*. Within its pages we find a number of definitions and a collection of unproved propositions employed to prove theorems, many of which are familiar (at least in name) from the history of geometry. Like many mathematicians before him, Hilbert does away with Euclid's distinction between common notions and postulates and instead speaks of all non-definitional, unproved propositions simply as axioms. In addition to individual axioms Hilbert also speaks of 'axiom groups', which are collections of axioms grouped together according to their specific subject matter. In the *Grundlagen*, there are five such groups, which deal with incidence, order, congruence, parallels, and continuity, respectively. But, despite these formal resemblances between Hilbert's work and that of Euclid, a rather wide conceptual gulf separates their respective uses of terms like 'point,' 'line,' 'plane,' 'between,' 'definition' and 'postulate' (or 'axiom').

### 1.1 Definitions

In order to see exactly how wide this divergence is, it is instructive to compare Euclid's definition of 'point' with that of Hilbert. Recall that Euclid defines a point as "that which has no part".[3] Hilbert, like Euclid, presents a number of definitions, some more in line with traditional Euclidean definitions, and some which seem quite different. This is partially due to the novelty of his axiomatic approach and partly due to confusions regarding the nature of definitions. Take his first attempt at a definition of the term 'point' for instance:

> DEFINITION. Consider three distinct sets of objects. Let the objects of the **first** set be called *points* and be denoted by *A, B, C, ...*; let the objects of the **second** set be called *lines* and be denoted by *a, b, c, ...*; let the objects of the **third** set be called *planes* and

---

[3] [Heath 1956, 155].

> be denoted by $\alpha$, $\beta$, $\gamma$, .... The points are also called the *elements of line geometry*; the points and the lines are called the *elements of plane geometry*; and the points, the lines and the planes are called the *elements of space geometry* or the *elements of space*.[4]

Now, postponing for the moment the question whether or not this paragraph forms Hilbert's complete definition of 'point', 'line', or 'plane', it is already apparent that his approach is quite different from that of Euclid. Rather than an attempt to fix our attention definitively upon a certain type of object (e.g., partless wholes), Hilbert begins his definition by situating points, lines, and planes within three geometric systems. There is no real attempt at reference-fixing here; instead what we have is a very loose indication of the relations in which the basic items of the system stand, or, perhaps more accurately, the bare indication *that* these items stand in some relations to certain systems and to each other. The pre-systematic, introductory role of the definition seems to have disappeared entirely. This is the case for Hilbert's definitions of other basic concepts and relations as well. He defines 'betweenness' for example, as follows: "The points of a line stand in a certain relation to each other and for its description the word "between" will be specifically used".[5] Now, taken on its own, this definition is hardly satisfactory. And, indeed, Hilbert does not seem to have wanted it, or any of his definitions, to be taken in isolation. Instead, he suggests that these concepts and their relations are to be further explicated by means of his axiom groups. Thus after (provisionally) defining 'point', 'line', and 'plane' by vague reference to the geometric systems which they form together, he suggests that his broad initial account will be narrowed through the axiomatic characterization of the *relations* in which these primitive concepts stand to each other:

> The points, lines and planes are considered to have certain mutual relations and these relations are denoted by words like "**lie**," "**between**," "**congruent**." The precise and mathematically complete description of these relations follows from the **axioms of geometry**.[6]

Here, again, we can see the difference between the traditional, Euclidean view of geometry (as a collection of intuitively grounded truths about idealized elements of physical space) and Hilbert's view. Rather than unprovable truths about the basic objects of geometry, the axioms in Hilbert's sense play a quasi-definitional role in which they help to determine the relations amongst the basic elements of certain geometrical systems; the nature of those elements beyond the relations in which they stand to each other seems to be a matter of no intrinsic geometric interest. As we've seen, within the traditional Euclidean dispensation, one type of definition played a reference-fixing role *prior* to the introduction of the axioms, and these axioms were understood to be unprovable truths about the objects toward which the definitions guided the reader. Within Hilbert's work, by contrast, the initial definitions of 'line', 'point', and 'plane' do not fix (or even attempt to fix) the references of these terms. Rather, they serve to indicate, in an extremely general way, that the objects stand in certain relations to each other. Neither do these initial definitions themselves determine specifically

[4] [Hilbert 1971, 3]; [Hilbert 1903, 2].
[5] [Hilbert 1971, 5]; [Hilbert 1903, 4] .
[6] [Hilbert 1971, 3]; [Hilbert 1903, 2].

which relations these are. It is instead left up to the subsequent axiom groups to 'fix' the reference of these terms and the relations in which their referents stand more definitively—though, as we will see, it is precisely the free-floating character of Hilbert's points, lines, planes, etc. which represents his biggest departure from the traditional view, so that, even later in the work, he is not terribly interested in pinning them down.

It is for these reasons that Hilbert's approach is remarkably different than the traditional Euclidean approach. While many traditional geometers considered the intention of their work to be the elaboration of a science of space and spatial objects, it is very difficult to construe Hilbert's work in this way. Instead of an emphasis on the objects of geometry, which are intuitively graspable according to the traditional view, Hilbert emphasizes the *relations* in which these objects might stand (though, as we will see, speaking of a determinate group of objects is, perhaps, not an ideal description of Hilbert's intention). This shift in emphasis had quite radical implications for the conception of geometry as a science.

1.2 Axioms and intuition

Despite this divergence from tradition regarding the definition of the primitive, intuitively given objects of geometry, Hilbert held on to some aspects of the Euclidean tradition. For instance, he still believed that a proper set of geometrical axioms (like his own) ought to be developed on the basis of our intuition of space. In this he resembled his colleague Klein, as both men favoured a highly abstract approach to geometry but remained insistent that, at bottom, geometry was the science of space (however expansive their notion of space), and that the ultimate properties of space were graspable by way of intuition. Indeed, more than having their certainty grounded in our intuition, Hilbert insisted on several occasions that any worthwhile mathematical investigations should have their origin in empirical concerns. At the outset of the *Grundlagen* he writes, for instance, that

> Geometry, like arithmetic, requires only a few and simple principles for its logical development. These principles are called the **axioms** of geometry. The establishment of the axioms of geometry and the investigation of their relationships is a problem which has been treated in many excellent works of the mathematical literature since the time of Euclid. This problem is equivalent to the logical analysis of our intuition of space.[7]

Thus, for Hilbert the basic material of geometry, given in its axioms, has its origin in our intuition of space. The remainder (i.e., the collection of theorems which are not axioms) stems from the logical analysis of this intuition, particularly, we may assume, from the analysis of the different aspects of this intuition given in Hilbert's axiom groups. He writes that "[e]ach of these groups expresses certain related facts basic to our intuition".[8] The axioms, then, are not arbitrarily chosen, with indifference to our intuition. Instead, they were intended by Hilbert to capture important features of spatial intuition with an eye toward the

---

[7] [Hilbert 1971, 2]; [Hilbert 1903, 1].
[8] [Hilbert 1971, 3]; [Hilbert 1903, 2].

development of a geometry which ought to be capable of delivering the appropriate set of deductive consequences implied by that intuition.[9] It is only at the stage of logical analysis—after intuition has been crystallized within the axiom groups—that the purely deductive, non-intuitive work so characteristic of 'modern' mathematics begins. Hilbert quite clearly outlines his intention to capture the important implications of our geometrical intuition of space in an 1899 letter to Frege:

> It was of necessity that I had to set up my axiomatic system: I wanted to make it possible to understand those geometrical propositions that I regard as the most important results of geometrical inquiries: that the parallel axiom is not a consequence of the other axioms, and similarly Archimedes' axiom, etc. I wanted to answer the question whether it is possible to prove the proposition that in two identical rectangles with an identical base line the sides must also be identical, or whether as in Euclid this proposition is a new postulate. I wanted to make it possible to understand and answer such questions as why the sum of the angles in a triangle is equal to two right angles and how this fact is connected with the parallel axiom. That my system of axioms allows one to answer such questions in a very definite manner, and that the answers to many of these questions are very surprising and even quite unexpected, is shown, I believe, by my *Festschrift* as well as by the writings of my students who have followed it up.[10]

We can see here that, instead of taking the content of Euclidean geometry as somehow originating in a set of arbitrary postulates, Hilbert views his axiom groups themselves as the components of a successful attempt to capture some pre-systematic intuitive content. In this case, the content is delivered by a history stemming from two millennia of geometric research, itself grounded (or at least thought to have been grounded) in our intuitive apprehension of space.[11]

The most novel elements of Hilbert's approach to the axiomatization of geometry had important precursors at least as far back as Lambert. One element of the *Grundlagen* which separates it from most prior work in geometry is Hilbert's insistence that the deductive

---

[9] Hilbert remarks, for instance, that "Axiom I, 7 expresses the fact that space has no more than three dimensions, whereas Axiom I, 8 expresses the fact that space has no less than three dimensions" [Hilbert 1971, 4]. Note that Hilbert writes simply 'space' and not 'Euclidean space'.

[10] [Hilbert to Frege, 29 December 1899; excerpt by Frege in Frege 1980, 38-39]; the final sentence refers, particularly, to the dissertation of Hilbert's student Max Dehn (1878-1952) on Legendre's theorem, subsequently published as [Dehn 1900].

[11] An illuminating account of Hilbert's understanding of the 'life cycle' of mathematical theories is given in [Corry 2004a]; Corry writes there that Hilbert distinguishes "three clearly separated stages that, in his view, mathematical theories usually undergo in the course of their development: the naive, the formal, and the critical" and that "For Hilbert, the need to introduce this method [i.e., the axiomatic method] arises precisely as a means to analyze already established theories, to criticize their basic assumptions, and to elucidate their logical deductive structure" [Corry 2004a, 20]. Given this view of the evolution of mathematical theories, it seems likely that Hilbert saw his own axiomatic endeavours as attempts to likewise clarify already-going concerns, rather than as efforts to investigate arbitrary domains of inquiry freshly sprung from the head of a mathematical Zeus. (Corry seems to suggest as much at [Corry 2004a, 93-94]). This approach to mathematical axiomatics is, accordingly, continuous with Hilbert's later attempts to axiomatize physics.

structure developed on the basis of intuitively given axiom groups does not require any *further* input from intuition once these groups are established. But as early as 1766 we find a remarkably similar view expressed by Lambert in a work on the theory of parallels.[12] There he writes the following rather prescient paragraph:

> The difficulties concerning Euclid's 11th axiom [i.e., the parallel postulate] have essentially to do only with the following question: Can this axiom be derived correctly from Euclid's postulates and the remaining axioms? Or, if these premises are not sufficient, can we produce other postulates or axioms, no less evident than Euclid's, from which his 11th axiom can be derived? In dealing with the first part of this question we may wholly ignore what I have called the representation of the subject-matter. Since Euclid's postulates and remaining axioms are stated in words, we can and should demand that no appeal be made anywhere in the proof to the matter itself, but that the proof be carried out—if it is at all possible—in a thoroughly symbolic fashion. In this respect, Euclid's postulates are, so to speak, like so many given algebraic equations, from which we must obtain x, y, z, etc. without ever looking back to the matter in discussion. Since the postulates are not quite such formulae, we can allow the drawing of a figure as a guiding thread to direct the proof. On the other hand, it would be preposterous to forbid consideration and representation of the subject-matter in the second part of the question, and to require that the new postulates and axioms be found without reflecting on their subject-matter, off the cuff, so to speak.[13]

As we can see here, Lambert, like Hilbert, paid due homage to the understanding of geometry as founded upon some specific content (e.g., the deliverances of spatial intuition). He also insisted that much of our subsequent reasoning about geometrical relations can and should be divorced from any intuitive or specifically geometrical content. We should think entirely abstractly, without reference to intuition, about, e.g., relations of independence between Euclid’s axioms.

Lambert suggests we can treat these axioms as mere symbolic formulae to be manipulated using acceptable techniques, just as we manipulate algebraic equations by performing allowable operations upon them, very often with indifference to the possible references of their terms. This view is, in germinal form, quite similar to Hilbert’s later treatment of proofs as mathematical objects in their own right. For Lambert, intuition plays an important role in the development of basic geometrical axioms. Once these axioms are established on the basis of intuition, however, geometry can and should be treated as an entirely abstract science, decoupled from reference to spatial intuition except at its basis.

A little over a century later, we find a more direct influence on Hilbert in the views of Moritz Pasch (1843-1930), particularly as expressed in his 1882 *Vorlesungen über neuere*

[12] Though completed by 1766, the work was only published in 1786.
[13] [Lambert 1786, quoted in Torretti 1984, 49].

*Geometrie*.[14] Pasch, like Lambert before him, held that geometry's "foundational concepts and propositions should be derived immediately from experience," but made the slightly different claim that "all other concepts and results must be reducible to these alone. ... [G]eometry should thus proceed from the beginning as a descriptive science of the world of appearances and even in its most complex conceptions it should retain this character."[15] Thus, while Lambert was content to treat axioms as the intuitive basis for a kind of symbolic geometric calculus which itself needn't be geometrically meaningful, Pasch sought a firmer experiential basis for geometry by insisting on the meaningfulness of both axioms and each step in any chain of geometric reasoning. This does not, however, mean that Pasch eschewed the abstract algebraic techniques hinted at by Lambert. He writes, for instance that

> If geometry is to be truly deductive, the process of inference must be independent in all its parts from the meaning of the geometrical concepts, just as it must be independent of the diagrams; only the relations specified in the propositions and definitions may legitimately be taken into account. During the deduction it is useful and legitimate, but in *no way necessary*, to think of the meanings of the terms; in fact, if it is necessary to do so, the inadequacy of the proof is made manifest. If, however, a theorem is rigorously derived from a set of propositions [...] the deduction has value which goes beyond its original purpose. For if, on replacing the geometric terms in the basic set of propositions by certain other terms true propositions are obtained, then corresponding replacements may be made in the theorem; in this way we obtain new theorems ... without having to repeat the proof.[16]

Here Pasch develops a conception of the deductive structure of geometric reasoning which is quite similar to Hilbert's later view. He suggests that, while we *may* employ intuitive thinking and diagrams when working our way through a geometric proof, if such devices are necessary the proof is insufficiently rigorous. Pasch's full view, then, is that the axioms of geometry must be grounded in our intuition of space, and that each subsequent argumentative step towards a given proof must itself be meaningfully construed on the basis of these axioms. But, despite this stepwise connection to intuition, the fully rigorous expression of geometric proof—if it is to be deductively valid—should not rely in any necessary way on intuitive reasoning or diagrammatic representation. So, while the meaningfulness of geometric formulae is guaranteed by their connection to intuitively given axioms, the deductive manipulation of those formulae ought to be performed in an algebraic manner with complete indifference to meaning. Thus, for both Lambert and Pasch, the deductive element of geometric reasoning is in important respects separable from the intuitive element required to establish a proper axiom set.

Hilbert's view—developed after the advent of differential and various non-Euclidean geometries—is almost identical with Pasch's and Lambert's in these respects. For Hilbert,

---

[14] For accounts of Pasch's work and its influence on subsequent work in geometry, see [Torretti 1984, 210-218] and [Corry 2004b, 155ff.]. For Pasch's influence specifically on Hilbert, [Corry 2004b, 157-159] and [Toepell 1986] are useful.
[15] [Pasch 1887, 129; translation in Rowe 2000, 72].
[16] [Pasch 1882, 98].

too, once the intuitive basis is given, the theorems of geometry can be deductively unfurled without ever returning to their intuitive ground. His proofs are all conducted (he would argue) using purely logical, non-intuitive means. None of these proofs should rely essentially upon a feature of a diagram or upon our accepting intuitions other than those formulated in the axiom groups. This view, exemplified in the *Grundlagen*, is more explicitly and generally stated in his well-known lecture on the most important mathematical problems for the twentieth century. There he states the following:

> To new concepts correspond, necessarily, new signs. These we choose in such a way that they remind us of the phenomena which were the occasion for the formation of the new concepts. So the geometrical figures are signs or mnemonic symbols of space intuition and are used as such by all mathematicians. Who does not always use along with the double inequality $a > b > c$ the picture of three points following one another on a straight line as the geometrical picture of the idea 'between'? Who does not make use of drawings of segments and rectangles enclosed in one another, when it is required to prove with perfect rigour a difficult theorem on the continuity of functions or the existence of points of condensation? [...] The arithmetical symbols are written diagrams and the geometrical figures are graphic formulae; and no mathematician could spare these graphic formulae, any more than in calculation the insertion and removal of parentheses or the use of other analytical signs. The use of geometrical signs as a means of strict proof presupposes the exact knowledge and complete mastery of the axioms which underlie those figures and in order that these geometrical figures may be incorporated in the general treasure of the mathematical signs, there is necessary a rigorous axiomatic investigation of their conceptual content. Just as in adding two numbers, one must place the digits under each other in the right order, so that only the rules of calculation, i.e., the axioms of arithmetic, determine the correct use of the digits, so the use of geometrical signs is determined by the axioms of geometrical concepts and their combinations.[17]

Beyond the clear similarity to Pasch's views expressed here, it is illuminating that Hilbert does not restrict himself to geometric examples. This is because he does *not* make any strict distinction between the type of work involved in geometric reasoning and that involved in, e.g., arithmetic or algebraic reasoning. There is for Hilbert a basic mathematical unity which is best expressed by the wide applicability of the axiomatic method to branches of mathematics which are, from the point of view of the 'objects' studied, quite distinct. We can establish Euclidean geometry axiomatically and prove its theorems purely deductively.

Here it is made clear by Hilbert that, while we *may* employ aids such as diagrams and pictures, even in absolutely strict proofs, our use of these sorts of aids must be licensed by the axiomatic presentation of the mathematical system in question. Thus, for instance, we may employ a picture of three points on a straight line to represent the notion of 'betweenness' in a strict proof only insofar as the intended conceptual content of the picture can be explicated in terms of the axioms of the system. The meaning of a specific theorem, proposition, or diagram formulated from within an axiomatized system is guaranteed by the

[17] [Hilbert 1900a, excerpt translated in Ewald 1999, 1100-1101].

intuitively grounded axioms themselves. But, just as with Pasch, the actual deductive linkage of propositions in geometric and other proofs should be treated with indifference to the epistemologically useful meanings of the propositions at issue. For Hilbert, it is only once we have established a given theory on an axiomatic basis that we can begin to speak of truly rigorous proof; in the case of non-axiomatic theories, our use of intuitive reasoning or diagrammatic representation cannot be guaranteed by an axiomatic basis, and can only very rarely be dispensed with. Without the exact knowledge engendered by the process of axiomatization, we cannot engage in the purely symbolic form of deductive reasoning because the meaningfulness of the steps in this formal reasoning cannot be guaranteed, and we may very well stray into absurdity or imprecision.

Though it seems clear from what we have just seen that Hilbert himself never held the purely postulational view of mathematics as the investigation of deductive structures derived from arbitrarily selected postulate sets, the way in which he developed his axiomatic approach to geometry was (largely) consistent with such an approach. Some of his remarks, particularly in his correspondence with Frege, certainly lead one in this direction, so that it is no surprise that such a strong opponent of this sort of project as Frege would have reacted accordingly.

But it was not only his critic Frege who read Hilbert in this manner; many of Hilbert's admirers within the mathematical community (e.g., Felix Hausdorff, and the school of American mathematicians surrounding E. H. Moore in Chicago and Edward Huntington at Harvard) developed their own approaches to axiomatics inspired by Hilbert's work, approaches which were avowedly 'arbitrary' in a way that Hilbert himself continued to resist throughout his career.[18] Hausdorff, inspired in large part by Hilbert, came very close to a purely arbitrary understanding of the nature of axiomatics. In the context of Hilbert's approach to geometry, Hausdorff wrote that

> In all philosophical debates since Kant, mathematics, or at least geometry, has always been treated as heteronomous, as dependent on some external instance of what we would call, for want of a better term, intuition, be it empirical, subjective or scientifically amended, innate or acquired. The most fundamental task of modern mathematics has been to set itself free from this dependence, to fight its way through from heteronomy to autonomy.[19]

The importance of completely detaching mathematics from intuition is further emphasized by Hausdorff in course notes from the same period: "Mathematics totally disregards the actual significance conveyed to its concepts, the actual validity that one can accord to its theorems. Its indefinable concepts are arbitrarily chosen objects of thought and its axioms are arbitrarily, albeit consistently, chosen relations among these objects. Mathematics is a science of pure thought, exactly like logic".[20] So, here we have Hausdorff, inspired by the *Grundlagen*, extending Hilbert's own view of deduction as indifferent to meaning to include the axiomatic basis of a theory as well. Hausdorff attempts to break entirely from the

---

[18] See, e.g., [Huntington 1902] and [Moore 1902b].
[19] [Hausdorff c. 1904, quoted in Corry 2004a, 117].
[20] [Hausdorff 1903-1904, quoted in Corry 2004a, 117].

Kantian understanding of ‘heteronomous’ mathematics tethered to pure intuitions of space and time by suggesting we examine *arbitrarily* developed postulate sets with no connection to intuition whatsoever. On this view, geometry, for instance, would cease to be coupled to spatial intuition and, instead, become the systematic exploration of a wide variety of structures of a certain type—perhaps those suggested by Riemann’s coefficient-determined classificatory scheme. (Though it is not possible here, a more detailed examination of Hausdorff’s influence on topology and measure theory would illustrate how powerful and widespread this view would later become in important branches of geometry).

While Hausdorff’s views seem at least partially opposed to the intuitively-grounded view developed by Lambert, Pasch, and Hilbert, some of Hilbert’s own remarks (particularly those made in correspondence to Frege) directly support this purely postulational approach to mathematics. Indeed, in one of the better known paragraphs of their correspondence, Hilbert writes to Frege that “if the *arbitrarily* given axioms do not contradict one another with all their consequences, then they are true and the things defined by the axioms exist. This is for me the criterion of truth and existence”.[21] Yet, elsewhere, Hilbert seems committed to the view that his axioms are anything but arbitrary—for to investigate arbitrarily postulated axioms without being guided by a pre-theoretic, pre-axiomatic intuition was, for Hilbert, a pointless exercise, and mathematics was certainly not meaningless for him, even in the later stages of his career when he began developing his well-known formalist programme. He writes later, for instance, that “We are not speaking here of arbitrariness in any sense. Mathematics is not like a game whose tasks are determined by arbitrarily stipulated rules. Rather, it is a conceptual system possessing internal necessity that can only be so and by no means otherwise.”[22] For Hilbert, mathematics is different from a merely formal game insofar as the *rules* of mathematics cannot be altered in the way that we might alter the rules of a game. For, in constructing a game, we can alter the rules however we like, even altering the nature of the ‘logical’ relations amongst the rules and components of the game. But, for Hilbert, once we have established some elements of a conceptual system (the axioms), the unfolding of the rest (the theorems) is a matter of necessity, not happenstance. Though by the time of the publication of the *Grundlagen* Hilbert had not yet developed a sophisticated picture of the nature of logic and its relation to mathematical reasoning, it is clear that he did not understand logic as a merely formal enterprise, but instead as an important and meaningful constraint to which mathematical thinking must answer.

## 1.3 Axioms, truth, and existence

Within the traditional understanding of Euclidean geometry, Euclid’s axioms (barring, perhaps, the troublesome parallel postulate) were thought to be basic, unprovable truths about physical space. As we’ve just seen, Hilbert’s understanding of the role of his axioms was partly traditional and partly novel. As within the traditional view, the content of Hilbert’s axiom groups was thought to be given by spatial intuition. Hilbert’s axioms, too, are unproven expressions of intuition on the basis of which a whole deductive structure is

[21] [Hilbert to Frege, 29 December 1899; excerpt by Frege in Frege 1980, 39-40].
[22] [Hilbert 1919-20 [1992], 14; quoted in Corry 2006, 138].

developed. But, in contrast to the tradition, the *truth* of his axioms is not guaranteed by the nature of physical space (though the content of the axioms is suggested by our intuition of physical space). Instead, their truth is held to result from their consistency with each other. While on the traditional dispensation the truth of axioms was not relativized to a particular axiomatic system, on Hilbert's view, we cannot speak of the truth of mathematical axioms except by reference to their mutual consistency.

In addition to the truth of the axioms, consistency also establishes for Hilbert the *existence* of the objects to which the axioms refer. Thus, for instance, a geometric point is said to exist if the axioms which describe points (and the relations between points and other elements of the axiomatic system) are consistent. This, perhaps even more than Hilbert's relativized conception of mathematical truth, is a radical claim. It shows that Hilbert, at this time, was willing to countenance the existence of an almost bewildering array of mathematical objects, with the only condition being the possibility of providing a consistent axiomatization describing the objects. Thus, for Hilbert, the realm of the mathematically existent is coextensive with the realm of the logically possible.[23] The exact relationship between mathematical existence and the consistency of axiom sets was not fully worked out in Hilbert's work of this time—a fact which became clearer in Hilbert's own later work on direct consistency proofs, and in Frege's criticisms of the *Grundlagen der Geometrie*.

Hilbert's views on mathematical existence and truth have a wide range of consequences. One of the most important is that the proof of the consistency of a set of axioms becomes a crucial question in the development of an acceptable axiomatic system. Within the traditional dispensation, the consistency of Euclid's axioms (again, excluding the parallel postulate) was, for the most part, not a matter for intensive geometric research. Since the axioms were understood to be truths based upon spatial intuition, the idea that the consistency of the axioms might require any special investigation simply did not occur. It didn't make much sense to ask whether a collection of obvious truths could all be true together—their being independently true of the world was a tacit guarantee that they were true when taken together, too. On Hilbert's view, however, the grounding role of intuition was more heuristic in character: our intuition suggests certain axioms, but their truth is not at all grounded in that intuition. Instead, once we have established some collection of axioms using our intuitive heuristic, we must investigate and, ideally, prove their consistency to justify their truth. Intuition on its own is insufficiently rigorous and requires precise and careful axiomatization, followed by the deductive elaboration of consequences. This alteration in the concept of mathematical truth goes hand in hand with Hilbert's shift away from an object-oriented approach to geometry. In contrast to geometrical truth understood as a pre-structural, absolute property determined by the nature of space and certain idealized

---

[23] Bolzano, in a seemingly offhand remark, offers a remarkably similar view as early as 1837. In §352 of the *Wissenschaftslehre*, while discussing how we might tell if an idea is objectual or not (i.e., whether any objects fall under it), he writes that "In order to determine [if an idea is objectual], we must also decide whether the attributes that the idea would ascribe to its object presuppose actual existence or not. If not, the idea is to be declared objectual. For if its object need not be something existing, it is obvious that such an object could only be lacking if the assumption that there were such an object contradicted a purely conceptual truth" [Bolzano 1837, §352]. In effect, Bolzano claims here that particular abstract objects 'exist' (in a restricted sense) just in case their existence is not contrary to a conceptual truth.

spatial objects, Hilbert turns to a notion of truth relative to a given axiomatization, determined by relations of consistency. This interest in questions of consistency was part of a larger shift in interest towards certain structural properties of axiomatic systems. In addition to consistency, Hilbert began to focus on other properties of axiomatic systems, particularly their simplicity, completeness, and the mutual independence of their axioms.

### 1.4 Desirable properties of axiomatic systems

One of the more innovative elements of Hilbert's axiomatic approach was his specific focus on relations amongst the axioms themselves. Within the Euclidean tradition, as we saw, there were some concerns about the independence of the axioms relative to each other. In particular, the potential dependence of the parallel postulate on the other axioms was a matter of intensive investigation. With Hilbert, this concern for the independence of his axiom groups became a matter for rigorous logical investigation. Alongside independence and consistency, Hilbert was also concerned that a suitable axiomatization be complete, simple, and (if possible) categorical. He writes at the outset of the *Grundlagen*, for instance, that "This present investigation is a new attempt to establish for geometry a complete, and as simple as possible, system of axioms and to deduce from them the most important theorems in such a way that the significance of the various groups of axioms, as well as the scope of the conclusions that can be drawn from the individual axioms, become as clear as possible".[24] Though these properties have since become the object of intense study throughout mathematics, formal semantics, and logic, at Hilbert's time the precise formulation of criteria for consistency, independence, completeness, categoricity, and simplicity had not yet been given. I will now turn to an examination of Hilbert's understanding of these notions around the time of the publication of the *Grundlagen der Geometrie*.[25]

#### 1.4.1 Consistency

As we've seen, for Hilbert the consistency of a set of axioms (or axiom groups) was of central importance for the establishment of the truth of a set of axioms, as well as the existence of the objects referred to by the axioms. Accordingly, proving the consistency of a set of axioms becomes a matter for direct mathematical investigation. The consistency of a collection of axioms, for Hilbert, amounts to the impossibility of deriving a contradiction from the consequences of those axioms taken together.[26] Although his conception of

---

[24] [Hilbert 1971, 2]; "Die vorliegende Untersuchung ist ein neuer Versuch, für die Geometrie ein vollständiges und möglichst einfaches System von Axiomen aufzustellen und aus denselben die wichtigsten geometrischen Sätze in der Weise abzuleiten, daß dabei die Bedeutung der verschiedenen Axiomgruppen und die Tragweite der aus den einzelnen Axiomen zu ziehenden Folgerungen möglichst klar zu Tage tritt" [Hilbert 1903, 1].

[25] Though Hilbert was instrumental in the subsequent development of proof theory and mathematical logic throughout the 1920s and 1930s, I limit my discussion here to his early work insofar as my interest is in highlighting the differences between Hilbert's early position and that of Frege. In his later work, Hilbert in fact took a number of Frege's criticisms to heart, so that the gap between Hilbert's highly structure-focussed position and Frege's point of view narrowed considerably.

[26] "The necessary task then arises of showing the *consistency* and *completeness* of these axioms, i.e. it must be proved that the application of the given axioms can never lead to contradictions, and, further that the system of

consistency is clearly proof-theoretic, his method of proving this sort of consistency is quite similar to the modern model-theoretic approach: an interpretation in which all the axioms hold (i.e., a model of those axioms) is constructed and this construction is taken to directly show the consistency of the axioms. Let us take a brief look at the interpretation Hilbert employs to prove the consistency of his particular set of axioms for Euclidean geometry.

Hilbert begins by interpreting his axioms within "the field Ω of all algebraic numbers that arise from the number 1 and the application of a finite number of times of the four arithmetical operations of addition, subtraction, multiplication and division, and the fifth operation $\sqrt{|1+\ \omega^2|}$, where ω denotes a number that results from these five operations".[27] Points are interpreted as pairs of numbers (x, y) within this field; lines are interpreted as "the ratios ($u$ : $v$ : $w$) of any three numbers from Ω [...] provided $u$, $v$ are not both zero."[28] Now, having thus interpreted the basic notions of his formalization of Euclidean geometry within the algebraic number field Ω, Hilbert proceeds to suggest that his axioms up to and including Axiom V, 1 are satisfied. The remaining axiom, the so-called completeness axiom V, 2[29] is satisfied, he claims, *only* if we employ the specific field of the real numbers in our interpretation in place of the subfield Ω. Having thus interpreted the basic notions of Euclidean geometry within the domain of the real numbers (and, initially, within the field Ω), Hilbert goes on to suggest that each of his axioms, thus interpreted, still holds. The first model of his axiomatization interpreted within the field Ω illustrates that his axioms from I-IV and including V, 1 are consistent with each other precisely because "Every contradiction in the consequences of [these axioms] would therefore have to be detectable in the arithmetic of the field Ω."[30] It is assumed, at this point, that no contradictions arise within the arithmetic of this field; likewise, it is assumed that no contradictions can arise within the axioms thus modelled. When Hilbert extends his model to the field of all real numbers, the claim is similar: his axioms for Euclidean geometry are shown to be consistent upon the assumption that no contradiction arises within the arithmetic of the real numbers. Today we see this immediately as an example of a relative consistency proof: the consistency of Euclidean geometry is proved only relative to the consistency of the real numbers. By the turn of the century, however, the forbidding difficulties of providing direct consistency proofs for

axioms is adequate to prove all geometrical propositions. We shall call this procedure of investigation the *axiomatic method*" [Hilbert 1900b, in Ewald 1999, 1093].

[27] [Hilbert 1971, 29].

[28] [Hilbert 1971, 29].

[29] This axiom was not included in Hilbert's *Festschrift* version of the *Grundlagen*, nor in the first edition, but appeared only in the first French edition and second German edition. Subsequent to those editions it was further narrowed and renamed the axiom of *line* completeness, and altered to read as follows: "An extension of a set of points on a line with its order and congruence relations that would preserve the relations existing among the original elements as well as the fundamental properties of line order and congruence that follows from Axioms I-III, and from V, 1 is impossible" [Hilbert 1971, 26]; in the 1903 edition it is phrased as follows: "Die Elemente (Punkte, Geraden, Ebenen) der Geometrie bilden ein System von Dingen, welches bei Aufrechterhaltung sämtlicher genannten Axiome keiner Erweiterung mehr fähig ist, d.h.: zu dem System der Punkte, Geraden, Ebenen ist es nicht möglich, ein anderes System von Dingen hinzuzufügen, so daß in dem durch Zusammensetzung entstehenden System sämtliche aufgeführten Axiome I-IV, V 1 erfüllt sind" [Hilbert 1903, 16]. We will have occasion to return to this axiom in the sections on completeness and categoricity below.

[30] [Hilbert 1971, 30].

axiomatized versions of arithmetic were not yet apparent to Hilbert, as is evident in his statement that he was "convinced that it must be possible to find a direct proof for the consistency of the arithmetical axioms, by means of a careful study and suitable modification of the known methods of reasoning in the theory of the real numbers".[31] The relativity of Hilbert's consistency proofs would, as we shall see, form one of the focal points of Frege's criticism, and would also become one of the chief motivators for Hilbert's formalist program of the 1920s.

1.4.2 Independence

After establishing the consistency of his axioms to his satisfaction, Hilbert moves on in the *Grundlagen* to prove the independence of some of those axioms from the others. He begins, unsurprisingly, with an illustration of the independence of his parallel axiom group. He opens this section of the work as follows: "Having seen the [consistency] of the axioms it is of interest to investigate whether they are all independent of each other. In fact it can be shown that no essential part of any one of these groups of axioms can be deduced from the others by logical inference".[32] At this time in his career, Hilbert (like most mathematicians of the time) had not expended much energy attempting to flesh out any systematic conception of basic logical notions like inference or consequence. He had, however, taken at least some interest in the quasi-logical work in elementary set-theory undertaken by Cantor and Dedekind. One thing that is quite clear about his conception of logic at the time, though, is that he, like Frege, took it to flow from a source of knowledge distinct from intuition. When we find the early Hilbert discussing things in a 'purely logical' manner, it is in order to draw an epistemological contrast between knowledge derived from (usually spatial) intuition and knowledge given without reference either to intuition or experience. It is only given this epistemological distinction that Hilbert can speak of the *logical* analysis of our intuition of space, and maintain the distinction between the intuitions which inform our choice of axioms and the subsequent logical work which gives us the deductive structure erected on the basis of those axioms. Thus, the independence of axioms from one another ought to be provable without reference to spatial intuition, or other forms of experiential knowledge, as suggested by both Lambert and Pasch. Let's take a brief look at the way in which Hilbert goes about proving the independence of his parallel axiom group from the remaining groups.

His procedure again seems to involve what we would now understand to be model-theoretic techniques. Having already provided a model in which all of his axioms excepting the completeness axiom hold (by interpreting his axioms and basic notions in terms of the field $\Omega$, or in the field of the real numbers), Hilbert now constructs a model in which all of his axioms except the parallel axiom are satisfied. For our purposes here, the details of this

[31] [Hilbert 1900b, in Ewald 1999, 1104]. In another paper of the same year on the concept of number, Hilbert writes, again optimistically, that "To prove the consistency of the above axioms [for arithmetic] one needs only a suitable modification of familiar methods of inference. In this proof I also see the proof of the existence of the totality of real numbers, or—in the terminology of G. Cantor—the proof that the system of real numbers is a consistent (finished) set" [Hilbert 1900a, in Ewald 1999, 1095]. A suitable modification and the resulting proof were both to remain elusive and, perhaps, impossible after Gödel's incompleteness results.

[32] [Hilbert 1971, 32].

model are unimportant.[33] What is important is the relativity of Hilbert's proofs. In order to prove that the uninterpreted axioms of geometry are independent of one another, Hilbert employs the following procedure. First, he interprets the primitive terms of his axiomatic system in such a way that a model of his axioms is produced, i.e., an interpretation in which *all* the axioms are satisfied. (This was his way of proving their consistency). He then proceeds to construct a different interpretation on which only some of the axioms are satisfied, while the axiom to be proven independent is not (or, within which all the axioms and the negation of the axiom to be proven independent are satisfied).[34] This procedure is meant to illustrate that there are certain structures within which the axioms as a whole are satisfied and other structures within which the axiom to be proven independent is not satisfied. Hence, the axiom to be proven independent is not a purely logical consequence of the remaining axioms insofar as both it and its negation are consistent with those axioms. That an axiom and its negation are both consistent with a set of axioms illustrates, for Hilbert, that it is not a deductive, logical consequence of those axioms and, thus, that it is independent of them. As we will see, Frege would criticize this approach to independence proofs as incoherent, given what he takes to be a proper understanding of the nature of logical inference and model-theoretic techniques. Hilbert, for his part, offers no discussion about what logic is or how it operates with indifference to intuitive content.

### 1.4.3 Completeness

In addition to independence and consistency, Hilbert also suggests that any decent axiom system should possess the property of completeness. Given the close resemblance between Hilbert's concepts of consistency and independence with their modern model-theoretic equivalents, we may be tempted to align his notion of completeness, too, with its model-theoretic cousin. But, as Leo Corry correctly notes,

> The "completeness" that Hilbert demanded for his systems of axioms should not be confused with the later, model-theoretical notion that bears the same name, a notion that is totally foreign to Hilbert's axiomatic approach at this early stage. Rather it is an idea that runs parallel to [physicist Heinrich] Hertz's demand for "correctness". Thus, Hilbert demanded from any adequate axiomatization that it should allow for a derivation of *all* the known theorems of the discipline in question. The axioms formulated in [*Grundlagen*], Hilbert claimed, would indeed yield all the known results of Euclidean geometry or of the so-called absolute geometry, namely that [geometry which is] valid independently of the parallel postulate, if the corresponding group of axioms is ignored. Thus, reconstructing the very ideas that had given rise to his own conception, Hilbert discussed in great detail the role of each of the groups of axioms in

[33] Hilbert's interprets the axioms of his geometry, aside from the parallel axioms, in much the same manner as Beltrami had prior to him: he limits the points, lines, and planes to the interior of a sphere and suggests we "define the congruences of this geometry by aid of such linear transformations of the ordinary geometry as transform the fixed sphere into itself" [Hilbert 1903, 19].

[34] "By suitable conventions, we can make this "*non-euclidean geometry*" obey all the axioms of our system except the axiom of Euclid (group III). Since the possibility of the ordinary geometry has already been established [by Hilbert's consistency proof], that of the non-euclidean geometry is now an immediate consequence of the above considerations" [Hilbert 1903, 19].

> the proofs of two crucial results: the theorem of Desargues and the theorem of Pappus [...]. Unlike independence, however, the completeness of the system of axioms is not a property that Hilbert knew how to verify formally, except to the extent that, starting from the given axioms, he could prove all the theorems he was interested in.[35]

In contrast, then, to the modern understanding of model-theoretic completeness, Hilbert's view of the concept was a rather more informal, historical one. In his view, guided no doubt by his understanding of the non-arbitrary, intuitive basis of any interesting axiom system (in our case, guided by his view that the axioms of geometry are based upon our intuition of space), an axiom system is complete just in case it captures the theorems one would have liked it to capture and no anti-theorems. Even if Hilbert's axioms for Euclidean geometry were consistent and mutually independent, they would still be unsatisfying if they did not, for instance, include Pascal's theorem among their consequences. Completeness in Hilbert's sense is guided by a pre-systematic understanding of the nature and extent of the mathematical field one aims to describe via a particular axiomatization. If some interesting or important element of this field is not deducible from a particular axiomatization, that is good reason to prefer another axiomatization. A complete axiom system is one that captures all the theorems that it is (at least tacitly) intended to capture. Which theorems are intended is a matter determined largely by the historical trajectory of the field in question up to that point.

### 1.4.4 Categoricity

The explicit model-theoretic notion of categoricity[36] was not yet developed at the time of the publication of the *Grundlagen*. Nevertheless, Hilbert's addition, in the French translation of the *Grundlagen* (and all subsequent editions) of a further axiom, the axiom of completeness, seems to reveal that he intended his geometrical axiom system to be something like 'categorical' in our modern sense. But, again, there are important differences between Hilbert's view and the standard model-theoretic view of today.

Hilbert introduces the early version of this axiom as follows:

> REMARK. To the preceding five groups of axioms, we may add the following one, which, although not of a purely geometrical nature, merits particular attention from a theoretical point of view. It may be expressed in the following form:
>
> *AXIOM OF COMPLETENESS* [Vollständigkeit]. *To a system of points, straight lines, and planes, it is impossible to add other elements in such a manner that the system thus generalized shall form a new geometry obeying all of the five groups of axioms. In other words, the elements of geometry form a system which is not susceptible of extension, if we regard the five groups of axioms as valid.*[37]

[35] [Corry 2004a, 95-96].
[36] In model-theoretic terms, an uninterpreted theory is said to be categorical iff there exists an isomorphism between any two models of the theory.
[37] [Hilbert 1950, 16].

Unlike the other axioms, each of which states some property of or relation among the elements of the system (e.g., lines, points, or planes), this axiom states something about the axiomatized system itself, namely that no further elements can be added to the system while still maintaining consistency. This axiom, in contrast to the others, has the appearance of what we would now call a metatheoretic claim. What is its function here? Hilbert's system was intended, as we've seen, to present a simple and complete axiomatization of the known results of Euclidean geometry, all of which seems perfectly accomplished prior to the addition of this further axiom of completeness, so what purpose does it serve?

The name itself sheds light on Hilbert's own conception of completeness, which, as we've seen, was not obviously identical with any modern concept of deductive or semantic completeness. Hilbert's intent was to capture all the theorems of geometry, but, further, to do so in a way which accurately corresponded to the intuitive understanding of those theorems throughout the history of geometry. We can understand Hilbert's goal as one of systematization: he attempts to clear away irrelevant features of mathematical history in order to uncover the underlying logical and mathematical structure—the really interesting parts. To do so in a purely abstract manner, as postulation theorists like Huntington, Veblen, Hausdorff, and others would later do,[38] was not Hilbert's goal at all; his interest was to formalize the properly mathematical part of geometrical history, and this history included analytic geometry and its concomitant connection between the real numbers and the line.

What Hilbert's axiom attempts to formalize is the (often tacitly) accepted isomorphism between the real numbers and the points on a line. By rejecting models of his axioms which are *not* isomorphic to the real numbers (e.g., the 'incomplete' model within the field $\Omega$ of algebraic numbers outlined above), Hilbert illustrates that his intention is not simply to capture the abstract formal structure of Euclid's geometry—which is perfectly well accomplished within the subfield $\Omega$, and without the completeness axiom—but, also, to capture the precise way in which that structure was actually understood and employed by the mathematical community since at least the time of Descartes. Without the axiom, we might add historically and geometrically irrelevant bits to an 'incomplete' model like the subfield $\Omega$. By adding the completeness axiom, Hilbert explicitly forbids such extension, deliberately pushing his reader toward $\mathbb{R}^3$ for historical and intuitive reasons. Here again we can see Hilbert's divergence from the contemporary notion of categoricity: his intent here is not to ensure that all models of the axioms are isomorphic with one another, but, instead, to insist as carefully as possible that we employ the historically established interpretation of analytic geometry. Thus, though Hilbert's axiom of completeness might push his axiomatization toward categoricity, this was clearly not his own intention.[39]

[38] Cf., e.g., [Huntington 1904], where a "purely mathematical or abstract point of view" is employed in order to "consider the construction of a purely *deductive theory*, without regard to its possible applications" [Huntington 1904, 288]. He then speaks of his postulates (for an 'algebra of logic') as "simply *conditions* arbitrarily imposed on the fundamental concepts" of the theory [Huntington 1904, 290].

[39] [Awodey and Reck 2002, 8-13] gives a more thorough account of Hilbert's views in relation to the modern notion of categoricity.

The axiom of completeness is different in kind from the other axioms and axiom groups. While the remaining axioms formalize properties and relations related to the elements of the geometrical system, the axiom of completeness dictates that the system itself should obey a certain criterion, namely, that any *interpretation* of the system must have the property of being non-extendible. As we will see in detail when we turn to the correspondence between Frege and Hilbert, Hilbert's use of this isolated metatheoretic statement within his axiomatic system is illustrative of a general lack of precision regarding logical and model-theoretic concepts. It is precisely this sort of precision that Frege, with his penetrating analyses of basic logical concepts, would bring to the table.

1.4.5 Simplicity

Unlike consistency, independence, and completeness, the property of simplicity has only rarely been formalized, even today. Though no clear criterion for simplicity has won the day even in the contemporary mathematical world, the American postulation theorist E. H. Moore came close to providing one. In a 1902 paper on the definition of abstract groups, Moore wrote that

> From the standpoint of abstract logic the canons of relative simplicity of equivalent definitions by sets of postulates are not well established. Perhaps the only established canon is this, that a definition is simplified by the omission of a group of postulates logically deducible from the remaining postulates. One is tempted to add this, that every postulate of a desirably simple definition shall be a simple statement, that is, a single and not a multiple statement. The difficulty here would arise in the precise formulation of the terms of this second canon, especially in view of the fact that the same statement may be made in various forms. At least, a definition is simplified by the substitution, for a postulate consisting of an aggregate of independent statements, of those statements as distinct postulates. Further, one may add as a third canon this, that of two definitions one with the smaller number of postulates is the simpler. As to this canon the case now in question [i.e., the definition of the abstract group concept] seems to show that the definition with the larger number of independent postulates may reveal more immediately the fundamental properties of the object of definition. It is, of course, evident that the task of proving the independence of the postulates presumably increases with the number of postulates.[40]

The sort of precision sought by Moore here was not much pursued by other mathematicians, and simplicity has since remained a largely informal matter within mathematics.

Despite the general lack of precision regarding the notion of simplicity, most mathematicians, Hilbert included, take simplicity to be an extremely desirable property of a theory or axiomatization. Simplicity, in this respect, is largely an aesthetic notion, though one still capable of directing research and methodology. Recall, for instance, that the relative complexity of the parallel postulate when compared to Euclid's remaining axioms was the

[40] [Moore 1902b, 488-489]. E.V. Huntington makes similar, though less detailed remarks at [Huntington 1904, 290].

chief reason that a proof of that postulate was so ardently sought; there is something aesthetically displeasing to the mathematical temperament about an unnecessarily complicated theory or technique. Leo Corry is again helpful in understanding Hilbert's particular view of this notion (which agrees broadly with Huntington's account). He writes that

> The [...] requirement [of] simplicity [...] complements that of independence. It means, roughly, that an axiom should contain "no more than a single idea." [...] Nevertheless, it was neither formally defined nor otherwise realized in any clearly identifiable way within [*Grundlagen der Geometrie*]. The ideal of formulating "simple" axioms as part of this system was present implicitly as an aesthetic desideratum that was not transformed into a mathematically controllable feature.[41]

One way of reading this view of simplicity into Hilbert's *Grundlagen* is to take each of his axioms as the expression of a single facet of spatial intuition, while the axiom groups are collections of axioms which express facets of spatial intuition related to some one more general concept (e.g., incidence or order). Thus, the axioms of Hilbert's third axiom group all express distinct aspects of space related to the concept of congruence. For instance, Axiom III, 1 states that "*If A, B are two points on a line a, and A' is a point on the same or on another line a' then it is always possible to find a point B' on a given side of the line a' through A' such that the segment AB is congruent or equal to the segment A'B'*."[42] And Hilbert himself writes that "This axiom requires the **possibility of** *constructing* **segments**. Its uniqueness will be proved later."[43] So, while we have here a single aspect of spatial congruence (i.e., that it is possible to construct congruent line segments in space), it is a further and separate matter to prove the uniqueness of segment construction. Similarly, Hilbert presents the additivity of segments and the transitivity of congruence in separate axioms. Taken together, these axioms represent the presentation of the single concept of spatial congruence, while taken singly they each represent a single aspect (a 'characteristic mark,' as Frege might say) of that concept. Assuming that our axiom system is complete in Hilbert's sense (i.e., it captures all the theorems we want it to and no anti-theorems), this criterion of simplicity ought to ensure that the axiomatization possesses only as many axioms as are required. For, if each axiom is required to clearly represent a single idea, it will (ideally) be clear when two axioms express the same idea, thus enabling the easy removal of redundant axioms. The result of such a complete and simple system should be a minimal collection of axioms from which the maximum number of theorems may be derived.[44]

In Hilbert's work we can see very clearly the shift away from a geometry which was oriented towards the properties of particular sorts of geometrical objects (points, lines, planes, etc.) and towards an examination of the structure of axiomatic systems. Hilbert's view is an interesting hybrid of these two approaches: he maintains the utility of intuition as a heuristic

[41] [Corry 2004a, 95].
[42] [Hilbert 1971, 10].
[43] [Hilbert 1971, 10].
[44] Bolzano offers a similar view at [Bolzano 1837, §221.2, note].

for determining which axiom sets are worthy of mathematical analysis, while eschewing any appeal to intuition in his attempt to provide satisfactory axiomatic characterizations for various geometric structures. The main focus of his work, however, was to prove theorems *about* certain geometric structures rather than proving theorems within them.

Hilbert's main interest is not in triangles, for example, but with the relationship between theorems pertaining to triangles and the postulates of a particular axiomatization. Similarly, he is not focussed upon the content of Archimedes' axiom as much as its relation to the other axioms of Euclidean geometry. As we will see in more detail below, this shift from an object-oriented approach toward a structure-oriented one was not without its difficulties.

From Hilbert's views I would now like to turn to those of Frege. Both men had complex relationships with prior mathematical and philosophical traditions: Hilbert took on some of the intuitive elements of geometry with his understanding of the nature of axioms, but insisted upon the purely deductive character of geometric proofs and theorems. For his part, Frege held more conservative views about the nature of elementary geometry, but was a radical innovator when it came to logical precision and mathematical rigor.

## §2. Frege

### 2.1 Frege's work in context

Frege had, as early as the late 1870s, become very interested in presenting complex mathematical concepts and proofs in a more perspicuous and rigorous manner. As we've seen, Hilbert's work in geometry attempted to provide a coherent axiomatic framework capable of representing a whole host of new and old geometrical results; Hilbert's work rather successfully eliminated (or, at least limited) geometry's reliance upon the intuition of space. This work was an evacuation of geometry's logically impertinent trappings, resulting in a pared-down version of the age-old truths of Euclid. Frege, for his part, had focussed more heavily on clarifying the nature and logical structure of certain basic mathematical concepts and modes of reasoning (e.g., the concept of number and mathematical induction).

While Frege was quite familiar with the revolutions taking place throughout geometry—he wrote a dissertation at Göttingen about geometry in the early 1870s,[45] as well as a few short reviews and a paper dealing with geometrical issues[46]—his views were for the most part those of a staunch geometric conservative. He was, for instance, convinced very late in his career of the intuitive basis of the truth of the axioms of Euclid's geometry, and he

---

[45] [Frege 1873], supervised by Ernst Schering (1824-1897) who, incidentally, was the editor of Gauss' collected works. Frege was also quite familiar with the work of the innovative algebraic geometer Alfred Clebsch (1833-1872), having attended his lectures on analytic geometry.

[46] The reviews are: [Frege 1877] (reviewing a work on linear geometry and its applications) and [Frege 1880], which reviews a textbook in analytic geometry; the paper is [Frege 1878], which discusses the possibility of conceiving of triangles as complex quantities. Prior to his exchanges with Hilbert and Korselt, we find Frege returning only briefly to geometrical matters in [Frege 1884a], the text of an 1883 lecture on plane geometry; though he did teach geometry, which likely kept him abreast of developments in projective geometry (in which he showed interest throughout his career).

characterized these axioms as necessary truths about the physical world. The following fragment on Euclidean geometry, which remained unpublished during his lifetime, is illustrative of his general approach to geometric questions: "In Euclidean geometry certain truths have traditionally been accorded the status of axioms. No thought that is held to be false can be accepted as an axiom, for an axiom is a truth. Furthermore, it is part of the concept of an axiom that it can be recognized as true independently of other truths."[47] Elsewhere, he offered some rather unconvincing arguments about the indispensability of Euclidean geometry for getting around in the world:

> Space, according to Kant, belongs to appearance. It would be possible for it to be differently represented by another rationality entirely different from us. True, we cannot even know whether it appears the same to one person as to another; for we cannot lay one intuition of space next to another in order to compare them. There is nevertheless something objective contained therein; everyone recognizes the same geometrical axioms even if only by the fact that he must do so in order to find his way in the world. What is objective in it is what is law-governed, conceivable, and capable of being judged, what is expressible.[48]

That Frege's claims for the experiential or pragmatic necessity of Euclidean geometry here miss their mark is shown by the fact that even his contemporaries (e.g., Poincaré and Helmholtz) were quite capable of illustrating the compatibility of everyday experience with geometries other than Euclid's.[49] In addition to this pragmatic consideration in favour of the traditional view of Euclidean geometry, Frege also conceived of physical space as Euclidean in nature and he believed we were able to grasp basic truths pertaining to that space (i.e., Euclid's axioms) *a priori* by way of pure intuition. Thus, in most respects, he was an orthodox Kantian when it came to the philosophy of geometry. Frege's fairly traditional views might partially explain his initial reaction to Hilbert's work on the foundations of geometry, but neither that reaction nor its continuation throughout his correspondence with Hilbert (and the subsequent series of articles written in response to Korselt) can be fully explained simply in light of those views. Indeed, the most interesting and far-reaching of Frege's criticisms of Hilbert's work have very little to do with his own geometrical views and quite a lot to do with his more impressive re-imagining of the nature of logic and the relations between mathematics and logic.[50]

Before we turn to Frege's criticisms of Hilbert I will discuss some elements of one of the main mathematical traditions which preceded and informed Frege's own work in logic and the foundations of mathematics. As we will see, it is the trajectory partially determined by this tradition, rather than Frege's own views of the nature of geometry, which informed many of the most important of his responses to Hilbert.

---

47 [Frege c. 1900, 169].
48 [Frege 1884b, §26].
49 Cf., e.g., [Helmholtz 1977].
50 See [Rusnock 1993] for a more detailed argument along these lines.

2.1.1 The rigorization of mathematics

Alongside the structure-focussed, axiomatic approach to geometry which had developed from the late 18th century and throughout the 19th century, there was another equally radical and influential movement within analysis and number theory which focussed on tightening up mathematical argumentation and providing stricter definitions of basic concepts. Many elements of Frege's work on the foundations of mathematics and logic belong squarely within this tradition and it is worthwhile to outline, at least briefly, some of the aims and methods which characterize it. This discussion should help to illustrate that Frege's criticisms of Hilbert on the foundations of geometry were not merely the result of his own somewhat outmoded, Kantian view of geometry but, rather, that they developed mainly as a result of his preoccupation with the central themes of a nineteenth century tradition of mathematical rigor.

The modern project of injecting rigor into mathematics began in earnest at the close of the 18th century, and was developed, for the most part, in response to the conceptual morass introduced into mathematics by the (then) obscure notions of the infinitesimal calculus. From the time of Newton and Leibniz, there had been an uneasy feeling surrounding the use of unclear notions like fluxions, derivatives, second and third differentials, vanishing quantities, etc. One of the better known (and certainly the most amusing) of the polemics deployed against these concepts can be found in Bishop George Berkeley's *The Analyst*. There Berkeley famously derides the use of fluxions in the calculus: "And what are these Fluxions? The Velocities of evanescent Increments? And what are these same evanescent Increments? They are neither finite Quantities nor Quantities infinitely small, nor yet nothing. May we not call them the Ghosts of departed Quantities?"[51] In conjunction with his mocking tone, Berkeley also offered a detailed critique of several of the central notions of the calculus. While he dismissed extant attempts to provide a conceptual basis for the successes of the calculus, he did not dismiss those successes themselves. Instead he explained them as the felicitous result of two competing errors. While considering the obscure solutions to several problems related to the determination of tangents for parabolas, Berkeley writes that "the two errors being equal and contrary destroy each other; the first error of defect being corrected by a second error of excess,"[52] and, further, that

> If you [i.e., the mathematician employing the methods of the calculus] had committed only one error, you would not have come at a true Solution of the Problem. But by virtue of a twofold mistake you arrive, though not at Science, yet at Truth. For Science it cannot be called, when you proceed blindfold, and arrive at the Truth not knowing how or by what means.[53]

[51] [Berkeley 1734, paragraph XXXV].
[52] [Berkeley 1734, paragraph XXI].
[53] [Berkeley 1734, paragraph XXII].

Despite the fact that the results of the calculus were quite often correct, Berkeley resisted treating the methods of fluxions and vanishing quantities as scientific ones for the simple reason that no one could explain how these methods managed to arrive at truth. Without some form of conceptual clarification, the success of the calculus appeared to result only from an obscure and inexplicable fumbling toward the truth, a kind of mathematical water-dowsing. Accordingly, its results could not be called scientific, as it was not at all clear how or even if those results could be grounded in scientific reason.

While many mathematicians were more than willing to employ the notions of the calculus as long as they delivered results, the scathing criticisms offered by Berkeley and others eventually began to weigh on the more philosophically inclined members of the mathematical community. Throughout the latter half of the 18th century, there was a great push to remedy the situation. Two general approaches were taken: one which attempted to deliver the same results as the calculus without employing its more troublesome concepts, and one which sought to clarify those concepts and place them on a sounder philosophical and scientific footing. Joseph Louis Lagrange (1736-1813), one of the foremost representatives of the former method, published an influential text on the theory of functions, whose subtitle boasted that it was "redeemed from all considerations of the infinitely small, from vanishing quantities, from limits and from fluxions, and reduced to the algebraic analysis of finite quantities."[54] Lagrange, as many others, characterized the new demand for rigor as a return to the spirit of precision exemplified by the ancients,[55] and was so concerned with the incoherent conceptual foundations of the calculus that, while serving as the director of the mathematical section of the Berlin Academy, he set the following prize problem for the year 1784:

> It is widely known that advanced geometry regularly employs the *infinitely great* and the *infinitely small* ... The Academy, therefore, desires an explanation of how it is that so many correct theorems have been deduced from a contradictory supposition, together with... a truly mathematical principle that may be substituted for that of the *infinite*.[56]

The feeling evidenced here and throughout the mathematical works and correspondence of the second half of the 18th century was that there had to be some explanation for the great successes, both mathematical and physical, of the calculus. The concepts of the calculus, however murkily understood, indisputably led to valid results, and no one could convincingly explain how or why. The attempts to avoid the issue by founding the calculus without the use of its most troublesome concepts ultimately had less appeal than those attempts based on conceptual clarification (it is indeed a rare case where mathematicians give up even shoddy tools if they produce beautiful results). It is largely to this work of clarification, begun in the late eighteenth century, that we owe the advances in mathematical

---

[54] [Lagrange, quoted in Grattan-Guiness 2000, 100].

[55] "L'objet de cet Ouvrage est de donner la théorie des fonctions, considérées comme primitives et dérivées, de résoudre par cette théorie les principaux problèmes d'Analyse, de Géométrie et de Mécanique, qu'on fait dépendre du Calcul différential, et de donner par là à la solution de ces problemes toute la rigueur des démonstrations des Anciens" [Lagrange 1867, tome 9, 20]

[56] [Lagrange 1784, quoted in Grattan-Guiness 2000, 101].

rigor made in the following century by Cauchy, Gauss, Abel, Dirichlet, Bolzano, and others. These mathematicians achieved a much stricter understanding of the demands to be met by any piece of rigorous mathematical reasoning, one that even Berkeley would likely have found unexceptionable.

Bolzano, whose work remained largely unknown during the period, nevertheless provides us with an exemplary philosophical account of proper mathematical reasoning in his essay on the mathematical method.[57] Bolzano, like Frege later, was interested in eliminating the psychological elements of mathematics in favour of a focus on the objective logical relations which held amongst its propositions (*Sätze an sich*), irrespective of their historical epistemological status. Thus, in the works of Bolzano and the other rigorizers, we find an insistence upon proving even obvious theorems[58]—a tradition which had slackened considerably since the time of the Greeks, perhaps because of the rapid advances in mathematical applications after the emergence of the integral and differential techniques of Newton, Leibniz, and the Bernoullis. Bolzano was critical of the lax treatment of definition within contemporary mathematical reasoning, and constructed a much stricter theory of definition than the relatively informal approach that had become popular by the early 19th century. Inferential gaps, too, were to be filled by the newly rigorized approach: hidden assumptions were to be uncovered and made explicit, superfluous appeals to intuition were to be dispensed with, and any propositions susceptible of proof (obvious or not) were to be proved. If a proposition was assumed but not proved, it was to be expected that its unprovability was to be demonstrated or, at least, that this would remain an open problem for research.

The appreciation for strict forms of definition and gapless chains of inference exhibited by Bolzano and others in the early part of the 19th century was extended considerably in the latter half of that century by Karl Weierstrass (1815-1897) and the talented circle of mathematicians which formed around him in Berlin. The members of the school were instrumental in the so-called arithmetization of analysis, a process which involved reconstructing many of the theorems and concepts of analysis in terms of a group of acceptable techniques which avoided as much as possible any appeal to (geometric) intuition and the murkier concepts of the early calculus. Weierstrass and his school developed and popularized rigorous definitions of a number of basic mathematical concepts, chief among them were the ($\varepsilon$, $\delta$) definition of a limit, of uniform convergence, and of the uniform continuity of a function.[59]

---

[57] [Bolzano 2004b].

[58] Bolzano's 'purely analytic' proof of the intermediate value theorem (i.e., [Bolzano 1817]) is an excellent example of the application of rigorous forms of reasoning and definition to a seemingly obvious theorem.

[59] Though Cauchy was the first to offer this form of the definition of a limit (see [Cauchy 1823, 44]), Weierstrass and his circle employed it frequently and in a variety of contexts, with the effect that it was disseminated more widely; for the Weierstrassian definition of uniform convergence (in a region), see [Weierstrass 1880, article 1]; for that of the uniform continuity of a function, see [Heine 1870, 361]. (Interestingly, an earlier, almost equally rigorous definition of uniform continuity was implicitly given by Bolzano in the 1830s in his *Functionenlehre* ([Bolzano 1930, §13; English translation available in [Bolzano 2004a]—some relevant emendations can be found in a manuscript, also available in translation in [Bolzano 2004a, 573-589]). For a discussion of Bolzano's priority in the matter, see [Rusnock 2000, 166-172] and [Rusnock and Kerr-Lawson 2004].

During the period of the Weierstrass school's influence at Berlin, a rival tradition arose at Göttingen centered around Bernhard Riemann. The Göttingers were noted less for their strictness and more for their mathematical creativity. Like the early workers in the calculus, those following in Riemann's wake were less interested in shoring up the rigor of their new topological and functional investigations than in extending as far as possible new mathematical vistas. Göttingen, in contrast to Weierstrass's Berlin, was noted for its more liberal use of geometrical and other forms of intuition as a fruitful generator of mathematical material. Much of German mathematics at the time was divided between the two schools, to the point that very little cross-pollination by way of publications, correspondence, or colloquia occurred. Only towards the end of the nineteenth century, when Klein and Hilbert were in ascendency at Göttingen did the two traditions begin to run together—though even then there were firm distinctions of style and method.[60]

By all appearances, Frege was well-disposed to become a card-carrying member of the Riemann school. He completed his doctoral work on geometry in Göttingen, he was quite familiar with the work of Clebsch (a highly influential Göttinger), and later worked on the geometric representation of complex quantities, a topic which combined two key elements of the Göttingen tradition. Despite these connections, however, Frege's more mature work focussed almost exclusively on the foundations of arithmetic and elementary number theory, topics considerably more Weierstrassian than Riemannian.[61]

Even while Frege was clearly inspired by the Weierstrassian project of rigor, he brought a peculiarly logic-centred approach to this project which was absent in the Berlin school. While the Weierstrassians were concerned to provide analysis with rigorous conceptual foundations, Frege was interested in going further by providing strict definitions which would reduce complex mathematical notions to purely logical ones. Thus we find, for instance, his attempt to define number in terms of the extensions of certain concepts, which he understood to be properly logical objects. While the details of Frege's efforts to clarify the nature of particular mathematical results are not directly relevant to our concerns here, his broad conception of the proper nature of mathematical reasoning and definition are. Many of his criticisms of Hilbert's views in the *Foundations of Geometry*, for instance, have to do with what Frege perceived as failures to observe the proper distinctions between definitions, axioms, and explanations, as well as Hilbert's improper understanding of the nature of definition more generally. Accordingly, let's take a look at what Frege has to say about the nature of logic, and, following that, what he has to say about Hilbert's allegedly inadequate understanding thereof.

---

[60] For an account of the Göttingen school with particular emphasis on Klein and Hilbert, see [Rowe 1989]; for the Berliners, see [Biermann 1973].

[61] Though Frege's earliest work (e.g., his dissertation of 1874) dealt with geometrical issues, and he taught several courses relating to various geometrical topics, it is apparent that the bulk of his own mental energy was expended on his project of rigorizing arithmetic by providing it with solid foundations in logic. This project was clearly related to (though Frege was highly critical of) similar efforts by the Weierstrassians. For a view which emphasizes Frege's connection to the Riemannian tradition at Göttingen, where Frege did his graduate work, (rather than the Weierstrass-Berlin tradition), see [Tappenden 2006].

## 2.2 Frege on the nature of logic

In a fragment from what appears to have been intended as a logic textbook, Frege writes that

> Logic is concerned only with those grounds of judgement which are truths. To make a judgement because we are cognisant of other truths as providing a justification for it is known as *inferring*. There are laws governing this kind of justification, and to set up these laws of valid inference is the goal of logic.[62]

Thus logic is a science which considers the laws of inference in their relation to truths. In the modern period, many logicians took this to involve distinguishing various types of judgments and the possible inferential relations between them. For Frege, by contrast, the aspects of inference in which logic is interested are not at all psychological, or even empirical, and hence logic proper is not concerned with the peculiarities of human judgment. (At which temperature humans best infer or with which mental images were not, for Frege, matters for logic to decide). In the same fragment, he writes the following:

> The subject-matter of logic is therefore such as cannot be perceived by the senses and in this respect it compares with that of psychology and contrasts with that of the natural sciences. Instincts, ideas, etc. are also neither visible nor tangible. All the same there is a sharp divide between these disciplines [i.e., logic and psychology], and it is marked by the word ‘true’. Psychology is only concerned with truth in the way every other science is, in that its goal is to extend the domain of truths; but in the field it investigates it does not study the property ‘true’ as, in its field, physics focuses on the properties ‘heavy’, ‘warm’, etc. This is what logic does. It would not perhaps be beside the mark to say that the laws of logic are nothing other than an unfolding of the content of the word ‘true’. Anyone who has failed to grasp the meaning of this word—what marks it off from others—cannot attain to any clear idea of what the task of logic is.[63]

The references to psychology here are not incidental, as a variety of psychologistic views of both logic and mathematics were popular during Frege’s time, and a considerable portion of his work was dedicated to undermining these views and building up his own alternative conception.[64] Psychologism in logic, as the name suggests, is the view that logical laws are subject to, deducible from, or in some way reducible to psychological laws, and that logic itself studies the particular ways in which human beings infer and make judgments. Instead of the laws regulating ideas and the association of ideas found at the heart of psychologistic accounts of logic, however, Frege turned to the study of the laws of truth itself. Instead of

[62] [Frege ?1879-1891, 3].
[63] [Frege ?1879-1891, 3].
[64] Psychologism in logic has been popular in a variety of forms throughout the history of the discipline; in Frege’s time, the *Logic* of John Stuart Mill ([Mill 1843]) was one of the major sources of the view. In addition to Mill, a number of German logicians held psychologistic views, notable among them were Christian von Sigwart (1830-1904) and Benno Erdmann (1851-1921). One of the better-known disputes regarding psychologism in mathematics can be found in Edmund Husserl’s psychologistic views on arithmetic in [Husserl 1891] and Frege’s highly critical review of Husserl [Frege 1894].

studying how humans actually infer (which is, more often than not, erroneously) Frege endeavoured to study the nature of correct inference and the relations between inference and truth. Since logic, properly construed, is "the science of the most general laws of truth,"[65] and since truth is not, for Frege, a subjective, mental property, logic in its proper form cannot be concerned with the subjective, mental realm.

Frege's own view of logic involved a shift away from the subjective ideas of psychologism towards what he (rather unfortunately) calls 'thoughts'. We will have occasion in the following chapter to return in more detail to Frege's 'thoughts'; here it is sufficient to characterize thoughts as those things which are true or false.[66] For Frege, truth was not at all a subjective property, nor, somewhat confusingly, were thoughts mental entities. Since logic deals with truth and the relation between thoughts and truth (neither of which are psychological or subjective in nature), it cannot be a subjective science. Nevertheless, Frege admits that *our* interaction with logic has to occur within the spatiotemporal world we inhabit and, thus, that any thoughts which we endeavour to study scientifically must be cloaked in something tangible, e.g., some language or mental process.

Because of its public character, language is invariably the means by which scientific results are communicated and formulated. Unfortunately for science, however, the relationship between thoughts and their linguistic garb is not always a clear or helpful one. Ordinary language can obscure the logical structure of thoughts; indeed, it often systematically misleads the unwary away from the true nature of proper inference. Completely different phrases may express thoughts which are, from a logical point of view, identical. Similarly, identical phrases, uttered or written in different contexts, may express different thoughts. In many cases, too, ordinary language is insufficiently precise to pin down a single meaning for a particular phrase, such that it may seem to express several different thoughts at once. Frege sought to remedy this situation through the creation of a logically perspicuous language. The result was his *Begriffsschrift*, or concept-script.[67] Frege's formal language was based upon a cluster of rigorously circumscribed logical concepts and rules of inference.[68] From this basic stock, Frege held that the logical structure of thoughts and their relations to each other could be clearly presented while most logically irrelevant properties (i.e., those which do not affect the truth or falsity of thoughts or the validity of inferences) could be ignored. In addition to the purely symbolic work, the construction of this language also required a significant amount of conceptual work: Frege had to isolate the logical concepts from those alien to logic, he also had to avoid the ambiguity and obscurity which prevented logical perspicuity

---

[65] [Frege 1897, 128].
[66] "I mean by 'a thought' something for which the question of truth can arise at all" [Frege 1918, 4]. Though Frege's idiosyncratic use of the word 'thought' to represent the objective content of a judgment has not caught on, the basic idea has persisted within contemporary philosophy under the heading 'proposition'. Frege reserves the German term '*Satz*' for what are now often referred to in English as sentences, i.e., verbal or written expressions of objective propositions (or Fregean thoughts).
[67] [Frege 1879].
[68] The basic symbols of Frege's first system, which express his basic concepts, are: the judgment or assertion stroke (§2), the conditional stroke (§5), the negation stroke (§7), the identity symbol (§8), letters which serve as variables (§§10-12), and the symbol for generality (§11). In addition to the specific notions expressed by these symbols, Frege also discusses the metatheoretic concepts of 'function' and 'argument' (§9).

within ordinary language. To this end, he developed some rather strict ideas concerning the relations between language and logic, and carefully distinguished between thoughts and their various possible expressions.

## 2.3 The role of a scientific language

In an interesting paper defending and elaborating upon the role of his *Begriffsschrift*, Frege opens with the following paragraph:

> Time and again, in the more abstract regions of science, the lack of a means of avoiding misunderstandings on the part of others, and also errors in one's own thought, makes itself felt. Both [shortcomings] have their origin in the imperfection of language, for we do have to use sensible symbols to think.[69]

Thus we find ourselves in something of a quandary. Logic, as noted, studies relations, objects, and properties which are not at all sensible. But, nevertheless, given that we are experientially bound to our spatiotemporal world, we must employ spatiotemporal means in our study of logic. That this is possible, Frege does not doubt, but neither does he explain the possibility of interaction between an objective, intangible, non-physical realm and our physical world. Despite the lack of any detailed metaphysical or ontological speculation on the nature of this relation, Frege does discuss in some detail specific ways in which slippage between the two realms can occur and lead us into error. Ordinary language, for instance, is very often ambiguous: sentences or individual words can possess multiple meanings, some of which can be determined from context, some of which cannot. In addition to being ambiguous, words and sentences may also be vague, such that it is not at all clear which thoughts (if any) they make tangible. In ordinary, everyday speech ambiguity and vagueness are mostly harmless. However, when it comes to scientific uses, where precision and the pursuit of truth are paramount, the situation becomes intolerable.[70]

Since ordinary language drives us away from the clarity and precision required to investigate the realm of thoughts and to communicate our findings to others, science must either put language aside or construct linguistic tools which minimize as much as possible the slippage between thoughts and their expressions. And, since Frege admits that (at least for humanity) language is a necessary accompaniment of our apprehension of thoughts, science is in need of a logically precise language. Indeed, Frege notes that, even despite the obvious difficulties of employing the sensible to represent the non-sensible, the use of symbols allows us

> [...] to penetrate step by step into the inner world of our ideas and move about there at will, using the realm of the sensible itself to free ourselves from its constraint. Symbols have the same importance for thought that discovering how to use the wind to sail against the wind had for navigation. Thus let no one despise symbols![71]

---

69 [Frege 1882, 83].

70 For an interesting analysis of Frege's views on vagueness in light of the modern differentiation between semantic, epistemological, and ontological vagueness, see [Ruffino 2003].

71 [Frege 1882, 84].

The importance accorded to symbols here is impaired by the imprecision in ordinary language. In order to better avoid this sort of imprecision in his *Begriffsschrift,* Frege developed a sophisticated account of the relations between language and the realm of thought. Most relevant for our purposes here are his views on the nature of definitions, axioms, and what he called ‘elucidations,’ to which I now turn.

## 2.4 Definitions

For Frege, within a regimented language suitable for science, the occurrence of a sign within a sentence should indicate that the sense of that sign plays a role within the sense of the sentence (i.e., within the thought that the sentence expresses). Thus, a sentence “is put together in such a way that parts of the thought correspond to parts of the sentence”.[72] Thus, in an ideal scientific language, the structure of thought is perfectly mirrored by the structure of the language itself. Of course, with the increasingly complex concepts and thoughts of modern science, such a goal is difficult (not to mention irritating) to maintain. To remedy the potential annoyance of perfect rigor, Frege allows for definitions as a kind of shorthand, though he does so in a way which he believes can maintain the intended correspondence between thought and language. He does this by treating definitions as arbitrary stipulations which give a simple sign a sense identical to the sense of a complex sign or group of signs.[73] Moreover, the sense of the complex sign must already be known. This apparently simple understanding of definition has a number of important consequences, particularly when framed in relation to Frege’s other views concerning the nature of logic.

Frege’s definitions are, for instance, tautologous; given that a definition stipulates that the sense of a previously-senseless, simple sign is to be equivalent to that of an already-known complex sign, and given that the sense of this complex sign can be ‘read off’ from the senses of its parts in a regimented language, there can be no doubt as to the truth of a definition. Nor can such a definition be informative in any positive sense. For all a definition says, on this view, is that a sign possesses its own sense, i.e. the sense that we have decided to give it. Once we accept a definition, it becomes a tautology.

Another important consequence of Frege’s understanding of the role of definition is that definitions are, at least logically speaking, superfluous. As we’ve seen, Frege’s definitions stipulate for some *sign* that it possesses a sense identical to the sense of another sign or group of signs. Definition thus occurs on a linguistic level, and definitions amount to linguistic conventions by which we agree to equate the senses of two distinct signs (or sign-complexes). Now, for Frege, any two expressions which express identical senses are logically equivalent; indeed, he calls definitions ‘identities’ or ‘logical identities’ on several occasions.[74] Thus, in any instance where we might employ one side of a definitional identity we could, without any change in truth conditions, employ the other. From a logical point of view, then, definition is a mere superfluity. Of course, psychologically or pragmatically,

[72] [Frege 1914, in PW 207].
[73] Cf. [Frege 1893, §27], [Frege 1914, in PW 210].
[74] At, e.g., [Frege 1893, §33], [Frege 1892, in PW 102], and [Frege 1914, PW 208].

definitions are often indispensible aids to the continued progress of a given science. This is particularly true in mathematics and physics, where linguistically simple terms like ‘limit’ or ‘tensor’ allow scientists access to the mental shortcuts required to produce results in the hyper-specialized fields of modern science. Without the ease of operation enabled by definitions and their linguistic condensation of complex concepts, the smooth functioning and forward progress of science would be greatly stunted. This is nowhere more evident than in contemporary mathematics and physics, where whole axiomatic systems are often condensed into a few words. Without the conceptual shorthand provided by well-defined terms for basic mathematical structures like groups or rings, it would be very difficult for mathematicians to work with more complex structures built up from these simpler constituents—e.g., cohomology groups or topological rings. Even at the level of group theory itself, it would be rather difficult, from both the typesetter’s and the mathematician’s point of view, to prove theorems pertaining to groups if one were forced to restate the group axioms at each occurrence of the term ‘group’. While Frege might take definitions to be logically inert, he does not deny their necessity from a pragmatic point of view.

An important corollary of the logical inertness of Fregean definition is that no such definition can be used to prove a proposition which would otherwise remain unprovable. Consider again that, for Frege, a definition is mere shorthand which should always be eliminable, however tedious such elimination might be in practice. On this view, in any instance where we employ a definition within a proof we could just as well employ the complex sign or group of signs the definition is intended to abbreviate. This view also tells against the informativeness of proper definitions, and against the use of ‘creative’ definitions within mathematics. On this score, Frege writes the following:

> In the development of science it can indeed happen that one has used a word, a sign, an expression, over a long period under the impression that its sense is simple until one succeeds in analysing it into simpler logical constituents. By means of such an analysis, we may hope to reduce the number of axioms; for it may not be possible to prove a truth containing a complex constituent so long as that constituent remains unanalysed; but it may be possible, given an analysis, to prove it from truths in which the elements of the analysis occur. This is why it seems that a proof may be possible by means of a definition, if it provides an analysis, which would not be possible without this analysis, and this seems to contradict what we said earlier. Thus what seemed to be an axiom before the analysis can appear as a theorem after the analysis. But how does one judge whether a logical analysis is correct? We cannot prove it to be so. The most one can be certain of is that as far as the form of words goes we have the same sentence after the analysis as before. But that the thought itself also remains the same is problematic. When we think that we have given a logical analysis of a word or sign that has been in use over a long period, what we have is a complex expression the sense of whose parts is known to us. The sense of the complex expression must be yielded by that of its parts. But does it coincide with the sense of the word with the long established use? I believe that we shall only be able to assert that it does when this is self-evident. And then what we have is an axiom. But that the simple sign that

> has been in use over a long period coincides in sense with that of the complex expression that we have formed, is just what the definition was meant to stipulate.[75]

Frege's worry here is the apparent contradiction between his view that a definition cannot occur *essentially* in a proof and the claim that "it may not be possible to prove a truth containing a complex constituent so long as that constituent remains unanalysed". Frege's explanation here, as I read it, is that, while certain definitions seem to perform an essential function within proofs, this is illusory, an after-effect of the mistaken belief that the complex sign expresses some simple sense. For, on Frege's view, either we grasp the sense of the complex sign—in which case the logical analysis of that sign's meaning only tells us what we already know and is thus superfluous—or we do not. If we do not grasp the sense of the sign, then we have strayed from Frege's vision of a perfectly scientific language in which every sign and sentence possesses a clear meaning. We might read Berkeley's attack on the 'unscientific' methods of the early calculus as an example of the sort of thing Frege wishes to highlight here: we may appear to be able to prove theorems with the modern definition of a limit that we might not have been able to prove while employing Newtonian fluxions, and yet we might view fluxions as somehow definable in terms of more precise modern notions. The problem, as Berkeley points out, is that the concept of a fluxion (among others) was insufficiently precise, such that we cannot really view the definition of integrals and derivatives in terms of more rigorous modern concepts to be the kind of 'logical analysis' which Frege discusses in this passage. Thus, we do not gain access to new theorems purely by means of logical analysis, as the concepts purported to be analysed in this case were, in a sense, not full-fledged concepts at all, and, therefore, incapable of being analysed in a purely logical manner. None of this, of course, tells against the historical or epistemological usefulness of these sorts of definitions, nor against the epistemological connection between vague 'concepts' like those of the early calculus and their more precise modern cousins.

But, even granting that we understand the sense of the complex sign without already grasping its identity with the sense of a simpler sign given by a logical analysis, Frege notes potential problems. If we fail to grasp immediately that two senses are identical, we either have a vague apprehension of the sense (which is a psychological or physiological problem, and not a logical one), the sense itself is vague (in which case it is no 'sense' at all in Frege's view), or we have mistakenly shifted the meanings of our terms, such that we take now one and now another sense to be the object of our logical analysis. If we are working within a scientific language constructed along Fregean lines, in none of these three cases does the definition, in its role as a definition, provide us with information that we could not otherwise have gained. For Frege, then, the appearance that a definition is essential or otherwise informative is an historical accident related to our misapprehension of a particular sense, which has nothing to do with the definition in its proper role.

On the face of it, this view seems rather puzzling: for, surely, we can employ a sign properly without grasping all the finer characteristics of its sense. On this Frege writes

[75] [Frege 1914, PW 209-210].

> The fact is that if we really do have a clear grasp of the sense of the simple sign, then it cannot be doubtful whether it agrees with the sense of the complex expression. If this is open to question although we can clearly recognize the sense of the complex expression from the way it is put together, then the reason must lie in the fact that we do not have a clear grasp of the sense of the simple sign, but that its outlines are confused as if we saw it through a mist. The effect of the logical analysis of which we spoke will then be precisely this—to articulate the sense clearly. Work of this kind is very useful; it does not, however, form part of the construction of the system, but must take place beforehand.[76]

Frege admits, then, that occasionally the use of a simple sign—a sign which really represents, indeed is defined as identical in sense with, some complex expression—becomes so entrenched that we can lose sight of the fact that it is identical in sense to some complex sign or group of signs. In effect, we fail to grasp the sense of the simple sign fully, and grasp only enough of this sense to employ it correctly in certain contexts. Still working under this veil, we then perform a logical analysis (i.e., we construct a definitional equality) which illustrates that the simple sign has the same sense as a complex sign whose sense is already known. But, for Frege, the correctness of the thought which establishes this identity is not and cannot be a positive discovery relative to the science at issue. For we can be neither correct nor incorrect in our logical analysis until both terms already possess some sense, and it is precisely the provision of the simple term with a sense that is performed by a definition. If we grasp the sense of the simple term clearly, then we will know immediately whether or not it is identical with the sense of the complex sign, which is already known. So, without a clear sense already in mind, the so-called logical analysis of the sense of a simple sign by means of a definition into a complex sense is entirely impossible—something like this kind of analysis is only made possible by *our* vague, unclear grasp of a sense which is, in itself, perfectly precise.

The discovery of logically efficacious definitions already presupposes a prior assumption about the equivalence of the senses of the two expressions. Such logical analyses, then, simply establish clearly the sense of the signs to be used within the system—they clarify what must already be known if we are to work and think scientifically, i.e., precisely.[77] Such work is undoubtedly important, and many of the most fruitful mathematical definitions perform this kind of clarificatory work. Though not at all in accordance with Frege's particular vision of the nature of definition, the early history of group theory developed along these lines. The group concept was first employed in the very specific context of attempting to illustrate the general insolubility of quintic equations. It was only through great mathematical excavation and rigorization that the abstract concept of the group fully emerged, such that the factor groups employed by Galois and others in their search for an answer regarding quintics were seen to be specific instances of a more general mathematical phenomenon. The eventual definition of the group concept, by way of the now well-known group axioms, occurred only after the concept had been sufficiently clearly grasped to be

---

[76] [Frege 1914, PW 211].

[77] A more extensive discussion of the ways in which Frege deals with this apparent 'paradox of analysis' can be found in [Horty 2007, 36ff.] and [Resnik 1980, 43-47; 181-185].

employed in a number of seemingly unrelated contexts.[78] So it is that many mathematical concepts are grasped only partially, or 'seen through a mist', and, with this tenuous grasp, come to be clarified through what appears to be purely logical analysis. Frege's point is that this is epistemological, historical, or psychological analysis, and not the sort of systematic logical activity it may on occasion appear to be. For Frege, then, this is not really *logical* work insofar as it has no real correlate within the objective realm of thoughts, nor, correspondingly, within a perfectly scientific language.

Another important quality of Fregean definitions is their 'completeness'. As a result of his views on the nature of sense (and the nature of regimented languages suitable for science), Frege requires that definitions of concepts[79] must be 'complete' in the following sense:

> A definition of a concept (of a possible predicate) must [...] unambiguously determine, as regards any object, whether or not it falls under the concept (whether or not the predicate is truly assertible of it). Thus there must not be any object as regards which the definition leaves in doubt whether it falls under the concept; though for us men, with our defective knowledge, the question may not always be decidable. We may express this metaphorically as follows: the concept must have a sharp boundary. If we represent concepts in extension by areas on a plane, this is admittedly a picture that may be used only with caution, but here it can do us good service. To a concept without a sharp boundary there would correspond an area that had not a sharp boundary-line all round, but in places just vaguely faded away into the background. This would not really be an area at all; and likewise a concept that is not sharply defined is wrongly termed a concept. Such quasi-conceptual constructions cannot be recognized as concepts by logic; it is impossible to lay down precise laws for them. The law of excluded middle is really just another form of the requirement that the concept should have a sharp boundary. Any object $\Delta$ that you choose to take either falls under the concept $\Phi$ or does not fall under it; *tertium non datur*. E.g. would the sentence 'any square root of 9 is odd' have a comprehensible sense at all if *square root of 9* were not a concept with a sharp boundary? Has the question 'Are we still Christians?' really got a sense, if it is indeterminate whom the predicate 'Christian' can truly be asserted of, and who must be refused it?[80]

As we've seen, a large part of Frege's reconstruction of scientific discourse via his work in logic and logical notation involves removing ambiguity and vagueness, and insisting upon strict divisions among distinct concepts. Remember, too, that, for Frege, definitions establish a logical equality between the senses of signs, i.e., they establish that the sense of a simple sign is to be identical with the sense of a complex sign. If the sense of the complex sign were indeterminate or ambiguous, we would be unable to determine whether or not we were associating the proper sense with the simple sign. Thus, the simple sign could not properly

[78] See [Wussing 1984] for an excellent account of the various historical sources leading to the emergence of the abstract group concept.
[79] Frege extends these considerations on the definition of concepts to the definition of relations in §62 of [Frege 1893].
[80] [Frege 1893, §56]; [Frege 1970, 159].

function as an abbreviation for the complex sign and the definition would be entirely useless for the purposes of science.

The requirement that definitions of concepts be complete in this sense also leads Frege to reject the relatively common mathematical practice of 'piecemeal' definition. Piecemeal definitions, within mathematics, are those which (partially) define a concept for some limited domain and then employ this limited, partially defined concept to extend the definition to other domains. Thus for instance, we might define addition for the natural numbers and employ this definition to extend the 'same' concept of addition to the rational, irrational, real, and complex numbers. In a continuation of the train of thought quoted above, Frege attacks this form of definition:

> Now from [the requirement that definitions be complete] it follows that the mathematicians' favourite procedure, piecemeal definition, is inadmissible. [...] Such piecemeal definition is a procedure comparable to drawing the boundary of a part of a surface in bits, perhaps without making them join up. But the chief mistake is that they are already using the symbol or word for theorems before it has been completely defined—often, indeed, with a view to further development of the definition itself. So long as it is not completely defined, or known in some other way, what a word or symbol stands for, it may not be used in an exact science—least of all with a view to further development of its own definition.[81]

Thus, the practice of defining a term only for the natural numbers and, subsequently, extending it bit by bit to other number domains is forbidden by Frege insofar as, strictly speaking, the initial 'partial definition' of a concept is senseless. A definition is a stipulation that the sense of one sign is to be equated with the sense of another; senses are unambiguous entities which populate a realm distinct from the logical imperspicuity of everyday speech and thought.[82] Hence a definition which purports to provide a sign with an ambiguous or incompletely determined 'sense' is no definition at all, because there can be no such senses.[83] If all senses are determinate, as Frege holds, and all definitions are mere assignations of already-known senses to previously unused signs, then the ascription of an indeterminate or incomplete sense is simply nonsense and not a form of definition at all.

In addition to logical superfluity and completeness, Frege's stipulative conception of definition also commits him to the view that there can only be a *single* definition for any given sign. For consider the two possible consequences of a contrary view. In the first case there might arise two distinct definitions of the same sign, i.e., a single sign would be assigned two distinct senses. The use of such a sign would be ambiguous. But, according to Frege's conception of a properly scientific language, which is his only concern when discussing definition, such ambiguity must at all costs be avoided. Indeed, one of the

[81] [Frege 1893, §57]; [Frege 1970, 159-160].
[82] On the possibility of ontologically vague or ambiguous entities in Frege's thought, see [Ruffino 2003, 262-264].
[83] Frege notes other problems with piecemeal definition which I shall not discuss here (e.g., their effect on the validity of alleged theorems employing the term thus defined); cf. [Frege 1893, §61].

primary reasons for constructing a regimented language at all is to eliminate the ambiguity present in both ordinary language and contemporary scientific discourse. Accordingly, we cannot allow multiple senses for a single sign. In the second case, two distinct definitions might assign the same sense to a single sign. In this case, we have not ambiguity but redundancy. Even more than redundancy, we are forcing upon ourselves a duty which otherwise would not present itself, namely the duty of proving that the two supposedly consistent definitions are in fact consistent—or in the case of equivalence, that they are equivalent.[84] For these reasons, then, Frege's views allow for at most one definition of any given term.

## 2.5 Elucidations

It is by stipulating, arbitrarily, that the sense of the new simple sign is identical to the sense of a complex sign or group of signs that Fregean definitions provide signs with senses. Such definitions give new signs the same sense as complex signs whose senses are already known. Now, it can happen that the sense of such a complex sign was itself given by means of a definition, in which case we can ask after the meanings of the simpler signs in terms of which it, too, is defined. But, on Frege's view, such a chain of definitions must eventually come to an end. A science must start with some primitive signs whose senses cannot be given in terms of definitions because, in the initial stages of system-construction, there just are no simpler signs in terms of which the primitives can be defined.

Given that Frege is interested in creating a scientific language in which every sentence and every sign is meaningful, it follows that even the primitives of the system must already possess senses. The role of elucidations [*Erläuterungen*] is to guide the reader toward a grasp of the senses which the system-builder intends to express by the primitive signs of the system. It is to be assumed that the system-builders themselves are always able to grasp the senses of their system-primitives, and that these senses do not fluctuate. As Frege's translator Kluge notes, however, "even [elucidations] presuppose the intelligibility of at least some of the terms involved in giving them; hence not even here can we escape the necessity of simply presupposing [the senses of] some terms as known."[85] For Frege, elucidations cannot play a role within the system itself. Instead, like Euclid's second type of definition (e.g., his definition of 'point'), they are relegated to the propaedeutic, which clears the way for systematic thinking by preparing the materials out of which the system is to be built. Once these elucidations have achieved their function (i.e., once they have managed to guide the reader toward the senses intended to be expressed by the primitive signs of the system) they can be safely discarded and the construction of the system from its primitive building blocks can begin.

---

[84] "...the same thing may never be defined twice, because it would then remain in doubt whether these definitions were consistent with each other" [Frege 1893, §33.2].

[85] [Kluge 1971, xxvii-xxviii].

## 2.6 Axioms

Frege held a traditional view of axioms as unprovable truths which lie at the foundation of a particular deductive edifice. Axioms, unlike definitions, offer indispensible support when attempting to prove theorems, for different sets of axioms will generally speaking lead to different deductive consequences. On Frege's view, the axioms of geometry present us with a small group of unprovable truths about space which, when combined with suitable modes of inference and a grasp of some primitive geometric concepts, allow us to elaborate the entire edifice of Euclidean geometry. In a more general attack on the notion of a purely formal mathematical system, Frege writes the following of specifically geometrical science:

> To be sure, that on which we base our definitions may itself have been defined previously; however, when we retrace our steps further, we shall always come upon something which, being a simple, is indefinable, and must be admitted to be incapable of further analysis. And the properties belonging to these ultimate building blocks of a discipline contain, as it were *in nuce*, its whole contents. In geometry, these properties are expressed in the axioms insofar as they are independent of one another. Now it is clear that the boundaries of a discipline are determined by the nature of its ultimate building blocks.[86]

Here we see Frege's view of the purpose of axioms in relation to his understanding of deductive theories in general. We begin with a stock of basic concepts and objects (which we have, ideally, grasped via elucidations), and the properties/characteristics of these are given to us by way of a cluster of axioms, which are truths about those basic concepts and objects. In contrast with Hilbert, for whom the truth of geometric axioms was intimately connected to their consistency relative to a particular axiomatic system, Frege saw the truth of the axioms of geometry as truth *simpliciter*. These axioms, if true, were true in just the same way that any other everyday thought was true.

Here we find one of Frege's most fundamental disagreements with the geometric trends of the 19th century. For Hilbert and others like him, the truth of Euclidean geometry was not at all incompatible with the truth of various other non-Euclidean geometries. This is because the truth of an axiomatic system simply results from its internal consistency. Insofar as Euclidean, non-Euclidean, Archimedean, non-Archimedean, and many other forms of geometry are all internally consistent, then, they are all equally 'true'. Frege, in stark contrast, takes Euclidean geometry to be the science of space, so that its axioms are truths *about* space, truths which directly contradict the axioms of other candidate geometries like the hyperbolic geometry investigated by Bolyai and Lobachevsky. Insofar as Frege construed Euclidean geometry to be a theory about space, then, its truth is in direct contradiction with the claims of the various non-Euclidean geometries.

---

[86] [Frege 1971, 143].

## §3. The Frege-Hilbert correspondence

Now that we have established some of the context and content of the views of both Hilbert and Frege, we can turn more profitably to their correspondence. The correspondence and subsequent articles by Frege are an interesting focal point for examining differing conceptions of the nature of mathematical structures. Hilbert, for his part, offers a more strictly relational point of view regarding axiomatics than he does elsewhere in his own work. The viewpoint developed there is similar in many respects to the modern postulational point of view, and (some of) Frege's criticisms are illustrative of larger problems within the more general postulational point of view as well. Frege's viewpoint, too, despite its idiosyncrasies shares general features with many modern accounts of the notion of structure which seek to embed the purely relational concept of structure within a fixed hierarchy of some sort (e.g., logic, set theory, category theory, some form of axiomatized structure theory, etc.). It is against Hilbert's position that Frege's own views come into sharpest relief, and it is in direct contrast with Hilbert's quasi-relational views that we can best understand the problems related to the concept of mathematical structure which will interest us in subsequent chapters. Let us now turn to the debate.

### 3.1 The course of the dispute

After having met Hilbert at a scientific convention in Lübeck in 1895, Frege instigated a brief exchange regarding the uses and abuses of formalization in mathematics, a topic in which Hilbert had apparently expressed some interest when they met. This initial exchange quickly petered out, and it wasn't until December of 1899 that Frege wrote to Hilbert again, having recently received and worked through a copy of the latter's *Festschrift* on the foundations of geometry. Having found much confusion in the work, Frege wrote to Hilbert for clarification (particularly regarding his use of the terms 'definition,' 'explanation,' and 'axiom'), he also took the opportunity to acquaint Hilbert with his own views on these matters. Hilbert's response was prompt—dated just two days after Frege's letter—and contained several attempts to respond to Frege's request for clarification. He restated in more detail his own understanding of the nature and purpose of axiomatic systems and, towards the close of the letter, attempted to illustrate the advantages of his own (loosely model-theoretic) understanding over the traditional picture he believed to be held by Frege.

Frege replied on the sixth of January, thanking Hilbert for his letter, and beginning by rephrasing his newly clarified understanding of Hilbert's approach to axiomatization in general, and geometry in particular. After agreeing, in broad terms, with Hilbert's dismissal of earlier forms of 'genetic' or piecemeal approaches to axiomatization, Frege turns once more to the perceived lack of clarity in the specifics of Hilbert's approach. His criticisms here again pertain to Hilbert's use of definitions, explanations, and axioms. The tone of the letter is admonitory, and Frege describes at length his view of the proper approach to the subject at hand.

After this point, there were a few more exchanges between mid-January 1900 and November 1903, but Hilbert seems to have lost interest in the correspondence, much to Frege's chagrin. Frege, who was not at all satisfied with the results of their exchanges, suggested that their correspondence be published—a suggestion which Hilbert apparently rejected, as his views had changed and his focus had shifted in the meantime. Undeterred, Frege published a series of articles relating to their dispute; Alwin Korselt, a proponent of Hilbert-style axiomatics, took up the gauntlet and Frege's final two papers in the series are directed at both Korselt and Hilbert. I will now turn to Frege's criticisms of the views proposed, first by Hilbert, and then by Korselt.

### 3.2 Frege's criticisms of Hilbert

#### 3.2.1 Definitions and 'explanations'

One of Frege's central complaints against Hilbert pertains to the way in which he employs the terms *Definition* and *Erklärung*.[87] As we've seen, Frege's work in logic and the foundations of arithmetic had led him to erect a strict boundary between pre-systematic elucidations and definitions properly speaking. A Fregean elucidation is similar in kind to Euclid's definition of the term 'point': it is a pre-systematic signpost toward the proper meaning of the basic elements of the system. Once we have grasped these basic elements and begun to work within the system (of logic, arithmetic, geometry, etc.) we can dispense with elucidations altogether. Definitions, for Frege, are a rather different matter. A definition only makes sense within a particular system, insofar as it is a proposition which declares that a simple term (e.g., 'continuous') is to have the same meaning as a more complicated term, proposition, or conjunction of propositions. In this sense, definitions too are dispensable for Frege, as we could just as well employ the complex expression in each case—nevertheless the role of a definition only makes sense within a particular system, whereas the role of an elucidation is fulfilled in the anteroom through which we must pass before entering the system. Given the precision with which Frege employed the term 'definition', it is not entirely surprising that he found Hilbert's treatment of geometrical definitions somewhat unclear.

His initial criticism is that Hilbert does not adequately distinguish between his own use of the terms 'definition' and 'explanation'—some explanations seem, to Frege, to be identical in character to definitions properly speaking, while at other times the two seem to diverge significantly. In connection with this confusion, he refers to Hilbert's axioms of order (found in Hilbert's axiom group II). Frege writes to Hilbert that "I [..] cannot take your [definition of the concept 'between' via the axioms of order] for a definition, but you also do not call it a definition but an explanation [*Erklärung*]".[88] Indeed, Hilbert does employ both *Definition* and *Erklärung* throughout, and Frege notes that it is difficult to establish the precise

---

[87] It is useful to note that the German term '*Erklärung*' is very often used as a synonym of '*Definition*'. The point Frege makes against Hilbert is not that he uses the *term* '*Erklärung*' incorrectly, but that his use of the term seems to fluctuate, such that Frege is unsure whether Hilbert's *Erklärungen* are (what Frege himself would call) elucidations or if they are definitions properly speaking. At several points throughout the *Grundlagen*, it appears as though Hilbert's *Erklärungen* are different in kind from his more usual definitions, so that Frege's confusion regarding Hilbert's use of the term is not surprising, nor as pedantic as it might seem at first glance. (In what follows I will occasionally employ the English term 'explanation' in place of the German '*Erklärung*').

[88] [Frege 1980, 35].

difference between the meanings of the two terms, if indeed one is to be found. To this end, Frege highlights the many different ways in which the terms are employed in Hilbert's text. He points, for example, to Hilbert's explanations in §4 ('Consequences of the Axioms of Connection and Order') as instances where the term '*Erklärung*' seems to be employed as a synonym for '*Definition*'.

Here is the first of these explanations:

> Erklärung. Es seien *A*, *A′*, *O*, *B* vier Punkte einer Geraden *a*, so daβ *O* zwischen *A* und *B*, aber nicht zwischen *A* und *A′* liegt; dann sagen wir: die Punkte *A*, *A′* liegen *in der Geraden a auf ein und derselben Seite vom Punkte O*, und die Punkte *A*, *B* liegen *in der Geraden a auf verschiedenen Seiten vom Punkte O*. Die sämtlichen auf ein und derselben Seite von *O* gelegenen Punkte der Geraden *a* heiβen auch ein von *O* ausgehender *Halbstrahl*; somit teilt jeder Punkt einer Geraden diese in zwei Halbstrahlen.[89]

Now, to see Frege's point here, contrast this explanation with the following, from the previous section:

> Erklärung. Die Punkte einer Geraden stehen in gewissen Beziehungen zueinander, zu deren Beschreibung uns insbesondere das Wort '*zwischen*' dient.
> Explanation. The points of a line stand in a certain relation to one another, which can be described by the word 'between'.[90]

In the first instance, we witness Hilbert providing a precise statement of the conditions under which we are licensed to say that a specific relation holds; in the second instance, however, we are given a very broad, almost vacuous, statement which indicates only that we can use the word 'between' to describe some relation in which points of a line may stand to one another, though what this relation is, precisely, is not given by the explanation. Frege writes of the latter sort of explanations that they "are apparently of a very different kind [from the explanations of §4], for here the meanings of the words 'point', 'line' 'between' are not given, but are assumed to be known in advance. At least it seems so. But it is left unclear what you call a point".[91]

So, if we are to understand some of Hilbert's explanations on the model of definitions, then they are by Frege's standards quite bad definitions and, really, no definitions at all, insofar as they fail to clearly fix the references of the terms which are supposed to be defined. Recall that, for Frege, the role of a definition, properly speaking, is to stipulate that the meaning of a complex term or proposition is the same as the meaning of a newly introduced, simple term. Thus, for example, if Frege's work on the concept of number had panned out, in place of a host of complex statements about equivalence classes and courses of values, we would

89 [Hilbert 1903, 6].
90 [Hilbert 1903, 4].
91 [Frege 1980, 35].

have been able to plug in the much simpler term 'number' while still preserving the precise meaning of statements which employed the more complex combination of terms. Recall, too, that for Frege 'meaning' is understood in terms of his basic distinction between functions (or concepts) and objects, and that only sentences which express 'fully saturated' thoughts (in his sense) can be considered meaningful or possess a truth value. If the references of Hilbert's terms are not fixed by way of elucidations or definitions, then the sentences which contain those terms cannot express complete thoughts, and hence cannot be true or false by Frege's lights.

In stark contrast to Frege's aims, one of the key features of Hilbert's approach to axiomatic systems is to be found in the fact that the terms of those systems are *not* given a definitive, fixed interpretation. At most, an axiom system should be 'categorical' in the sense that any model of the system will be isomorphic to any other—thus, even a categorical axiom system should only ever determine a class of isomorphic models. For Hilbert, there appeared to be no point in moving beyond this class of models to a single, uniquely determined model, with a particular fixed set of objects (i.e., *the* Euclidean points, lines, etc.). Precisely the point of formalizing geometry was to escape the vagaries required by the elucidatory definitions of, e.g., lines as breadthless length—precision and generality were gained by escaping the demand for unique, fixed reference and examining the more important structural relations determined by a particular axiomatization. For Frege, however, the indeterminacy of reference here makes Hilbert's definitions useless. Frege asks 'What do you mean by point?' and by this question he means 'To what does your term 'point' refer?'. Hilbert, on the other hand, can only answer Frege's question by directing him toward the relations determined by the axioms of his system. The further requirement that he precisely fix the reference of his basic terms once and for all directly contradicts the very purpose of his axiomatization of geometry, as he saw it. Fixing the reference of the terms in more than a merely relational way would restrict the system to one particular interpretation, and, accordingly, greatly limit the applicability of that system to areas of inquiry not already intended by the developers of that interpretation. He writes in response to Frege that:

> [...] it is surely obvious that every theory is only a scaffolding or schema of concepts together with their necessary relations to one another, and that the basic elements can be thought of in any way one likes. If in speaking of my points I think of some system of things, e.g., the system: love, law, chimney-sweep... and then assume all my axioms as relations between these things, then my propositions, e.g. Pythagoras' theorem, are also valid for these things. In other words: any theory can always be applied to infinitely many systems of basic elements. One only needs to apply a reversible one-one transformation and lay it down that the axioms shall be correspondingly the same for the transformed things. This circumstance is in fact frequently made use of, e.g. in the principle of duality, etc., and I have made use of it in my independence proofs.[92]

We can see here that the idea of fixing the reference of his terms any more determinately than up to isomorphism seemed mathematically uninteresting to Hilbert. As we saw above, Hilbert's interest in geometry was not at all in the properties of a specific collection of

[92] [Hilbert to Frege, in Frege 1980, 40-41].

objects, but in the structural properties (consistency, simplicity, categoricity, etc.) of axiomatic systems themselves. He took a mathematical theory to be a schema which can be exemplified in a variety of different concrete systems. For instance, because there exists an isomorphism between the basic entities and relations of Euclidean geometry and those of the set of ordered triples of real numbers,[93] the consistency of Euclidean geometry follows from the (tacitly assumed) consistency of the theory of $\mathbb{R}^3$. But this model-theoretic, structure-focussed view of Hilbert's is predicated upon the belief that the axiomatic structure of Euclidean geometry can be discussed with complete indifference to any of its particular interpretations, be they number-theoretic or more traditionally geometric in nature. To demand to know which *things* were picked out by his axioms prior to interpretation seemed to him almost nonsensical and, at the very least, beside the point. From Hilbert's point of view, then, Frege's remarks seemed, first, to have missed the point of axiomatizing geometry entirely and, second, to have demanded a massive step backwards in mathematical thinking.

The conflict between two demands—Frege's for rigor based upon a precise understanding of certain logical concepts, and Hilbert's for mathematical fruitfulness based on generality and applicability—was sharpened considerably by Hilbert's reply to Frege's initial criticisms. In his very direct response, Hilbert writes:

> I do not want to presuppose anything as known. I think that with my explanation in §1 I give the definitions of the concepts point, straight line, and plane if one again adds to these all the axioms of axiom-groups I-V as characteristics. If someone is looking for other definitions of "point," etc., perhaps by means of circumscriptions like extensionless, then of course I would most decidedly have to oppose such an enterprise. One is then looking for something that can never be found because there is nothing there, and everything gets lost, becomes confused and vague, and degenerates into a game of hide-and-seek. If you prefer to call my axioms characteristics of the concept that are posited in the "explanations" and consequently exist, I should have no objections to this except, perhaps, that it contradicts the custom of mathematicians and physicists. Of course I must also be able to do as I please in the matter of positing characteristics; for as soon as I have posited an axiom, it will exist and be "true".[94]

Here Hilbert attempts to defuse Frege's objections regarding his apparently unclear concept of definition in two ways. First, he suggests that if Frege is looking for something like Euclid's definitions of 'point' and 'line', then he is barking up the wrong tree—such definitions play no real mathematical role and have, to Hilbert's mind, been rightly dispensed with in most modern mathematical works. Their obsolescence is due to their lack of clarity, the seeming impossibility of scientifically providing a pre-mathematical definition of point, and, most importantly, to the lack of intrinsic mathematical interest in uniquely fixing geometric concepts beyond the limits imposed by a categorical formalization. Hilbert's interest here is in the structure determined by the axioms of Euclidean geometry, and not in the nature of the entities which might serve as the reference for the basic terms of the axiomatization on one *particular* interpretation, traditional though it might be. Second,

---

[93] Either $\mathbb{R}^2$ or $\mathbb{R}^3$, for the cases of plane and solid geometry respectively.

[94] [Hilbert to Frege, in Frege 1971, 11-12].

Hilbert argues that if Frege still has problems with his understanding of definition, he can look to the axiom groups for answers: the full meaning of the concept of point, for example, is not given solely through the explanations and definitions of that term which are given by Hilbert. Instead, this meaning is given when those definitions and explanations are taken in conjunction with the axiom groups. If Frege is still unsatisfied by this characterization of the concept of a Euclidean point, then, to Hilbert's mind, he is simply asking for something which cannot (and, really, should not) be given.

### 3.2.2 Defining axioms

As we've just seen, in his attempt to respond to Frege's criticisms of his definitions, Hilbert suggested that, in addition to those propositions marked off by the headings 'definition' and 'explanation,' Frege should consider his axiom groups themselves as part of the definition of basic concepts like point, line, etc. Even more strongly, he writes that "[...] it is impossible to give a definition of a point in 3 lines, since it is only the whole axiom-structure that gives the complete definition, and therefore each new axiom alters the concept. 'Point' is always something different in Euclidean, non-Euclidean, Archimedean and non-Archimedean geometry respectively".[95] Two things are worth noting in this remark. First is Hilbert's view—dramatically different from Frege's—that axioms can play a role in the definition of the basic terms of a particular system. Second, is Hilbert's admission that different collections of axioms result in different basic concepts, such that the term 'point' is employed within different axiomatizations to designate quite different objects (i.e., Euclidean, non-Euclidean, Archimedean, and non-Archimedean 'points'). Both of these aspects of Hilbert's view have become a mainstay of the model-theoretic conception of axiomatics. In any introductory text in group theory or abstract algebra, for instance, the term 'group' is defined by the group axioms, and different types of group (Abelian, non-Abelian, Galois, cyclic, etc.) are given by altering the axiom set in various ways.[96] This procedure is now common throughout other branches of mathematics, various forms of geometry included. Frege, however, found Hilbert's particular version of this view logically imprecise at best, and an example of meaningless formalism at worst.

Frege himself viewed axioms as unprovable truths, and definitions as a kind of logical shorthand which helped to fix the meaning of new signs in relation to already established signs. Axioms and definitions thus play very different logical roles, and have quite different relations to the inferential systems of which they are parts. Fregean definitions are true by virtue of stipulation: we simply decree that a new sign is to have the sense of some other sign or combination of signs. An axiom, by contrast, was thought by Frege to be true in virtue of the way that the world is—in the case of geometry, by virtue of the nature of space itself. In his first essay dealing with the foundations of geometry, Frege writes of Hilbert's understanding of axioms that:

> Once the explanation [of the meaning of the term 'rectangle'] including the two axioms has been posited, the latter may be asserted as true; however, their truth [i.e.,

[95] [Hilbert to Frege, in Frege 1971, 12-13].
[96] [Scott 1987, 9ff.] and [Deskins 1995, 192ff.] are representative.

> the truth of the axioms] will not be founded on an intuition, but on the definition. And it is precisely because of this that no real knowledge is contained in [Hilbert's axioms]—something which undoubtedly is the case with axioms in the traditional sense of the word.[97]

The truth of Hilbert's axioms, as well as their meaningfulness, depends essentially upon the complex definitions of which they are a part. Because the supposed truth of Hilbert's axioms is not determined by an external check (like our intuition of space, some objective conceptual structure, or the very nature of physical space itself), Frege argues that the entire edifice constructed upon these axioms is empty and ill-founded.

Furthermore, since for Frege definitions are, at least in principle, always eliminable from any chain of logical inferences, they therefore ought not to serve as premises in a proof. Nor can they make possible proofs which are otherwise impossible. Axioms, by contrast, *must*, at some point, serve as the premises of some proofs within a system, as they are the foundation upon which the entire deductive structure of a particular system is based. Given that the axioms of a system must play some inferential role, it is important (for Frege at least) that they express complete thoughts.[98] But, if the sense of a term occurring within a sentence expressing an axiom is not fixed, then the axiom itself is not a complete thought, and thus can be neither true nor false. Thus, if each axiom or axiom group adds some new element in the definition of the basic terms of the system, it is only with an extraordinarily complex conjunction of all the axioms, explanations, and definitions that we get a complete thought. The result, Frege suggests, is a "logical edifice [which is] puzzling and opaque in the highest degree".[99] In "On the Foundations of Geometry" Frege remarks that

> Insight into the logical nature of a mathematical theory is frequently made more difficult by the fact that what really ought to be represented as a unitary propositional complex is torn apart into apparently independent grammatical propositions. This often happens for stylistic reasons, in order to avoid a propositional monstrosity; but this cannot be permitted to obstruct one's insight into the nature of the case. [100]

It seems clear that Frege believed Hilbert's stylistic choices to have greatly obstructed logical insight in this case. So, taken individually, Frege suggests that Hilbert's axioms do not possess a meaning, nor are they true or false, hence they can play no inferential role in the subsequent development of the science of geometry.

---

[97] [Frege 1971, 27].

[98] Moreover, the axioms must express *true* thoughts as Frege maintained that proper inferences can only be drawn from truths. In a posthumously published fragment, he writes: "Now if *A* is the concept of a right-angled equilateral pentagon, 'this *A*' is an inadmissible—a meaningless—proper name, and nothing whatever can be inferred either from the sentence 'This *A* has the property *B*' or from 'This *A* does not have the property *B*; for, strictly, we do not infer from sentences at all, but from thoughts. And only true thoughts are admissible as premises of inferences" [Frege 1906a, in PW, 177].

[99] [Frege to Hilbert, in Frege 1971, 17].

[100] [Frege 1971, 72].

### 3.2.3 Equivocations in Hilbert's independence and consistency proofs

Hilbert happily accepts that different collections of axioms help to define different basic concepts. Thus, despite the similarity between the Euclidean and non-Euclidean concepts of 'point', because these different geometrical structures are given by different collections of axioms, their basic notions are also different. Hence, theorems proven for non-Euclidean points are not the same as theorems proven for Euclidean points. Hilbert insists that the removal or addition of a single (independent) axiom results in a completely new system of basic concepts, a system which will have its own set of deductive consequences.

In his replies to this aspect of Hilbert's view, Frege notes two difficulties, both of which involve dangerous equivocations. The first sort of equivocation is most apparent in Hilbert's independence proofs. Recall that Hilbert's method of proving the independence of a particular axiom from a collection of other axioms involved modelling two distinct sets of axioms: one in which all the axioms including the independent axiom hold true, and one in which all the axioms save the independent one hold true. This is intended to prove something about a particular axiom as well as a particular set of axioms. But, notes Frege, using Hilbert's own reasoning, such proofs simply cannot achieve this goal. For, in the two separate models we are not working with the same basic concepts at all. By removing an axiom (as is required for the independence proof to go through) we are altering the *meaning* of the basic terms, if indeed they possess a meaning in Frege's sense at all. For example, if our intent is to prove the independence of the parallel postulate from the other axioms of Euclidean geometry, when we remove the parallel postulate from the larger set, the meaning of the basic terms occurring in the axioms (e.g., point, line, etc.) changes. But Hilbert's proof seems to require that the meanings of these terms remain stable throughout the proof, otherwise we are simply constructing two completely unrelated models. It seems, then, that Hilbert's own construal of his methods directly thwarts his aim of providing independence proofs. From Frege's point of view, addressing this problem requires either that Hilbert alter his methods such that different systems are not treated as one, or he must alter his goal of providing an independence proof of the *Euclidean* version of the parallel postulate from the remainder of the *Euclidean* axioms.

The second, more subtle, form of equivocation is best illustrated by Hilbert's consistency proofs. Here Hilbert constructs a single model of his axioms, i.e., an interpretation in which each axiom holds. The existence of a model is meant to show that the axioms *can* be true together, and are thus consistent. Since Hilbert takes his axioms to be a mere scaffolding whose basic terms do not refer to specific objects, from his point of view there appears to be no problem with this procedure. To see where the equivocation lies here, we must turn to Frege's idea of concepts of different levels. A first-level concept is one under which objects fall. For instance, all cows fall under the first-level concept '$x$ is a cow' and all Euclidean points fall under the first-level concept '$x$ is a Euclidean point'. Of first-level concepts, Frege writes that

> [...] between objects and (first-level) concepts there obtains a relation of subsumption: an object falls under a concept. For example, Jena is a university town. Concepts are generally composed of component-concepts—the characteristics. *Black silken cloth* has the characteristics *black, silken,* and *cloth*. An object falling under this concept has these characteristics as its properties. What is a characteristic with respect to a concept is a property of an object falling under that concept.[101]

Analogous to (though still distinct from) the relation of subsumption which obtains between first-level concepts and objects is the relation between second- and first-level concepts. Frege writes that in this case

> [...] one could speak of subsumption; but this relation, although indeed similar, nevertheless is not the same as that of the subsumption of an object under a first-level concept. I shall say that a first-level concept falls (not under, but) within a second-level concept. The distinction between concepts of the first and second levels is just as sharp as that between objects and concepts of the first level; for objects can never substitute for concepts. Therefore an object can never fall under a second-level concept; such would be not false but nonsensical. If one tried something like this linguistically, one would get neither a true nor a false thought, but no thought at all. [...] A different feature of first-level concepts is expressed by the proposition that if an object falls under such a concept, another object distinct from the preceding one also falls under it. Here we have a second concept of the second level. From both, as second-level characteristics, we can form a third second-level concept within which fall all those first-level concepts under which fall at least two distinct objects. The concepts *prime number, planet,* and *human being* would be such as fall within this second-level concept.[102]

Frege notes here that the meaningfulness of our language depends upon our subsuming entities of appropriate type under concepts of appropriate type. Thus, we can meaningfully state that a certain entity falls under a certain concept only if the first entity is one level lower in Frege's ontological hierarchy. Any confusion in this regard results in nonsense, which is neither true nor false. The second equivocation which Frege examines in Hilbert's work results from just this sort of confusion between concepts of different levels.

Hilbert's proof that Euclidean geometry is consistent, for instance, involves constructing an interpretation of the axioms of Euclidean geometry. To Frege's mind, this process of constructing a model already requires an equivocation between first- and second-level concepts. For instance, Euclid's first axiom—taken in the traditional sense—is a claim which pertains to objects (i.e., Euclidean points and straight lines). Similarly, Hilbert's model (constructed, either in the field Ω or within the real numbers) requires a discussion of objects, insofar as Hilbert treats 'points' in this interpretation as pairs of numbers. By illustrating that the real numbers can serve, given an appropriate interpretation, as a model of

[101] [Frege to Liebmann, in Frege 1971, 4].
[102] [Frege to Liebmann, in Frege 1971, 5].

Euclidean geometry, Hilbert intends to show that Euclidean geometry is consistent, under the reasonable assumption that the real numbers themselves are consistent.

What Frege takes Hilbert to have shown, however, is that both the Euclidean concept of 'point' and the concept of 'point' understood in terms of the real numbers or the field Ω fall under a distinct, second-level concept (which Hilbert also calls 'point'). More generally, Frege believes that Hilbert's consistency proofs pertain, not to Euclidean geometry and its first-level concepts, but to a system of second-level concepts which is entirely distinct from Euclidean geometry, but within which Euclidean geometry and Hilbert's interpretation of Euclidean geometry both fall.

Frege's reasoning runs as follows. Hilbert presents us with two systems of objects and first-level concepts: one in which we deal with points, lines, and planes, and the other in which we deal with numbers and sets of numbers. From the mere presentation of two systems of objects, we cannot reach the conclusion Hilbert wishes to make. Frege rightly notes that we must, at the very least, ascend to a higher level from which *both* the system of Euclidean geometry and the system of the real numbers appear as particular cases of a more general system. What Frege takes Hilbert to have shown with his consistency proof, then, is not the consistency of Euclidean geometry, but the consistency of a second-level system under which Euclidean geometry falls. And, most importantly for Frege, from the consistency of this more general theory, we cannot derive the consistency of Euclidean geometry itself. In one of his replies to Hilbert, Frege writes of this problem as it relates to the more limited case of individual propositions:

> If a general proposition contains a contradiction, then every particular proposition included under it will do likewise. Therefore, from the consistency of the latter, we can infer that of the general [proposition], but not vice versa. Suppose we proved that in a right triangle with two equal sides, the square on the hypotenuse is twice as large as that on one of the other two sides—which is easier than proving the general Pythagorean theorem. From this we can now infer that the proposition contains no contradiction; that the square on the hypotenuse of a right triangle having two [equal] sides is equal to the sum of the squares on the other two sides. From this we can deduce further that the general Pythagorean theorem contains no contradiction. But can we conclude still further that the Pythagorean theorem is therefore true? I cannot admit such an inference from consistency to truth.[103]

Frege's claims here hold more generally for systems of propositions as well. Let's see how this applies in the present case. There are three distinct systems at issue here. First is the system of Euclidean geometry itself, second is the 'model' of Euclidean geometry interpreted within the real numbers, and third is the more general system of propositions within which are included both traditional Euclidean geometry and Hilbert's model. For ease of reference, I will call this more general system *L*. Now, from the consistency of Hilbert's model, it does indeed follow that the system *L* does not contain a contradiction. It does not, however, follow that *all* of the particular systems which fall within *L* are consistent.

[103] [Hilbert to Frege, in Frege 1971, 21].

More specifically, it does not follow that Euclidean geometry itself is consistent. For Frege, Hilbert's mistake is to tacitly assume that his claims about *L* also apply directly to the Euclidean system. To Frege's mind, the claims made about 'points' in *L* have nothing whatsoever to do with Euclidean points, other than the unfortunate use of the same word to describe them. The dangerous equivocation, then, is found in Hilbert's fluctuation between the general system *L* and the specific system of Euclidean geometry which falls within *L*. By Frege's lights, it is only via this equivocation that Hilbert is able to 'prove' the consistency of Euclidean geometry. What he in fact proves is that *L* is consistent, given that the system of the real numbers is consistent.

A further criticism stems from Frege's non-relativized conception of truth, in which truth does not directly follow from consistency as it seems to for Hilbert. Hilbert takes his consistency proof for Euclidean geometry to be a sufficient proof of the *truth* of Euclidean geometry, whereas Frege, who views consistency as merely possible truth, does not at all authorize the inference from the merely possible to the actual. Here the crucial difference between Hilbert's point of view and Frege's is that Hilbert does *not* understand Euclidean geometry in Frege's more traditional sense. Instead, he views the axioms of Euclidean geometry as an uninterpreted conceptual schema which determines a formal structure. This structure (or scaffolding) can then be applied to a variety of different systems by way of the process of interpretation. Both Hilbert and his defender Korselt insist upon this. Korselt, for his part, emphasizes the potential emptiness of the uninterpreted concepts of a given formal theory. To this effect he writes:

> A purely formal system is noteworthy as long as the propositions that might occur in it do not lead to contradiction; indeed, it remains interesting even as a contradictory formal system, as long as this contradiction becomes apparent only at the end of a long sequence of inferences about the objects of the formal system. Therefore certain concepts of a "purely formal system" may even show themselves to be "empty". But on the other hand, a formal system may be "applied" to a *given* domain only after we have assured ourselves of the validity of the principles for that domain.[104]

Here Korselt expresses a key element of the postulational point of view: namely, that axiomatized mathematical theories are objects of investigation in their own right, and not simply tools employed to study their potential interpretations. To make contact with a concrete system, an *application* is required, and this application lies outside the mathematical theory properly speaking. This emphasis on the abstractness of formal theories is intended to illustrate that Frege's insistence upon fixing precise references for basic terms is not at all mathematically important, and that it is in a sense also misguided. For Korselt and Hilbert, a more precise characterization of the objects of a theory is only possible once we descend from the high generality of the axiomatic theory itself to one of its possible interpretations. Within a particular interpretation we can certainly ask after the precise reference of a given term, but from the more general formal perspective, this sort of question makes no sense.

[104] [Korselt 1903, in Frege 1971, 40-41].

From Frege's point of view, however, this formal stance requires a great degree of logical perspicuity which is not at all evident in either Hilbert's or Korselt's responses. From his own point of view, it appears impossible to claim (as Hilbert does), that we can prove the independence of, e.g., the *Euclidean* parallel postulate by means of these model-theoretic or 'formal techniques'. It is difficult for Frege to see, that is, how a proof pertaining to one system of objects and first-level concepts can show anything whatsoever about a completely distinct system of objects and first-level concepts without a clear understanding or presentation of the higher-level system which acts as their intermediary.

Moreover, if Hilbert or Korselt wish to highlight some of the features of a more general conceptual system (e.g., the system *L*), under which Euclidean and other systems might fall, their method of doing so is both misleading and unclear. In response to Korselt (who defends the apparent indeterminacy and ambiguity of Hilbert's presentation as a feature resulting from a formal point of view), Frege writes that, when referring to the increased generality of Hilbert's conception of a point, Korselt seems to have

> [...] in mind a second-level concept within which, aside from the Euclidean *point-concept* still other concepts fall. Of course this second-level concept must also be a completely determinate one; but it behaves toward the first-level concepts falling within it in a way similar to that in which a first-level concept behaves towards the objects falling under it. When we consider the multiplicity of these concepts of the first level (point-concepts), we get the notion that we are faced with an indeterminacy or ambiguity. This need not be the case here any more than it is in the case of the first-level concept *prime number*, where not only 2 but also 3 falls under the latter.[105]

In other words, if Korselt and Hilbert wish to discuss concepts from this higher point of view, that does not exempt them from making sure the reference of their basic terms is precisely fixed. A second-level concept is just as determinate and precise as a first-level concept. For Frege, rigor is *lost*, not gained, by the logical fiction of ambiguous or indeterminate concepts. What Frege objects to here is not the very idea of a higher level geometry—to which he is occasionally quite sympathetic[106]—but to the 'free-floating' character of that geometry as presented by Hilbert and defended by Korselt. Frege's logical view of the world involves (as we will discover in more detail in the following chapter) an ontological hierarchy within which each entity belongs to a distinct ontological level.[107] Hilbert's work attempts to avoid this hierarchy by characterizing a structure which is not determinately linked to any particular ontological level. To be sure, Hilbert does not intend his formal characterization to describe a concrete system of objects and first-level concepts. But neither does he intend his work to outline a system of second-level concepts. Instead,

---

[105] [Frege 1971, 68-69].

[106] He writes to Hilbert, for instance, that "At any rate, your idea of considering Euclidean geometry as a special case of a more inclusive theory is valuable even without [successful independence proofs]" [Frege to Hilbert, in Frege 1971, 15-16].

[107] Before our more detailed discussion in the following chapter it might be helpful to note that Frege's hierarchy begins with *all* objects at the lowest level, and then ascends to first-level concepts and functions, second-level functions, etc.

Hilbert's work describes a structure which is itself free from the fixed references of any one interpretation, but which may thus be applied to systems of a variety of levels. The usefulness of such a free-floating picture of mathematical structures is apparent in the many applications of, e.g., abstract group theory, which range from the study of the symmetry groups of various particles in quantum physics to the use of groups as a tool to understand other abstract algebraic structures (e.g., polynomial rings). As Frege would surely note, however, this apparent usefulness does not preclude the emptiness, vagueness, or even the inconsistency of the concepts at issue.

### 3.2.4 Blanchette on Frege's conception of analysis

Many commentators[108] on the Frege-Hilbert correspondence have sided against Frege in his estimation of Hilbert's project as an almost complete failure.[109] A common thread has been that Frege was short-sighted, too conservative, and blind to the obvious advances evidenced in Hilbert's work. Patricia Blanchette, in an illuminating paper on Frege, Hilbert, and consistency, outlines the reasons why this evaluation of their correspondence has seemed attractive:

> Frege was clearly of the view that principles of logic can be expressed as syntactic transformation rules; his own formal system used precisely the kinds of topic-neutral, syntactically specified rules [that Hilbert employed in his *Grundlagen*]. Given a set of sentences in Frege's *Begriffsschrift* notation, a Hilbert-style consistency proof will demonstrate that no contradiction is derivable from that set via Frege's own rules of logic. Hilbert-style independence proofs, similarly, would successfully demonstrate the nonderivability in Frege's system of one of Frege's sentences from a set thereof. It looks, therefore, as if Frege's assessment of Hilbert's demonstrations as "failures" is simply wrong.[110]

On this view, Frege's own work had set out to achieve the same goals as Hilbert's and by the same means, so that his rejection of Hilbert's views seems simply to have missed their point. But, as Blanchette notes,[111] one crucial difference between Hilbert's approach to consistency and Frege's is that Frege takes it to be the case that *thoughts* (and not sentences) are the things which are proven or disproven, and it is the capacity of sentences to represent the logical structure of thoughts which is the subject of analysis. Thus, the derivability or non-derivability of a *sentence* from a collection of other sentences is by no means the whole story

[108] See, for instance [Freudenthal 1962] and [Torretti 1984]. Freudenthal writes that "Frege, rebuking Hilbert like a schoolboy, also joins the Boeotians. (I have never understood why he is so highly esteemed today" [Freudenthal 1962, 618]; Torretti writes that "To demand like Frege that the meaning of [Hilbert's basic, uninterpreted concepts] be intuitively elucidated shows a lack of understanding of the nature of logical consequence that is indeed astonishing in the founder of modern logic [Torretti 1984, 251] and that "Frege's failure to understand abstract axiomatics comes out very clearly in his criticism of Hilbert's independence proofs" [Torretti 1984, 251]. More recently, Stewart Shapiro suggests of Hilbert's view that "Frege did not get it, or did not want to" [Shapiro 2005, 66].

[109] "Clever and inventive as [Hilbert's *Grundlagen*] is in many points, I think that it is on the whole a failure and in any case that it can be used only after thorough criticism" [Frege to Liebmann, in Frege 1980, 91].

[110] [Blanchette 1996, 322].

[111] [Blanchette 1996, 322-323].

for Frege. If we ignore this difference, then several of Frege's criticisms seem considerably weaker, pedantic, or even absurd. Keeping this difference in mind, however, we get quite a different picture.

Let's take a closer look at exactly what is going on here, beginning with an examination of Frege's understanding of logical analysis. As Blanchette points out in her recent book, *Frege's Conception of Logic*, the idea of analysis is central to Frege's overall logicist project.[112] So what exactly is analysis for Frege, why is it important for his logicist project, and why is it relevant for our present purpose of understanding his opposition to Hilbert?

While Frege's views of analysis shift in relation to his broader philosophical and logical development,[113] it seems clear that for him analysis is never merely a linguistic affair, but essentially involves the realm of thoughts (or conceptual contents). Analysis often involves the reduction (in some way) of a complex concept to simpler concepts. Often, such analysis is made quite easy by language. Thus, for instance, we might wish to analyze the concept '*x* is an orange house'. In this case, our language makes it very clear that the concept is likely not a simple one, and therefore an appropriate subject for continued analysis. Part of the reason the concept strikes us as complex is because of our use of the two terms 'orange' and 'house' in the expression above. A straightforward analysis of this complex concept would be to treat it as a concept under which fall all and only those entities which fall under both the concept '*x* is orange' and the concept '*x* is a house'. Here the underlying logical complexity is obviously represented by the linguistic complexity of our expression.

But language is not always so cooperative, and we can (and often do) represent complex concepts with linguistically simple terms. Mathematics, in particular, is filled with this sort of thing: the definition of the linguistically simple term 'group,' for instance, involves at least four logically complex axioms. This isn't a problem if the complexity is recognized by those employing the term (as is certainly the case for group theorists using the term 'group'). But we can also be misled into thinking that linguistic simplicity mirrors logical simplicity. Thus, for example, we might think that the basic notions of arithmetic or Euclidean geometry are logically simple when they might be 'simple' only relative to a given linguistic rendering of those theories.

As Blanchette points out, a significant portion of Frege's work involves illustrating that some of the apparently simple mathematical concepts of arithmetic are in fact complex. She gives a useful example of this phenomenon in relation to Frege's analysis of the successor function.

She suggests that, for Frege, the relatively simple statement that "One is the successor of zero" ought to be analyzed into the obviously far more complex claim that

[112] This is the key theme of [Blanchette 2012, chapter 1].

[113] See [Blanchette 2012, 7-23] for an examination of the role of analysis in Frege's works from the *Begriffsschrift* to the *Grundgesetze*.

(A) For some concept *F* and some object *a* falling under *F*, the extension of the concept equinumerous with the concept identical with the extension of the concept (equinumerous with the concept (not self-identical)) is the extension of the concept equinumerous with the concept *F*, and the extension of the concept equinumerous with the concept not self-identical is the extension of the concept equinumerous with the concept falling under *F*, but not identical with *a*.[114]

The point of Frege's analysis is to show that these two markedly different *statements* ultimately express the same *thought*. The real logical structure of the thought that one is the successor of zero is actually better represented by the more complex statement.

As Blanchette points out, this kind of analysis is crucially important for Frege's project of showing that arithmetic can be developed on the basis of logic alone. For much of Frege's project involves illuminating the hidden logical structure beneath the linguistic surface of arithmetic. But, as anyone familiar with Frege's technical work can attest, such analyses are rarely obvious, easy, or even intuitive. Now, this again prompts a certain question: why should we take the result of Frege's analysis (i.e., the statement (A)) to be a better representation of the actual logical situation we are interested in when we speak of the successor function? Why should we accept his analysis at all? Given their importance for his project, one would imagine that Frege would provide a thoroughgoing account of the nature of his analyses and give us reasons for accepting the accuracy of their results. But, as Blanchette notes,

> Notoriously […] he offers nothing of the sort. Though he repeatedly affirms the central importance of analysis, Frege never defines this notion, and—most importantly—never gives an account of the necessary and sufficient conditions for successful analysis. We are left to look at the details of the analyses he gives, and at his various scattered remarks about the nature of analysis, to determine what he took himself to be providing under the heading of "analysis" or "elucidation" and to assess the plausibility of his general approach.[115]

In the present context, our goal is not to provide a general account of Frege's notion of analysis. What is relevant for us here is how Frege's conception of analysis bears on his criticisms of Hilbert. There are two areas in his critique which seem to directly involve his broad conception of analysis. The first is in relation to Hilbert's consistency proofs, the second in relation to his independence proofs. (In the following section, we will also see that Frege's capacity to provide logically accurate analyses plays a significant role in his suggested emendations of Hilbert's system).

### 3.2.5 Blanchette on derivability and provability

Frege's claim is that if Hilbert wishes to employ a purely uninterpreted system (i.e., one with no semantics) then he may discuss relations of derivability between the sentences of his

[114] [Blanchette 2012, 24; emphasis removed].
[115] [Blanchette 2012,24].

system, but not between thoughts. For uninterpreted sentences do not possess any determinate reference and are thus incapable of expressing thoughts for Frege.

Now, none of this would matter if Hilbert's syntactic brand of consistency always exactly overlapped with Frege-consistency, i.e., if the relations between the sentences of his formal language perfectly mirrored those between thoughts. But, as Blanchette notes, this is not at all the case, and, for Frege, simply cannot be the case. Indeed, she writes that "The inconsistency of a set of thoughts despite the syntactic consistency [i.e., Hilbert-consistency] of a set of sentences expressing those thoughts is not simply a farfetched possibility for Frege. It is of the very essence of his views on analysis and thoughts that such cases are to be expected."[116] Even supposing that Hilbert's sentences somehow express thoughts—which, for Frege, was already quite a dubious supposition—the problem of equivocating between consistency at the level of sentences and consistency at the level of Fregean thoughts would have remained. The problem stems from the fact that sentences which are syntactically different can yet express the same thought (as we've just seen in relation to Frege's analysis of the successor function). Blanchette explains that

> Because each thought is generally expressible by a number of different sentences, there is much more to the relation of *provability* [which holds between thoughts] than is evidenced by the relation of *derivability* [which holds between sentences]. Where $p$ is a thought and $s$ a sentence expressing it, $\Pi$ a set of thoughts and $\Sigma$ a set of sentences expressing $\Pi$: the derivability of $s$ from $\Sigma$ guarantees that $p$ is a consequence of $\Pi$, but the fact that $s$ is not derivable from $\Sigma$ is no guarantee that $p$ is not a consequence of $\Pi$. For $s$'s nonderivability from $\Sigma$ is entirely compatible with the existence of some $s'$ and some $\Sigma'$ expressing $p$ and $\Pi$, respectively, such that $s'$ is derivable from $\Sigma'$. Relations between sentences of a well-designed formal system can always, for Frege, provide a positive test of provability and consequence, but never a negative test.[117]

(Blanchette's terminological distinction here between the relation of provability between thoughts and the relation of derivability between sentences is a useful one, which we shall employ in what follows). The key point here is that multiple different sentences can express the same thought. This is an essential component of Frege's philosophy of language and (allowing for differences of opinion regarding the nature of 'thoughts') a presupposition shared by many philosophers and non-philosophers alike. Simply put: there is, in general, more than one way to say the same thing.

But, importantly, the relation of derivability can only hold between particular groups of sentences. In the case at hand, derivability claims will pertain to the sentences of Hilbert's axiomatization in the *Grundlagen*. At best, for Frege, what Hilbert is capable of showing is that a given sentence $S$ is derivable from a sentence or collection of sentences within his formal theory, using the allowable syntactic transformations. But, given the fact that the same thought can be expressed in different ways, we can very easily imagine a different sentence, $S^o$, which expresses the same thought as Hilbert's $S$ but which is not derivable

[116] [Blanchette 1996, 325].
[117] [Blanchette 1996, 325].

from the same sentence or collection of sentences. For example, in the absence of definitions of the terms ‘sine’, ‘continuous’ etc. the sentence ‘the sine function is continuous at zero’ is not derivable from the sentence ‘the sine function is differentiable at zero’; with appropriate definitions (i.e., ones that adequately express the usual conceptual content), however, the derivation will go through. What this shows is that derivability and provability are different relations, and, subsequently, if Hilbert wishes to equivocate between the two, he will have claimed more than he has actually proved.

Frege’s criticism, then, is not simply that Hilbert’s sentences fail to refer for a whole host of reasons, but that, *even if they did manage to express thoughts*, still Hilbert would be incapable of proving what he believes he has proved. For Hilbert has not paid sufficient attention to the differences between derivability of sentences and the provability of thoughts. Ultimately, as Blanchette notes

> Frege was right to hold that the method of interpretations holds no promise for demonstrating what he took to be consistency and independence results. The point here is not merely terminological: Frege takes it that the consistency of a set of thoughts, and the independence of thoughts from one another, have important philosophical ramifications, none of which is demonstrable via Hilbert’s method.[118]

3.2.6 Independence and logical simplicity

We’ve seen why Frege’s views of analysis and provability led him to reject Hilbert’s interpretation of his consistency proofs (or, better, his derivations of consistency within his particular formal system). Let’s now turn our attention to Blanchette’s suggestion (in the above quotation) that there are important implications for Frege’s understanding of Hilbert’s independence proofs as well.

As we saw above in Frege’s example of the successor function, relatively simple expressions can go along with considerable logical complexity. Thus, we may very often be incorrect when we take linguistically simple expressions to straightforwardly represent logical or conceptual simplicity. The goal of many of Frege’s analyses is to bring out exactly this hidden logical complexity, so that its role in inferences can be directly examined. And, in fact, one of the chief goals of Frege’s logicist project is to unearth the hidden logical complexity of mathematics in order to prove that basic arithmetic can be founded on pure logic alone. Ultimately, this view has two implications. The first is that we can be wrong about logical simplicity: we may, for example, take a linguistically simple term like ‘number’ to express true logical simplicity, when in fact (at least according to Frege) it contains hidden logical complexity. And, second, the revelation of this logical complexity can allow us to prove things which we otherwise might not have been able to prove.

This view has direct implications for Hilbert’s independence proofs. These proofs, recall, have the aim of showing that “no essential part of any of these groups of axioms can be

[118] [Blanchette 1996, 336].

deduced from the others by logical inference."[119] Recall, too, that each of the axioms contains 'primitive' or (putatively) logically simple terms, e.g., 'point'. We can see the difficulty posed for these proofs by Frege's views on simplicity: if the putative logical simples of Hilbert's system are not in fact simple, but (upon further analysis) have additional logical structure, then it may in fact be the case that we *can* deduce some of the supposedly independent axiom groups from one another.

Thus, for instance, if, as Frege later suggests, the putatively simple term 'point' in Hilbert's axiomatization actually refers—if it is to refer to anything at all— then it refers to a second-level concept ranging over all first-level point concepts. Though he does not go through with the attempt, the implication is that such hidden complexity may very well allow us to prove the dependence of some of Hilbert's putatively independent axioms. Here, in the end, it seems that Frege's criticism ultimately amounts to the claim that he offers a better analysis of the actual logical situation than Hilbert does. As Blanchette has shown us, however, the success of this and other of Frege's criticisms relies on the *actual* superiority of his analyses—and Frege himself provides no systematic means of establishing this superiority.

In any case, as Blanchette's work helps us to see very clearly, Frege's criticism is not simply that Hilbert's sentences fail to refer for a whole host of reasons, but that, *even if they did manage to express thoughts*, still Hilbert would be incapable of proving what he believes he has proved (at least by Frege's lights). For Hilbert has not paid sufficient attention to the relations between derivability from sentences vs. provability from thoughts, nor of the linguistic relativity of his own notion of logical simplicity. Against Torretti's view of Frege as astonishingly blind to the nature of logical consequence, then, Frege's criticisms of Hilbert seem to indicate an extremely subtle appreciation of the logical relations between sentences and what they are meant to represent (i.e., thoughts in Frege's case)—one which even modern commentators have failed to notice. And it is just this kind of appreciation which Frege lambastes Hilbert for lacking when he claims more than Frege believes he is capable of proving.

### 3.3 Frege's proposed emendations of Hilbert's project

By 1900 Frege had already expended considerable effort developing his logical and semantic concepts: he had re-organized and formalized logic by way of his *Begriffsschrift* and had developed a detailed theory of meaning in a series of papers, most notably *Über Sinn und Bedeutung*. The result of his logical and linguistic work was a theory of meaning based upon the notion of a complete thought. For Frege, sensible discourse was only possible when our words manage to express what he called complete thoughts, which are those objective entities which can be either true or false. If geometry and mathematics are to be understood as sciences, then, Frege argued, they must aim at truth. Since truths are a particular kind of complete thought, the mathematical sciences (like all the others) must be composed of complete thoughts. Complete thoughts result only from the saturation or completion of a

[119] [Hilbert 1971, 32].

concept[120] by an entity of an appropriate level (i.e., an object if the concept is a first-level concept, or a first-level concept if the concept is of the second-level, etc.). For Frege, then, the very conditions of meaningfulness hinge upon carefully distinguishing between entities of different ontological levels. This is particularly evident in modern mathematical discourse, which deals with relations between entities of different levels more frequently than most other areas of human enquiry.

Both Hilbert and Korselt insisted upon the free-floating, schematic aspect of the uninterpreted axiomatic system of geometry presented in the *Grundlagen*. For Frege, with his careful semantic and logical distinctions, this system and its lack of fixed references could only result in meaningless discourse. Since meaning attaches only to complete thoughts, and complete thoughts require the combination of entities of appropriate ontological types, any system of sentences which ignores distinctions between entities of different ontological types must fail to deliver meaningful claims. Since Hilbert's suggested reading of his system of geometry as a mere scaffolding ignores precisely these kinds of distinctions, it must be meaningless in its current form, according to Frege.

Again, however, this is not to say that Frege rejected the notion that a higher-level system of geometry was possible at all. On the contrary, the idea that Hilbert's geometry might describe a higher level structure was perfectly acceptable for Frege, and he even expended a considerable amount of effort attempting to provide a clear logical foundation for just such a structure. His key objection to Hilbert, for our purposes, was that Hilbert's free-floating view required the description of a structure which was completely indifferent to any ontological level. Such a system would be meaningless for Frege. More importantly, without fixing the references of even 'structural' terms, Frege took it to be the case that Hilbert's geometric work (e.g., his independence and consistency proofs) ultimately relied on a kind of logical sleight of hand whereby claims pertaining to a determinate system of second-level concepts were substituted for those pertaining to an equally determinate first-level system, all under the guise of an uninterpreted structure which was supposedly free from any particular ontological level whatsoever.

Though Frege was dubious of the possibility (and usefulness) of successfully altering Hilbert's picture of geometry to answer his criticisms, he nevertheless attempted to provide a reformulation of this picture in terms that he himself could accept. This reformulation entailed embedding Hilbert's supposedly free-floating mathematical structures within his ontological hierarchy. This embedding first required, however, a method of giving a determinate sense to Hilbert's axioms, definitions, and theorems. In order to overcome this rather serious stumbling block, Frege suggested a way in which Hilbert's 'pseudo-propositions' could be reformulated to express complete thoughts, and thereby attain the truth values required of a proper science. Let us examine how he believed this could be done.

---

[120] More generally, complete thoughts result from the saturation of a *function* of a certain kind by one or more entities of the appropriate type(s); Frege viewed (first-level) concepts as functions which take objects as arguments and deliver truth values as outputs. I will discuss Frege's ontological hierarchy in more detail in the following chapter.

For Frege, pseudo-propositions result from the tendency to break up conditional propositions into antecedent and consequent parts, neither of which expresses a complete thought. Take the following conditional proposition, for instance:

If $a$ is a whole number, then ($a$ x ($a$ – 1)) is an even number.[121]

This propositional complex expresses a complete thought, and utilizes the dummy letter '$a$' as a means of indicating generality. Now, we can—and often do—split up such conditional propositions for stylistic, pedagogical, or other reasons. For example, the above statement can be split up into two pseudo-propositions as follows:

$a$ is a whole number.
($a$ x ($a$ – 1)) is an even number.

In the complete conditional propositional complex above, we express the thought that if a whole number is of the form ($a$ x ($a$ – 1), then it falls under the concept of being even. Here the sign '$a$' does not refer to a specific number, but serves only to lend generality to the proposition; it allows us to claim that *any* whole number of the appropriate form possesses the property of being even. (Frege says that such signs *indicate* but do not refer). It is important that, without the stipulation that $a$ is a whole number, the phrase '($a$ x ($a$ – 1)) is an even number' does not express the same thought, or, indeed, any thought at all. Without carrying over the antecedent condition that $a$ is a whole number into the consequent claim that ($a$ x ($a$ – 1) is an even number, we fail to express a complete thought. For in both isolated pseudo-propositions, the apparent function of the dummy letter '$a$' is not to lend generality to the thought expressed. In the consequent pseudo-proposition, for instance, its function (or at least its purported function) is to refer to an object and predicate of that object that it possesses the property of being an even number. But the expression '($a$ x ($a$ – 1))' does not refer to any specific object and so the truth or falsity of the pseudo-proposition remains indeterminate for Frege. The result of splitting up a conditional proposition into separate expressions of its antecedent and consequent components is not two propositions, each possessing a truth value, but two pseudo-propositions, neither of which is a complete thought and neither of which can be true or false. Frege writes that

> [...] we must not let ourselves be deceived by the fact that for stylistic reasons, an antecedent pseudo-proposition occasionally occurs in such a form that upon cursory examination it appears to be an explanation of one or more letters. For in fact neither these putative explanations nor the proposition in which the conclusion is asserted are real propositions. Rather, being antecedent and consequent pseudo-propositions, they belong inseparably together, so that only the whole constituted of them is a real proposition. It would greatly facilitate insight into logical structure if what is a single real proposition according to its subject matter were also a unitary propositional complex according to its grammar and did not break down into independent propositions. To be sure, in our word-languages such propositional complex[es] would

[121] This example is found in [Frege 1971, 72].

> sometimes attain a monstrous length, whereas, because of its perspicuous nature, the Begriffsschrift is better suited to the representation of the logical fabric.[122]

Frege came to the conclusion that something like the confusion between real and pseudo-propositions occurs in Hilbert's (and Korselt's) attempt to outline a formal axiomatic system for geometry. The logical structure of Hilbert's axiomatic system is greatly obscured by his use of the familiar terms 'point,' 'line,' 'lies on,' 'between,' etc. By employing these familiar terms, Frege argues that Hilbert unwittingly engages in a dangerous equivocation. At one moment these basic terms are to be understood as referring to specific entities (the concepts and objects of traditional Euclidean geometry, or their correlates within Hilbert's model of Euclidean geometry within the real numbers). At another moment, however, these same terms are intended to function as purely uninterpreted dummy signs, meant only to lend generality to the propositions in which they occur. Through this confusion the fact that Hilbert's claims are either meaningless or possess quite different meanings than he intended is simply glazed over, and the real logical situation quite obscured.

Frege, though clearly critical of Hilbert's endeavour, attempts to clarify the logical situation by employing the notion of a pseudo-proposition. Instead of Hilbert's 'defining axioms'—which, by Frege's lights, do not define at all, as they do not fix the references of their respective *definienda*—we have a system of antecedent pseudo-propositions, none of which expresses a complete thought. Similarly, in place of the theorems Hilbert proves on the basis of his axioms and definitions, we have a system of consequent pseudo-propositions, none of which, again, expresses a complete thought. A theorem or axiom, taken in isolation, does not express any thought insofar as it contains signs which do not have a fixed reference. Moreover, taken in isolation, these empty signs do not lend generality to the pseudo-propositions within which they occur any more than did the sign '$(a \text{ x } (a - 1)$' in our previous pseudo-proposition.

Frege's method for salvaging Hilbert's work from this type of meaninglessness involved combining his axioms and theorems into appropriate propositional wholes which expressed complete thoughts, and whose terms properly indicated or referred. Unfortunately, in the case of Hilbert's work, these wholes are exceedingly complex: for in order to give sense to a particular theorem (consequent pseudo-proposition) we must combine it with an antecedent pseudo-proposition which contains all the relevant 'axioms' for the system we are interested in. In other words, each actually meaningful theorem will contain, as an inseparable part, an entire collection of 'definitions' and axioms connected together via conjunction. Of course, expressing this complex logical edifice in ordinary language would be extremely unwieldy—hence Frege's recommendation of his own, more perspicuous, logical formalism in the quotation above. For any theorem of Hilbert's, reformulated according to Frege's specifications, we will have a conjunction of conditions (the axioms) followed by the statement of the appropriate theorem. Ideally, once we've done this, the appearance of a purely relational, or referentially indeterminate system should disappear.

[122] [Frege 1971, 73-74].

Let's examine how this would work in the case of the first 'theorem' (really three theorems) of Hilbert's *Grundlagen*. These are stated as follows:

> THEOREM 1. Two straight lines of a plane have either one point or no point in common; two planes have no point in common or a straight line in common; a plane and a straight line not lying in it have no point or one in common.[123]

Now, in the traditional Euclidean context, these theorems are perfectly meaningful and require no further interpretation. They describe first-level relations amongst objects, and the claims they express are inferred directly from the (ostensibly true) axioms. Importantly, too, the basic terms of traditional Euclidean geometry are supposed to refer to particular objects and concepts: we grasp the reference of a particular term by means of definitions or (what Frege would call) elucidations. Hilbert's view is quite different, as is Korselt's; for both, the theorem does not pertain (or at least does not solely pertain) to particular objects or concepts. Instead of referring, the basic terms of Hilbert's system are intended to lend generality to the propositions in which they occur, such that these propositions can (by means of a suitably outlined interpretation) hold true of a variety of domains. Until such an interpretation is outlined, however, these theorems are of a purely formal character, having no determinate reference and no need of one.

Frege was quite opposed to the notion that purely formal, or completely uninterpreted signs could be of any use in science—as science concerns itself with truth, and truth adheres only to complete thoughts. Moreover, he argued that the appearance that these terms are actually completely uninterpreted in Hilbert's work was simply an after-effect of poor terminological choices. More specifically, Frege argued that the use of the traditional Euclidean terminology (e.g., 'point,' 'lies on,' etc.) in the place of proper variables for the purely formal system had the effect of obscuring the true logical status of Hilbert's claims. By replacing this misleading Euclidean terminology in the first of the above theorems with appropriately differentiated variables ('$x$,' '$y$,' and '$z$' are variables for objects; '$P$,' and '$R$,' variables for relations between objects), we get something like the following:

> (H1) Two $x$ standing in relation $P$ to $y$ either stand in relation $R$ to one $z$ or to no $z$.

Here it becomes quite apparent that, without further clarification of the references of the terms at issue, we not only do not have a theorem, we have no thought at all. When the purported theorem is more perspicuously laid out to illustrate its generality, it becomes quite clear (as Hilbert would accept) that, unlike Euclidean geometry, the basic terms of Hilbert's system do not seem to have fixed references. But neither is it clear exactly how (or if) they serve to lend generality to the theorem. Frege writes that "here, unless we are to understand the words in the Euclidean sense, just about everything is unknown".[124] Frege's contention is that, either we make the Euclidean claim and Hilbert's work does not have the import it is supposed to, or we accept that the basic terms are intended to lend generality, in which case the result is at best an incomplete or partial thought, and at worst no thought at all.

---

[123] [Hilbert 1950, 3].

[124] [Frege 1971, 90].

Clearly Hilbert's intention is not to make the traditional claim that two straight Euclidean lines which lie in the same Euclidean plane either have one Euclidean point or none in common. Rather, his intention seems to be to claim that for any specific system which satisfies his axioms on a particular interpretation, the theorem (whose terms will have references tailored to the interpretation) will come out true. Frege's ultimate point is that statements like this do not require—and indeed are not possible on the basis of—indeterminate reference. Instead, they are statements which (when suitably reworked along his lines) make very clear and fixed reference to second-level concepts and the first-level concepts which fall within them. His re-imagining of Hilbert's work involves construing the supposedly purely formal system as one containing clusters of pseudo-propositions, themselves meaningless, which can be made to express genuine thoughts by proper combination. Once this has been achieved, the resulting propositions do not contain indeterminately referring terms but, instead, terms that refer to second-level concepts and the first-level concepts which fall within them. Thus, for instance, supposedly 'uninterpreted' statements about points, lines, etc. are taken to express determinate claims about second-level 'point-' and 'line-' concepts, i.e., concepts within which fall first-level concepts like those of traditional Euclidean geometry or Hilbert's interpretation of the real numbers. In place of Hilbert's own statement of his first theorem, then, we obtain a lengthy conjunction of axioms and definitions as antecedent pseudo-propositions combined with the re-configured statement (H1). By characterizing Hilbert's axioms as conditions pertaining to a system of second-level concepts, Frege gives them a fixed reference. By including these conditions as the antecedent portion of every theorem, he also makes it possible for the variables in (H1), for instance, to properly lend generality to the complete second-level proposition. Finally, in place of interpretations, *inferences* are made from these general propositions to particular propositions, e.g., from second-level claims about *all* point-concepts to claims about particular first-level point concepts.

In Hilbert's work (and others like it), many geometers believed they had found the answer to their questions about the nature of geometry. Geometry was a science which studied the deductive elaboration of particular types of axiomatic structures, irrespective of their relationships to physical space. One underappreciated effect of Frege's criticisms of Hilbert (and Korselt) was to draw in sharper relief the shortcomings of the purely relational or free-floating understanding of geometrical structures. In Frege's opinion, the answers offered by Hilbert were merely chimerical, and relied upon various logical confusions in order to maintain the appearance of working properly. In place of this relational view, Frege offered a picture wherein various mathematical structures are embedded within a single ontological hierarchy. One key benefit of such a view was an increase in the precision of claims regarding the various relations between different structures. Thus, for instance, by careful attention to distinctions between Hilbert's axiomatic theory and its different possible 'interpretations', Frege is able to draw out the logical connections (or lack thereof) between the second- and first-level propositions of these respective theories. Of course Hilbert would not agree with Frege's view that his independence proofs were useless, but (as his later work would show) the logical perspicuity Frege brought to the table was an early indicator to Hilbert that understanding the logical relations between structures required a careful and attentive eye. In the following chapter, I will begin to examine problems with Frege's own attempt to characterize his ontological hierarchy and employ it as a means of explaining the nature of (certain elements of) mathematics.

# Chapter 3: Problems with Frege's views

## §0. Introduction

In the previous chapter we examined a number of problems pertaining to Hilbert's understanding of the nature of mathematics and mathematical structures. Hilbert's account of axiomatics was predicated on the view that we needn't fix the references of the basic terms of an axiomatic system in order to prove theorems in that system. As we saw, however, Frege raised a number of difficult philosophical questions for Hilbert. These chiefly pertained to Hilbert's purported inability to make it clear precisely what he is talking about in his discussions of mathematical structures like those required for his independence proofs. As we saw in that chapter, Frege's criticisms are themselves predicated on the view that mathematics (and, indeed, any meaningful speech whatsoever) is *about* something in particular, i.e., that it have a determinate content. While Hilbert's views took it for granted that we could speak about a purely uninterpreted system (or 'scaffolding of concepts,' in Hilbert's words), the basis of Frege's attack was his understanding of the importance of fixity of reference and determinately circumscribed concepts. These criticisms of Hilbert seem to require that Frege himself have a clear and defensible understanding of how it is that we can fix the references of (at least) mathematical terms.

In this chapter we will examine problems relating to Frege's own attempts to fix the reference of mathematical and logical terms within his ontological hierarchy. The first such problem is the so-called concept 'horse' problem. While this problem can be solved using resources available to Frege, the solution depends entirely upon his ability to clearly pin down the references of basic terms within his ontological hierarchy. The second problem, the Julius Caesar problem, is rather more difficult, and (I will argue) Frege never provides an adequate solution to it. This problem cuts to the heart of Frege's fixed-reference theory of mathematical structures, and, as we shall see, his inability to solve this obscure difficulty makes many of his criticisms of Hilbert fall flat. In the end, despite Frege's best efforts, I will argue that his account fails to explain how we can pin mathematical structures down within his ontological hierarchy of objects and functions.

Moreover, in the final sections of the chapter, I will argue that the inability of Frege to fix the reference of certain mathematical concepts is illustrative of a more widespread difficulty still facing the philosophy of mathematics today. While the specific forms of the problems faced by Frege are not necessarily faced by the contemporary philosophy of mathematics, the general problem of fixing the references of mathematical terms (or explaining why we don't need fixed reference) is shared by all. The upshot of this chapter, when combined with my analysis of Hilbert's views, is that neither the free-floating account of structures nor the general view of structures as embedded within a single ontological hierarchy is satisfying, and that neither picture is capable of satisfactorily explaining what mathematics is about.

Let's begin with an examination of Frege's ontological hierarchy.

## §1. Frege's Ontology

For Frege, the entities[1] proper to logic have a special ontological status, and are not reducible to (or otherwise eliminable in terms of) linguistic, mental, or physical entities. In particular, Frege takes the semantic content of meaningful speech and thought to possess independent ontological status, such that the things which are true or false (what he calls 'thoughts') are not dependent upon the human mind or even the empirical world. This is a central aspect of his logical views, and one of the key theses of early analytic philosophy.[2] Though Frege often insisted upon the strict separation of the science of logic from psychological concerns, his fundamental ontology is, by and large, indifferent to the distinctions between the mental, physical, and logical realms. Numbers, chairs, and ideas, for example, all occupy the same basic ontological niche: they are all objects. While Frege's ontology is not in this sense ultimately divided into the physical, mental, and the logical, it is not therefore unstructured. In fact, it is hierarchically organized into an infinite number of strictly separate levels. In what follows I will examine the organization of this hierarchy.[3]

### 1.1 Objects

At the most basic level of Frege's hierarchy, there are *objects*. In addition to the everyday entities we are used to calling objects (e.g., tables, dogs, windows, etc.) Frege includes a number of less familiar entities under this heading. For example the extensions of concepts, the True, the False, numbers, thoughts, ideas, etc. are all taken to be objects. Despite the extreme generality of the category, there are some characteristics that all objects have in common which distinguish them from other types of entity. On the linguistic level, for instance, objects are referred to by way of proper names or declarative sentences; this distinguishes them from functions and concepts, which are only (properly) referred to via a grammatical predicate or function-name. Another important (if somewhat metaphorical) way in which Frege distinguishes objects from other entities is through their completeness, self-subsistence, or 'saturation'. In contrast to other entities (i.e., functions), objects do not require and, indeed, do not accept any supplementation or saturation from other entities. They are already complete entities. In a slightly less metaphorical sense, objects have the property that they can only combine with other entities (objects or functions) if there is some function serving as combinator or tie. Frege writes, for instance, that "An object—e.g., the number 2—cannot logically adhere to another object—e.g., Julius

---

[1] A terminological note: here, and throughout the thesis, I employ the term 'entity' to indicate only the barest ontological status. All entities possess being, in some sense, but not all entities are 'things' or 'objects'. I reserve these latter terms to refer to a specific type of entity. Thus, I take functions or concepts (as Frege construes them) to be entities but not objects.

[2] In his earlier *Wissenschaftslehre* [Bolzano 1837], Bolzano presents a very similar view. There Bolzano employs the term 'proposition in itself' [*Satz an sich*] for the objective content of our spoken, written, or thought propositions. Bolzano also held that, though *there are* such propositions in themselves, they are not actual, and, further, that the being of these propositions did not depend at all upon their being thought, or even upon their being thinkable. For a comparison of Frege's 'thoughts' with Bolzano's 'propositions in themselves', see [Künne 1997]. G. E. Moore and Bertrand Russell, too, dealt with entities similar in some respects to Fregean thoughts (see, e.g., [Moore 1953] and [Russell 1903]).

[3] [Wells 1951, 542] gives a succinct table which partially articulates the structure of this hierarchy, though it fails to include functions of more than two variables. [Mendelsohn 2005, chapter 5] also develops a type-system for Frege's ontology.

Caesar—without a binding agent. But that may not be an object, but must be something unsaturated. A logical combination can only become a whole by having an unsaturated part saturated or supplemented by one or more parts".[4] Though slightly misleading, we might picture objects as the bricks required to build a logical edifice (a thought), and functions as the mortar binding them together—without both components, we end up with something less than the building we sought.

## 1.2 Thoughts, subjective and objective

For Frege, a very important distinction amongst objects is that between thoughts [*Gedanken*] and what we might call subjective thoughts. The latter, which most of us would call simply 'thoughts', are possessed by individual thinkers. We *have* subjective thoughts, and they, like us, come into and pass out of being. These thoughts are more like ordinary objects (mice, rugs, etc.) than the abstract entities Frege calls thoughts. Depending upon one's theory of mind, such subjective thoughts may or may not be physical in character. For Frege, in any case, it does seem to be the case that the subjective element of thinking is limited (in humans at least) to the mind of an individual thinker, such that no one other than myself can have direct access to my subjective thoughts, for example.

To understand why Frege posits the existence of non-subjective thoughts, we need only recall that it often seems to be the case that two individuals share the same thought—perhaps we both think that 2+2=4 or that it is a beautiful day. Though our internal picture or subjective grasp of these thoughts may differ significantly—I may, for instance, imagine four oranges arranged in a certain pattern, while you may imagine a sequence of squares when thinking that 2+2=4—there is in these claims some content that we seem to share. The possibility of sharing thoughts (or at least part of their content) is central to human communication in general, and crucial to the progress of science more specifically. Frege's belief was that the objectivity of scientific and mathematical truth (which he did not question) required that there be something over and above our individual, subjective thoughts which guaranteed the possibility of communication. In his many arguments against psychologism in logic,[5] for instance, Frege contends that, if we restrict thoughts to the minds of individuals, then there can be no rational adjudication between seemingly contradictory claims. If, for example, you hold that 2+2=5 and I hold that 2+2=4, the appearance of contradiction between our views evaporates if our understanding of the meaning of terms like '2' and '=' is limited to our own minds. There will be *your* 2 and *my* 2, each possessing its own distinct properties, and there will be no possibility of our deciding which of our statements is true, *simpliciter*. Yours will be 'true for you' while mine will be 'true for me'. As a succinct statement of Frege's objections against this type of psychologism, take the following two sentences: "A criterion for whether a mode of connection constitutes a thought is that it makes sense to ask whether it is true or untrue. Associations of ideas are neither true nor untrue".[6]

---

[4] [Frege 1903, 372].

[5] For a clear exposition of Frege's position against psychologism, as well as a fair presentation of the doctrine itself, see [Resnik 1980, 25-53]. The first nine chapters of [Husserl 2001] also contain an extremely detailed account and refutation of the works of the psychological logicians Frege inveighed against in his own work.

[6] [Frege c. 1906b, 174]. It is worth noting that Frege does not think that judgments are simply associations of ideas.

Because Frege believed in the objectivity of scientific truths, he argued in favour of a realm of objective, non-empirical, atemporal entities which are represented or grasped by individual thinkers with the aid of various tangible devices like spoken and written language or mental processes. In one of the most unfortunate terminological decisions in philosophical history, Frege chose to call these entities 'thoughts' [*Gedanken*].[7] These objective thoughts differ markedly from their subjective cousins. For one, no one 'thinks' a Fregean thought—instead, as he says, we *apprehend* or *grasp* them. These thoughts are not the possessions of the individuals who apprehend them, nor do they come into or pass out of being. For Frege, objective thoughts do not even require the possibility of being grasped by a thinker in order to possess being. They are, in the strictest sense, independent entities, with properties quite distinct from those of subjective thoughts.

Now, such remote entities would be entirely useless for the purposes of science if humans, as the limited beings we are, could not somehow encounter them. Frege's claim is that we *grasp* these entities and, through various media (e.g., mental processes and languages) we are able to represent and communicate their structure to ourselves and others. The nature of the process of apprehension is never satisfactorily clarified by Frege and at one point he refers to it as "perhaps the most mysterious of all"[8]. Nevertheless, after grasping a thought we are, somehow, capable of representing or expressing aspects of its structure by way of language or subjective thinking. The idea seems to be that certain tangible objects (e.g., written inscriptions or mental ideas) represent or express these intangible, causally inert[9] thoughts, which are then apprehended. Natural languages like German and English help us to grasp thoughts by mimicking their logical structure as well as the inferential relations between them. But often they do so only imperfectly, such that language is as much a barrier to the representation of thoughts as it is a facilitator. The intention of Frege's *Begriffsschrift* was to construct a formal language which would help eliminate the logically extraneous elements of natural language so that only the inferentially important elements of the realm of thoughts were represented. Unfortunately, the relation of

---

Rather, by the 1890s, at least, Frege understands judgment to be a mode of connection which does involve truth. He thinks of judgments as the admission or acknowledgement [*Anerkennung*] of the truth of a thought, and also as an advance "from a thought to a truth-value" [Frege 1892b, 35; in Beaney 1997, 158]. But, importantly here, the judgment itself is not true, it is only a mental act whereby the truth-value (a non-mental entity) of a thought (also a non-mental entity) is admitted.

[7] I will occasionally refer to Fregean thoughts as 'objective thoughts' when the distinction between thoughts and their subjective, mental counterparts is at issue or in cases where the term 'thought' seems especially misleading.

[8] [Frege 1914; in Beaney 1997, 246].

[9] Though Frege insists upon the non-empirical nature of his 'thoughts', in a later paper he seems inclined to attribute them at least some form of causal efficacy. Frege himself was troubled (as are many contemporary philosophers of mathematics) by the deep epistemological problems posed by purely abstract entities, as can be seen in the following passage: "A thought, admittedly, is not the sort of thing to which it is usual to apply the term 'actual' ['*wirklich*']. The world of actuality is a world in which this acts [*wirkt*] on that and changes it again and undergoes reactions itself and is changed by them. All this is a process in time. […] And yet what value could there be for us in the eternally unchangeable, which could neither be acted upon nor act on us? Something entirely and in every respect inactive would be quite unactual, and so far as we are concerned it would not be there [*nicht vorhanden*]. […] Thoughts are not wholly unactual but their actuality is quite different from the actuality of things. And their action is brought about by a performance of the thinker; without this they would be inactive, at least as far as we can see. And yet the thinker does not create them but must take them as they are" [Frege 1918, 76-77; in Beaney 1997, 343-345]. (See [Dummett 1982] for a thorough analysis of Frege's use of the term '*wirklich*').

representation or expression between the causal realm of written and spoken inscriptions and the causally inert logical realm, though central to Frege's project of developing a logically perspicuous language, is very little explored or explained by him. What we do get from Frege's remarks on the topic is that language can help us to clarify the logical structure of the realm of thoughts, but that it is rarely, if ever, a perfect mirror of that structure.

Despite the flaws in language, it can be a useful logical tool. All thoughts are structured, and Frege seems to think that their structure is best represented by declarative sentences (perhaps because these sentences seem to express truths and falsehoods most directly). Frege's interest is in logic, and hence in the *logical* structure represented by declarative sentences. This is not to say that thoughts do not possess non-logical structure; simply that, in his work, Frege is only interested in thoughts construed as logical entities. For him logic is that science which is concerned primarily with the most general laws of truth.[10] Hence, he is interested not in the emotional or aesthetic qualities of thoughts, but primarily with their relationship to truth. In his logical work, Frege characterizes thoughts specifically as those entities which are either true or false; whether or not a particular thought is true or false depends largely upon its internal structure, i.e. upon the particular way in which its elements are combined. Frege's analysis of the structure of thoughts is intimately connected to his ontological hierarchy: each thought is decomposable into at least two parts: one which is unsaturated, and one which saturates the unsaturated part. This latter part may itself be either saturated (i.e., an object) or unsaturated (i.e., a function, concept, or relation of appropriate level), depending upon the nature of the argument places of the unsaturated component.[11] The unsaturated component of a thought must be a function or a concept. The bulk of Frege's ontological work is spent clarifying the ways in which various entities can or cannot combine to form the saturated wholes he calls thoughts. As we will see below, it is Frege's attempt to develop mathematics on the basis of this ontological work which leads him into trouble.[12]

## 1.3 Truth and truth values

As noted, for Frege thoughts are those things which can be true or false. But how does he understand the notions of truth and falsity? Like many of his modern descendants, he treats truth and falsity as *values* taken by certain types of functions (e.g., concepts) when entities of appropriate type are input as arguments, and these values are themselves objects. There are, as one might expect, two such values, namely the True and the False. The concept '*x* is a bottle,' for example, maps all and only those objects which are bottles onto the True, while all objects which

---

[10] "[L]ogic is the science of the most general laws of truth" [Frege 1897, 139; in Beaney 1997, 228].

[11] Though only objects are *themselves* saturated, functions can complete or saturate other functions if they are of the appropriate level. The result of such saturation is, again, an object—often a thought. Thus, both objects and functions can saturate or complete other functions.

[12] Perhaps the most direct attack on the coherence of Frege's basic distinction between complete and incomplete entities can be found in [Resnik 1965]. There Resnik's aim is to illustrate that it is Frege's metaphysical assumptions which lead to the difficulties of his logicist program. In many ways, the present chapter aims to arrive at the same claim by different means. Unlike Resnik's paper, however, my aim is to extend these metaphysical troubles to include post-Fregean work in the philosophy of mathematics (including Resnik's own account (s) of structures in [Resnik 1981] and [Resnik 1999]).

are not bottles are mapped onto the False.[13] Though we are not primarily interested in Frege's linguistic views here, it is interesting to note that he takes the role of declarative sentences to be logically identical to (what we would normally call) proper names, and the objects which they name or refer to are the True (if the sentence is true) and the False (if the sentence is false).[14]

Fixing precisely which objects the True and the False are is something of a thorny issue for Frege and his commentators, but he does seem willing to treat the True and the False as courses-of-values of functions. In §10 of *Grundgesetze*, for example, he presents an elaborate argument supporting the claim that "it is always possible to stipulate that an arbitrary course-of-values is to be the True and another the False".[15]

Aside from these arbitrary stipulations (which are themselves quite problematic),[16] Frege says little about the properties of either truth value, or even his reasons for construing truth and falsity in terms of objects (i.e., courses-of-values). This is not entirely surprising, as he notes in several places that he takes the concept of truth itself to be something indefinable. He writes, for instance:

> [...] it would be futile to employ a definition in order to make it clearer what is to be understood by 'true'. If, for example, we wished to say 'an idea is true if it agrees with reality' nothing would have been achieved, since in order to apply this definition we should have to decide whether some idea or other did agree with reality. Thus we should have to presuppose the very thing that is being defined. The same would hold of any definition of the form '*A* is true if and only if it has such-and-such properties or stands in such-and-such relation to such-and-such a thing'. In each case in hand it would always come back to the question whether it is true that *A* has such-and-such properties, or stands in such-and-such a relation to such-and-such a thing. Truth is obviously something so primitive and simple that it is not possible to reduce it to anything simpler.[17]

Embedded in this paragraph are the beginnings of two distinct arguments in favour of the claim that truth is indefinable. These are developed further in one of Frege's last works, a 1918 essay entitled "The Thought". The first argument is rather limited in scope, addressing as it does only a very narrow form of the correspondence theory of truth. Both Wolfgang Künne and Michael Dummett have suggested for different reasons that the apparent narrowness of this first argument

---

[13] For Frege, entities which are not objects (i.e., functions) take neither the True nor the False as a value when input as argument into a (first-level) concept like '*x* is a bottle'; instead, they fail to produce a logical whole at all, resulting in nonsense. See [Diamond 1991, chapter 2] and [Goldfarb 1997] for competing accounts of exactly what is meant by nonsense here.

[14] As can be imagined, the view of sentences as names of truth values is a fairly controversial one which has given rise to considerable debate. [Heck 1997] gives a fairly reasonable explanation of this peculiar view, suggesting that this move expedites Frege's treatment of relations in terms of double courses-of-values.

[15] [Frege 1893, §10].

[16] Below, in our discussion of the Julius Caesar objection, we will treat in more detail problems pertaining to the association of a function with a unique course-of-values, as well as problems directly related to Frege's arguments in *Grundgesetze* §10.

[17] [Frege 1897, 128-129].

should not be taken to undermine Frege's broader claims about the indefinability of truth.[18] Let's take a look at this first argument and then see how it connects up with his larger claims about the indefinability of truth.

This first argument is pitted against what Künne calls the "traditional object-based correspondence theory" of truth.[19] On such a view, a statement is true if it corresponds to the object which it is about. Thus if I claim that my hat is yellow, the claim is true if it (somehow or other) corresponds with the object in question (viz. my hat). But what is meant here by correspondence, and how can such a correspondence give us truth? Frege writes

> Now a correspondence is a relation. But this goes against the use of the word 'true', which is not a relative term and contains no indication of anything else to which something is to correspond. If I do not know that a picture is meant to represent Cologne Cathedral then I do not know what to compare the picture with in order to decide on its truth. A correspondence, moreover, can only be perfect if the corresponding things coincide and so just are not different things. […] It would only be possible to compare an idea with a thing if the thing were an idea too. And then, if the first did correspond perfectly with the second, they would coincide. But this is not at all what people intend when they define truth as the correspondence of an idea with something real. For in this case it is essential precisely that the reality shall be distinct from the idea. But then there can be no complete correspondence, no complete truth.[20]

Here Frege suggests that the very idea of a correspondence contains a kind of imperfection which is foreign to the nature of truth. One of Frege's guiding assumptions throughout his work is that truth is not a matter of degrees, nor is it relative to a given context, situation, or theory. Truth is an unchanging feature of thoughts. Thus, if we suggest that the only means of getting at perfect truth is via a correspondence between our claims and reality, then we are doomed to failure from the very start. For a picture is, by its very nature, something other than what it pictures. But this argument is only really effective against a fairly naïve form of the correspondence theory. Frege uses this naïve view as the starting point for a broader criticism of correspondence as a means of defining truth.

Immediately following the above-quoted passage he writes:

> But could we not maintain that there is truth when there is correspondence in a certain respect? But which respect? For in that case what ought we to do so as to decide whether something is true? We should have to inquire whether it is *true* that an idea and a reality, say, correspond in the specified respect. And then we should be confronted by a question

---

[18] Cf. [Dummett 1973, 443] and [Künne 2008, §2]. Dummett holds that, while Frege's arguments do not convince us of his ultimate claim, they do provide us with sharp new limitations on any attempted definition of truth. Künne suggests that Frege's arguments about truth are subtler than many commentators have noted, a fact owing, perhaps to their enthymematic character.

[19] [Künne 2008, 12].

[20] [Frege 1918, 59-60; Beaney 1997, 326-327].

> of the same kind, and the game could begin again. So the attempted explanation of truth as correspondence breaks down.[21]

This is the crux of his first argument against the definability of truth in terms of correspondence. The claim here is that any attempt at defining truth in terms of correspondence leaves us no better off than when we started. For in order to apply any definition of this form (say, '*x* is true iff *x* corresponds in respect *y* with reality') in the first place, we would have to know whether it was true that our idea corresponded with reality in the relevant respect. But to establish that our idea corresponds to reality in the relevant respect, we can supply another claim or give another idea (the idea of my idea corresponding to reality in the relevant respect), and so on down the line. The problem here is not that we encounter an infinite regress, but that our attempts to apply the definition never actually succeed; we are always left with the same demand to apply the definition again. We find ourselves unable to get anywhere with the definition.

But this narrower argument only attacks certain features of correspondence theories of truth, which are far from the only game in town. Following this focussed attack, Frege generalizes the argument to include any attempt at defining truth as a property or relation of any kind. He writes:

> And any other attempt to define truth also breaks down. For in a definition certain characteristics[22] would have to be specified. And in the application to any particular case the question would always arise whether it were *true* that the characteristics were present. So we should be going round in a circle. So it seems likely that the content of the word 'true' is *sui generis* and indefinable.[23]

In this more general argument, nothing hinges upon the particularities of correspondence relations. Instead, the general nature of definitions is brought into play. A definition involves isolating a particular notion by providing characteristics which only that notion possesses. On Frege's view, when providing such characteristics for 'true', we would have to ask again whether it was true that these general characteristics were present in the particular case at hand. But then we would have to apply our definition again, resulting in the same question once more. And round and round we go. As Künne points out, this isn't exactly a vicious infinite regress, but

> a kind of vicious circle objection. This does not mean that [Frege] condemns all definitions of 'true' as circular definitions A definition is circular if the concept expressed by the definiendum is […] expressed by a proper part of the alleged definiens. […] The circle Frege is complaining about is not of this kind. Rather, it is a circle one gets into as soon as one tries to *apply* the alleged definiens of 'true'.[24]

The success of this argument, then, seems to hinge at least in some part on the capacities of those who would attempt to apply it. More specifically, as Künne notes, it hinges on the general

[21] [Frege 1918, 60; Beaney 1997, 327].

[22] The German term here is '*Merkmale*' which can also be translated 'marks'. Below we will have occasion to discuss Frege's technical understanding of these characteristics or marks in more detail. [Künne 2008, 13] gives good reasons for thinking that Frege is *not* using '*Merkmale*' here in his technical sense.

[23] [Frege 1918, 60; Beaney 1997, 327].

[24] [Künne 2008, 13].

principle "One cannot decide whether things are thus and so without deciding whether it is *true* that things are thus and so", which, he claims, "Frege seems to regard as self-evident".[25] Ultimately, Frege does not claim that these arguments are apodeictically certain or even terribly convincing, he only holds that given these considerations "it seems likely" that we cannot define truth at all.

Despite his acceptance of the impossibility of a proper definition of the concept, he goes on to suggest that we can "bring out the peculiarity of our predicate [i.e., 'is true'] by comparing it with others" and that "What, in the first place, distinguishes it from all other predicates is that predicating it is always included in predicating anything whatever".[26] It would seem, then, that Frege's many remarks about truth and the predicate 'is true' are to be taken as pre-systematic elucidations of the concept. None of these elucidations are particularly helpful in deciding upon the relationship between Frege's primitive concept of truth and the two objects which serve as truth values. Presumably, the extension of the concept of truth is identical with the object Frege calls the True, but this assumption is never made explicit and there remains confusion as to the relations between the True, and the predicate '$x$ is true'.[27]

## 1.4 Functions

In addition to the saturated entities Frege calls objects, he also discusses unsaturated entities. These he calls functions.[28] Frege's view of functions is closely related to the informal mathematical notion of a numerical function. A function in this sense is a rule for, or way of, determining one quantity (the value of the function) by means of another quantity (the argument, or input).

Within most areas of mathematics functions are only defined for a particular domain, e.g., on the natural, real, or complex numbers. Each entity within the domain of a function determines a specific value for that function. The range of the function is composed of the collection of entities which act as its values. The range and domain may be disjoint, co-extensive, or partially overlap with each other. The successor function on the natural numbers, for instance, takes natural numbers as inputs and delivers natural numbers as outputs. The square root function defined on the natural numbers as arguments, however, takes natural numbers as inputs but includes real numbers amongst its outputs. Though, historically, many mathematical domains have been defined by the types of objects of which they are composed (e.g., Euclidean space as the collection of Euclidean points, the natural number domain as the collection of all natural

---

[25] [Künne 2008, 14].

[26] [Frege 1897, 129]. [Künne 2008, §3] discusses the arguments in favour of a version of this claim in more detail.

[27] For more detailed examination of Frege's views on the indefinability of truth and their development, see the remainder of [Künne 2008] as well as [Sluga 2002]; for an account of the relationship between Frege's anti-empiricist epistemology and his conception of truth, see [Ricketts 1996] and [Levine 1996].

[28] It should be noted here that the ontological or metaphysical characterization of functions which I outline in this section was not always present in Frege's work. As [Heck and May 2013] convincingly shows, in the *Begriffsschrift* of 1879, Frege's distinction between functions and objects was (at least tacitly) a *linguistic* one. Later, in the *Grundgesetze* of 1893, Frege is very explicit about the ontological character of this distinction. My remarks here concern only this later period in Frege's thinking, which he took to be an improvement upon his earlier position (see, again, [Heck and May 2013] for detailed argumentation to this effect).

numbers), some mathematical domains have been partially characterized in terms of closure under certain types of functions. For a domain to remain closed under a given function the range must be a subset of the domain. Thus, for instance, an important part of the definition of a group is that a group is closed under the binary group operation. Similarly, the natural numbers are closed under addition, but if we include the operation of subtraction, we must extend the domain to include all the integers to maintain closure.

Historically, many new mathematical domains—perhaps most notably within number theory—have been investigated in an effort to understand what type of domain is required to maintain closure under a particular operation or type of function. This is one way in which the concept of a function has been extended. Beginning with the birth of analysis during the time of Leibniz and Newton, the concept of a function has been almost continuously expanded, such that it has come to include functions defined for wider and wider number domains, as well as functions which employ a much wider array of allowable operations. Frege saw his work as contributing to both these ways of extending the notion of a function. First, he extended functions to include a much greater variety of possible inputs. Instead of simply the natural, real, or complex numbers, Frege viewed functions as accepting any objects whatsoever as arguments and values.[29] In effect, Frege demanded that functions be defined exclusively within the single domain of his ontological hierarchy. Second, he extended the allowable operations which were taken to define mathematical functions. Just as the Irish mathematician William Hamilton's development of quaternions required him to countenance non-commutative forms of multiplication, Frege, similarly, sought to expand the domain of allowable mathematical functions by including, e.g., the symbol for equality in the construction of function-expressions.[30]

Alongside his two-fold extension of the mathematical notion of a function, Frege also attempted, in both his formal and informal works, to deepen our understanding of the nature of functions and their relations to objects. On the linguistic level, there is the obvious difference between symbols which refer to functions and those which refer to objects that function-symbols tend to indicate, in one way or another, the need for completion by object-symbols, often of a particular type. Frege believed that this orthographic feature of mathematical discourse represented a deeper, ontological difference between functions and objects. One key aspect of this difference is that, for him, functions possess argument places. Frege writes that he "is concerned to show that the argument does not belong with a function, but goes together with the function to make up a complete whole; for a function by itself must be incomplete, in need of supplementation, or unsaturated. And in this respect functions differ fundamentally from numbers".[31] Though Frege is specifically concerned here with differentiating functions from numbers, his remarks are guided by the more general view that functions differ fundamentally from *all* objects. The argument places of functions dictate which form or forms of supplementation a particular

[29] After Frege introduces concepts as functions whose values are always truth-values (at [Frege 1891, 139]), he writes that "we must go further and admit objects without restriction as values of functions" [Frege 1891, 140].

[30] Though Frege's remarks on the nature of functions always presume an extremely liberal understanding of the allowable types of functions, in practice he rarely refers to the stranger types of function which were gaining in mathematical popularity during his career. [Burgess 1997] presents an interesting examination of Frege's actual use and treatment of specifically mathematical functions throughout his work against the backdrop of broader developments in the concept of function of the time.

[31] [Frege 1891, 133].

function will accept. There are different levels and degrees of function determined by the nature and number of argument places. Each of these levels is entirely distinct from all others, such that a function of one level is (according to Frege) as ontologically separate from every other type of function as it is from objects.

For Frege, the most basic functions are first-level functions of one argument place. Such functions possess a single argument place which can be supplemented or saturated only by an object. The result of a first-level function with an object saturating its argument place is another object. Thus, for instance, we might have the square function $x^2$ (where '*x*' indicates an argument place). This function takes objects as arguments and delivers objects in turn: if we input the number 2 as argument, the value returned by the function is the number 4, another object. There are, similarly, functions of the first level with multiple argument places, each of which can take any object as an argument.[32] Second-level functions take first-level functions as arguments, and can possess one or more argument places. Third-level functions take second-level functions as arguments, and so on.

Perhaps one of Frege's most lasting contributions to modern thought was his understanding of concepts as particular types of functions, namely those of one argument place whose values are always truth-values. Similarly, he viewed relations of all types as functions of two or more arguments whose values are always truth-values. Thus, for instance, the concept of a natural number is a function which takes objects as arguments and gives truth values as outputs. All objects which are indeed natural numbers give the True as value when input into the function, while any objects which are not natural numbers give the False as value. Concepts, like other functions, can be of different levels. Thus, the concept '*x* is a horse' is of the first-level because it takes objects as arguments, while the concept '*F* is a concept easily attained' is a second-level concept insofar as it takes first-level concepts as arguments (we will discuss this in more detail below in the section on the concept *horse* problem).

In order to completely specify what type of function we are dealing with, Frege believes that we must carefully indicate the nature and number of its argument places. We may, for instance, have two second-level functions which, despite this similarity, are nevertheless fundamentally different insofar as they possess a different number of argument-places. For Frege, a second-level function of two argument places is, despite certain similarities, as ontologically distinct from a second-level function of three argument places as it is from an object.[33] Because this hierarchy of functions is (at least in principle) infinitely extendible, and each level of function is ontologically separate from every other, Frege's ontology ends up being quite rich despite the relative simplicity of the general notions involved. We have both an infinite number of levels of function and, at each level, an infinite number of different types of functions determined by the number of argument places they possess. To further add to this ontological panoply, there are also various functions of *mixed* level which possess two or more argument places of different levels. Thus, for instance, we might have a function which possesses one argument place which

[32] E.g., $x + y$, $x = x$, $((x + 3) - x)/3$, etc. where '*x*' and '*y*' represent argument places for objects.

[33] I have found Warren Goldfarb's description (at [Goldfarb 1997, 58] ) of these distinctions in terms of different types of Velcro particularly useful.

takes objects as arguments and another argument place which takes first-level functions as arguments. This does not mean, however, that there are single argument places which accept entities of different level, only that a function may possess argument places of multiple different levels.[34]

As we've seen, Frege describes objects as saturated and self-standing or self-subsistent. We've also seen that for him a key feature of concepts (and functions more generally) is their unsaturatedness. In this they differ fundamentally from objects. But are concepts also self-subsistent for Frege? That is, do concepts possess independent existence in the same way that Fregean objects do?

Given the centrality of functions for Frege's ontology, the somewhat surprising answer is that they do not. In 1882, in a letter to the Swiss philosopher Anton Marty (1847-1914) outlining his logical views, Frege writes

> A concept is unsaturated in that it requires something to fall under it; hence it cannot exist on its own. […] Now I do not believe that concept formation can precede judgement, because this would presuppose the independent existence of concepts, but I think of a concept as having arisen by decomposition from a judgeable content.[35]

And later, in one of his 1903 articles on the foundations of geometry, he states the view again:

> It is clear that we cannot put a concept forward as self-subsisting like an object. Rather a concept can only occur in a combination. It can be said that the concept can be distinguished in the combination, but cannot be separated out of the combination.[36]

This is in some ways a puzzling view. For Frege certainly does not want to deny that concepts and functions have a kind of ontological status. They are not in any sense reducible to objects, though it does seem, as Thomas Ricketts convincingly argues, that objects are, in a certain sense, logically prior to concepts.[37] But, as the metaphor of unsaturatedness is meant to elucidate, it is of the essence of concepts that they are isolable only in the context of the analysis of the logical wholes Frege calls 'thoughts'. To treat them as self-standing entities, as already-completed wholes, would be to misconstrue them as yet another type of object. And, as we will see in greater detail below, one of Frege's three guiding principles in his logical and mathematical work was "never to lose sight of the distinction between concept and object".[38]

1.5 Relations between functions and objects

---

[34] E.g., the functions $\int_0^x f(t)dt$ and $\left[\frac{d}{dx}f(x)\right]_{x=y}$, both of which take objects and functions as arguments.

[35] [Frege to Marty, August 29, 1882; in Beaney 1997, 81].

[36] [Frege 1903, 372, note 5].

[37] Cf. [Ricketts 2010, 163-170]. Ricketts argues that this sort of priority is a self-conscious (if complex) aspect of Frege's view; [Bergmann 1958] suggests that, self-conscious or not, it leads Frege away from the sort of Platonic realism about functions with which he is often associated and towards a 'hidden nominalism' about functions.

[38] [Frege 1884b, xxii].

In addition to outlining the properties of functions and objects, Frege also discusses various relationships between these ontological categories. One of the most important logical relationships between objects and concepts, as noted in Frege's letter to Marty, is that of an object 'falling under' a concept. Frege says that an object falls under a concept just in case it produces the value the True when input into the concept as argument. As a human being, I fall under a number of concepts. For instance, I fall under the concept "*x* is a corporeal being". Frege's understanding of properties is directly connected to this relationship between concepts and the entities which fall under them. The properties (*Eigenschaften*) of an entity are, for Frege, just the concepts under which that entity falls.[39]

In his attempt to understand the nature of number, Frege found it useful to note a distinction between the properties of an entity and the *marks* or *characteristics* (*Merkmale*) of the concept under which that object falls. Let's take a look at what he says, and follow it up with a few examples. In §§52-53 of the *Foundations of Arithmetic*, Frege writes:

> […] our ordinary language does assign number not to concepts but to objects: we speak of "the number of the bales" just as we do of "the weight of the bales". Thus on the face of it we are talking about objects, whereas really we are intending to assert something of a concept. This usage is confusing. The construction in "four thoroughbred horses" fosters the illusion that "four" modifies the concept "thoroughbred horse" in just the same way as "thoroughbred" modifies the concept "horse." Whereas in fact only "thoroughbred" is a characteristic [*Merkmal*] used in this way; the word "four" is used to assert something of a concept. […] By properties which are asserted of a concept I naturally do not mean the characteristics [*Merkmale*] which make up the concept. These latter are properties of the things which fall under the concept, not of the concept. Thus, "rectangular" is not a property of the concept "rectangular triangle"; but the proposition that there exists no rectangular equilateral rectilinear triangle does state a property of the concept "rectangular equilateral rectilinear triangle"; it assigns to it the number nought.[40]

We can see here that the distinction between a mark and a concept is quite a useful one for Frege's dual project of disambiguating natural language and garnering a proper understanding of the nature of number. Natural languages like German and English tend to blur the distinction, particularly when it comes to number, and Frege believes that this can (and does) lead us to incorrect accounts of the actual nature of number. By mistaking the properties of a concept for the properties of the object or objects which fall under that concept, we are led to a misunderstanding about the role concepts play in numerical applications like counting.

But the distinction is not only useful in the specific context of Frege's logicist program in the *Grundlagen*. Eight years later, in his attempt to overcome Benno Kerry's criticisms of his distinction between concepts and objects, he returns to the vocabulary of marks and properties. In 'On Concept and Object' he writes:

[39] [Frege 1892a, 201; in Beaney 1997, 189].
[40] [Frege 1884b, §§52-53].

> The words ['*Eigenschaft*' and '*Merkmal*'] serve to signify relations, in sentences like 'Φ is a property of Γ' and 'Φ is a mark of Ω'. In my way of speaking, a thing can be at once a property and a mark, but not of the same thing. I call the concepts under which an object falls its properties; thus
>
> 'to be Φ is a property of Γ'
>
> is just another way of saying
>
> 'Γ falls under the concept of a Φ'.
>
> If the object Γ has the properties Φ, X and Ψ, I may combine them into Ω; so that it is the same thing if I say that Γ has the property Ω, or, that Γ has the properties Φ, X and Ψ. I then call Φ, X and Ψ marks of the concept Ω, and, at the same time, properties of Γ. It is clear that the relations of Φ to Γ and to Ω are quite different, and that consequently different terms are required. Γ falls under the concept Φ; but Ω, which is itself a concept, cannot fall under the first-level concept Φ; only to a second-level concept could it stand in a similar relation. Ω is, on the other hand, subordinate to Φ.[41]

Here we see the distinction laid out more explicitly. Frege employs the term 'mark' when encountering what we might call a conjunctive or a complex concept. A complex concept, in this sense, is one composed of a number of component concepts. Thus we might think of the complex concept '*x* is a large orange cat' as a conjunction of three component concepts, '*x* is large,' '*x* is orange,' and '*x* is a cat'. Such complex concepts needn't be given in such an obviously conjunctive form. We might, for instance, have previously defined the phrase '*x* is a Garfield' to mean '*x* is a large orange cat,' in which case we can still think of '*x* is large,' '*x* is orange,' and '*x* is a cat' as component concepts of the complex concept '*x* is a Garfield'.

Let's take another example and work through the relationship between the properties of an entity and the marks of a concept under which that entity falls.

Blue guitars—objects in Frege's dispensation—fall under two distinct component concepts: '*x* is blue' and '*x* is a guitar'. For Frege, this is another (and perhaps a more accurate) way of saying that these objects possess the properties of being blue and of being a guitar.

Just as guitars have properties, so too do the concepts under which they fall. Importantly, and perhaps obviously, the properties of concepts are usually quite different from the properties of the objects which fall under them. For instance, it is *not* a property of the concept '*x* is a blue guitar' that it is blue. One property of our complex concept would be that it is not an empty one; there have in fact been objects in the world which have been both blue and a guitar.

Frege uses the term 'mark' or 'characteristic' to refer to a quite different relationship between concepts and the objects which fall under them. For Frege, some of the properties of the objects which fall under a complex concept are to be understood as the *marks* of that concept. Take a

[41] [Frege 1892a, 201-202; in Beaney 1997, 189-190].

particular blue guitar as an example. Such an object falls under a plethora of concepts, not all of which will count as marks of the complex concept (‘*x* is a blue guitar’) we are interested in here. For instance, it might fall under the concept ‘*x* is destructible’. Now, it is in fact the case that *all* the objects which fall under our complex concept ‘*x* is a blue guitar’ also happen to be destructible. But for properties to count as marks of a given concept, they have to be component concepts of that concept, and ‘*x* is destructible’ is not one of the concepts conjoined to form our complex concept. The notion of a mark has much more to do with the parts or components of the concept than it has to do with the objects which fall under the concept.

Let’s see how this plays out.

We have an object (say, Rollie Fingers), a complex concept (‘*x* is a moustachioed baseball player), and two component concepts (‘*x* is a baseball player’ and ‘*x* is moustachioed’).

Our object, Rollie Fingers, has the following three properties:

1. He is a Cy Young award winner.
2. He is a baseball player.
3. He is moustachioed.

None of these are properties of the concept ‘*x* is a moustachioed baseball player’, which is to say that the concept ‘*x* is a moustachioed baseball player’ does not fall under (indeed, cannot fall under) any of the concepts underlying these three properties. The latter two of these properties are, however, *marks* of the complex concept ‘*x* is a moustachioed baseball player’. They are marks of this complex concept because they are the components or parts out of which it was constructed. The first property and many other properties of Rollie Fingers are not components of the complex concept and are not, therefore, marks of that concept.

All of this might seem obvious enough (even pedantic), but when it comes to complex concepts of higher levels, or highly abstract ones like those we find in mathematics, it becomes more difficult to distinguish between properties and marks without such a clearly defined vocabulary. The distinction between a property and a mark is not exactly an ontological one, as a single concept can serve as both a mark and a property,[42] though it cannot be a mark and a property of the same entity. It is nevertheless a useful terminological distinction.

Another relation that emerges in Frege’s discussion of properties and marks is that of subordination (*Unterordnung*). For Frege, a concept *F* is subordinate to a concept *G*, when the extension of *F* is contained within the extension of *G*. Note that this relation need not be asymmetrical: co-extensive concepts can be said to be mutually subordinate. A complex concept is subordinate to its component concepts. Thus, for instance, the concept ‘*x* is a positive even whole number greater than 7’ is subordinate to the component concepts ‘*x* is a positive number’, ‘*x* is a whole number’ ‘*x* is an even number’ and ‘*x* is greater than 7’. The relation of subordination, however, needn’t be restricted to complex concepts (i.e., concepts which are given as conjunctions of simpler concepts).

[42] Cf. [Frege 1892a, 201; in Beaney 1997, 189].

In contrast to the relations between concepts and marks, the relation of subordination is concerned more directly with the entities which fall under the concept than with the components of the concept itself. Let's take two first-level concepts as an example, say '*x* is a banjo' and '*x* is a physical object'. The former is subordinate to the latter not because being a physical object is part of a conjunction of concepts which make up the concept '*x* is a banjo' but because of certain facts about the objects which fall under these two concepts. All the objects which fall under the concept '*x* is a banjo' also fall under the concept '*x* is a physical object', but the converse is not true. It is this relation between the objects which fall under the concepts which determines subordination.

Frege stresses at a number of points that this relationship of subordination which holds only between concepts of the same level is not to be confused with similar relations between entities of different levels.[43] We've already seen the relation of falling under which holds between objects and first-level concepts. But Frege also distinguishes a similar relationship between first- and second-level concepts which he calls 'falling within'. The distinctions between 'subordination,' 'falling under,' and 'falling within' are of particular importance for understanding his remarks regarding Hilbert's work on Euclidean geometry. He writes that

> Second-level concepts, which [first-level] concepts fall under, are essentially different from first-level concepts, which objects fall under. The relation of an object to a first-level concept that it falls under is different from the (admittedly similar) relation of a first-level to a second-level concept. To do justice at once to the distinction and to the similarity, we might perhaps say: an object falls *under* a first level concept; a concept falls *within* a second-level concept. The distinction of concept and object thus still holds, with all its sharpness.[44]

So, for instance, the object Jumbo falls under the first-level concept '*x* is an elephant'. The first-level concept '*x* is an elephant' falls within the second-level concept '*F* is not an empty concept'. And the first-level concept '*x* is an elephant' is subordinate to the first-level concept '*x* is a mammal'. In case of Hilbert's geometric researches, Frege would say that Euclidean lines are objects which fall under the first-level concept '*x* is a Euclidean line', while Hilbert's supposedly uninterpreted notion of a line is better understood as a second-level concept within which fall various first-level concepts, including the concept of a Euclidean line. This is different from the relation of subordination which holds between, e.g., the first-level concepts '*x* is a Euclidean line' and '*x* is a Euclidean parabola'. Thus, from Frege's point of view, much of the confusion in Hilbert's work stems from the fact that Hilbert confuses the properties of Euclidean points with the marks of his second-level concept of point. Though Frege does not often remark upon concepts of levels higher than the second, it is presumably the case that the difference between falling under and falling within can be analogously extended to each ontological level. Thus the relationship between second- and third-level concepts (falling into?) will be similar, though ontologically distinct, from that which holds between second- and first-level concepts. The

[43] [Frege 1882; in Beaney 1997, 81].
[44] E.g., at [Frege 1884b, §53], [Frege 1892a, 201; in Beaney 1997, 189].

difficulty of making such distinctions in ordinary language can of course be surmounted through various formal devices (e.g., subscripts).

## 1.6 Courses-of-values and extensions

Frege believed that each function was associated with a (unique) object. These objects are called *courses-of-values* (*Wertverläufe*) or, if the function happens to be a concept, *extensions*. Let's look at the special case of extensions first. Though Frege assumes that it is known what the extension of a concept is,[45] there are, I think, at least two ways of understanding this notion. The first is to take as the extension of a concept simply the collection of all entities which produce the value 'the True' when input into the concept as arguments, i.e., the collection of all those objects which fall under the concept at issue. The second would be to take the extension as a set of ordered pairs, where the first term of each pair is any object whatever, and the second either the True or the False, depending on whether the object falls under the given concept. For the concept '$x$ is a natural number', accordingly, the extension would thus be either:

{0, 1, 2, 3, …}

or (using the usual notation for ordered pairs):

{...<Scipio Africanus, the False>, <2, the True>, <3, the True>, ...}

The first way of understanding extensions of concepts is a fairly common one. But it will not do as an account of the nature of courses-of-values in general. This first option cannot help us explain courses-of-values for functions which are not concepts, as in such functions there is no pairing between the arguments of a function and the True and thus no obvious way to generate a list of entities. In the general case, if we are to attempt to pick out a unique course-of-values, we need to know which arguments produce which values, a task which can be accomplished if we model courses-of-values on the second approach using ordered.

Since Frege claimed that no entity can stand in the relation of identity to any entity of a different ontological level, he did not identify functions with their courses of values (as is often done, for example, in most versions of contemporary set theory). Given the loose characterization of functions as rules for determining quantities by means of other quantities, this refusal seems plausible; for a rule is not identical to the pairings it determines. For Frege, too, this makes sense, as his ontological hierarchy is built largely upon the fundamental distinction between functions and objects. This understanding of the notion of function has the consequence that distinct functions might have the same course of values,[46] or distinct concepts the same extension. Thus, for instance, the square function will possess the same course of values as the function $x^2 + 0$. Similarly, when programming a computer we might have three distinct programs which achieve the same goal (i.e., deliver the same output given the same inputs) though at wildly different computational speeds and levels of complexity. There are also well-known examples of co-extensive concepts which seem 'intensionally' quite different (a standard example is the pair of

[45] [Frege 1884b, §69, note 1].
[46] Frege admits as much in the specific case of concepts at [Frege 1884b, §68 note 1].

concepts '*x* is a creature with kidneys' and '*x* is creature with a heart'). Below we will discuss in greater detail the way in which Frege believes the course of values of a given function (or the extension of a particular concept) is determined—we will also see that this seemingly simple concern posed a major problem for his entire philosophical program.

Now that we have a clearer understanding of Frege's ontological hierarchy, let's begin our examination of problems pertaining to that hierarchy.

## §2. The concept *horse* problem

The first problem with Frege's philosophy we will examine is the so called concept *horse* problem. This problem relates directly to Frege's basic distinction between objects and functions. As we've seen above, Frege took it to be the case that the world was exhaustively divided into objects and functions,[47] and that no object could be identical to a function. The concept *horse* problem, first raised by Benno Kerry (1858-1889), seems to attack the possibility of maintaining this absolute distinction. In a series of eight articles published between 1885 and 1891, Kerry examined the role of intuition within reasoning. At several points in these articles, Kerry mentions Frege's works, sometimes favourably, sometimes critically. Kerry agreed with Frege, for instance, that the relationship between concepts and the objects which fall under them is likely irreducible, but argued that "this [is] in no way bound up with the view that the properties of being a concept and of being an object are mutually exclusive".[48] Further, he argued that, on Frege's own account, we can construct meaningful sentences expressing thoughts which sensibly treat concepts as objects, and, thus, that the distinction between the two types of entity is not absolute. His arguments against Frege's claim that the world is exhaustively divided into these two disjoint categories center upon the following putative counterexample:

> The concept *horse* is a concept easily attained.

This relatively simple example seemed, to Kerry, to illustrate a serious problem with Frege's distinction. For, by Frege's own lights, the use of the definite article generally indicates that we are speaking of an object.[49] Thus, if the phrase "the concept *horse*" refers to anything, it would seem to refer to an object and not a concept. Hence, Kerry's sentence seems to express a thought in which an entity that is obviously a concept (namely, the concept *horse*) is treated as an object. So, following Kerry's reasoning, Frege must either drop his insistence that concepts and objects are mutually exclusive, or admit that his system contains a contradictory element.

[47] "[A]n object is anything that is not a function" [Frege 1891; in Beaney 1997, 140].

[48] [Kerry, quoted in Frege 1892a, 193; in Beaney 1997, 182].

[49] Frege notes of Kerry's claim that "the concept 'horse'" refers to an object that it is quite correct, and that it "is in full accord with the criterion I gave—that the singular definite article always indicates an object, whereas the indefinite article accompanies a concept-word" [Frege 1892a, 195; in Beaney 1997, 184]. This criterion is also discussed in *Die Grundlagen der Arithmetik*, where Frege writes, for instance, that "A general concept word just designates a concept. Only with the definite article or a demonstrative pronoun does it function as a proper name of a thing, but it then ceases to function as a concept word" [Frege 1884b, §51; in Beaney 1997, 102].

A third rather easy way out of the problem for Frege would be to alter his views so that they allow for instances where the use of the definite article does not accurately reflect the logical structure of the thought expressed.[50] Frege might then treat Kerry's putative counter-example as an exception to a rule which is applicable only in *most* cases. Frege does admit as much, and provides a few such examples of his own. But he subsequently dismisses these rare exceptions as unimportant and irrelevant to Kerry's main point.[51] As we will see when we begin to analyze Frege's response to the problem, this refusal to alter his views of the logical function of the definite article is one of the most puzzling aspects of this episode.

Before we get to Frege's response, however, let's take a closer look at the exact sources of the problem in Frege's philosophy. The first and most fundamental is his refusal to allow any identification of concepts/functions with objects, and vice versa. This means that no function can be an object, and no object can be a function. And, since Frege's ontological distinctions are intended to cut deeper than mere psychological, linguistic, or temporal considerations, there is no sense to claims that the very same entity can be considered an object from one point of view and a function from another. Putative counterexamples to the contrary are, on Frege's view, simply the result of confusion or equivocation. This insistence that ontological distinctions are absolute is essential both to his own project of reducing arithmetic to logic and to his criticisms of Hilbert's views.

The second, related, source of the problem lies in the fact that Frege takes his ontological distinction to be exhaustive. Every entity is either an object or a function with a determinate number of argument places of determinate types. There is no third type of hybrid entity able to partially fulfill the roles of both saturated and unsaturated entities. Hence the option of placing troublesome entities somewhere in the middle of the function/object distinction is simply not available to Frege.

Third, the problem arises because of certain features in Frege's theory of language, notably his view that proper names of any kind (including those indicated via the use of the definite article) can only refer to objects if they refer at all.

<u>2.1 Mendelsohn's analysis of the problem</u>

We can understand, I think, Frege not wanting to alter his views regarding the fundamental distinction at the heart of his entire theory, so it makes sense that he does not simply invent an ad

---

[50] An example of such an instance might be the statement 'The concept *unicorn* is empty'. Here the surface grammar suggests that we are attributing a property (emptiness) to an object (the concept *unicorn*), but it seems relatively uncontroversial to symbolize the actual logical structure of this sentence, in modern notation, as follows: $\forall x{\sim}Ux$, i.e., as a straightforward claim that nothing falls under the concept unicorn.

[51] Frege admits that there are certain difficulties pertaining to the plural usage of the definite article, as well as rare cases of the singular. He writes: "In the singular, so far as I can see, the matter is doubtful only when a singular takes the place of a plural, as in the sentences 'The Turk besieged Vienna,' 'The horse is a four-legged animal'. These cases are so easily recognizable as special ones that the value of our rule is hardly impaired by their occurrence. It is clear that in the first sentence 'the Turk' is a proper name of a people. The second sentence is probably best regarded as expressing a universal judgement, say 'All horses are four-legged animals' or 'All properly constituted horses are four-legged animals'" [Frege 1892a, 196; in Beaney 1997, 184-185].

hoc third ontological category or relax the strictness of the distinction. But it does seem rather puzzling that he is so steadfast in his merely linguistic doctrine regarding the definite article. What reasons can he possibly have for sticking to his guns here? Frege himself is of very little help in answering this question. In a rather infamous passage from "On Concept and Object," for instance, he writes that

> By a kind of necessity of language, my expressions, taken literally, sometimes miss my thought; I mention an object, when what I intend is a concept. I fully realize that in such cases I was relying upon a reader who would be ready to meet me half-way—who does not begrudge a pinch of salt [52]

As Richard Mendelsohn (among others)[53] notes, this appeal to readers to simply accept the *prima facie* preposterous claim that the concept *horse* is not a concept is "much too dogmatic: a reasonable doubt about the theory remains as long as Frege is unable to dispel the puzzle."[54] One way of doing so would be to explain exactly what is meant by 'a kind of necessity of language' here.

He does briefly try to explain this necessity in a footnote, but the explanation itself is confusing and quite misleading. In the footnote, Frege writes:

> A similar thing happens when we say as regards the sentence 'this rose is red': the grammatical predicate 'is red' belongs to the subject 'this rose'. Here the words "The grammatical predicate 'is red'" are not a grammatical predicate but a subject. By the very act of explicitly calling it a predicate, we deprive it of this property.[55]

Now presumably what Frege is trying to show here is that we encounter exactly the same difficulty with phrases like 'The concept $\eta$' as we do with phrases like 'The predicate $x$'. As his wording indicates, he takes this difficulty to stem from the subject/predicate structure of (some) natural languages. But, as Richard Mendelsohn points out, Frege is just wrong here. And, as a result, the analogy that he wishes to construct as a means of explanation here is actually rather misleading and unhelpful. Mendelsohn writes that "contrary to what Frege says, we don't, simply by calling a given expression a grammatical predicate, deprive it of the property of being a grammatical predicate".[56]

To see why, consider the following two sentences from Mendelsohn:

(A) The grammatical predicate 'is red' is not a grammatical predicate.[57]
(B) 'The grammatical predicate "is red"' is not a grammatical predicate.[58]

---

[52] [Frege 1892a, 204; in Beaney 1997, 192].
[53] See, for instance, Crispin Wright's aptly titled chapter "Why Frege Does not Deserve His Grain of Salt" in [Hale and Wright 2001].
[54] [Mendelsohn 2005, 73].
[55] [Frege 1892a, 196 note H; in Beaney 1997, 185 note H].
[56] [Mendelsohn 2005, 73].
[57] [Mendelsohn 2005, 73]. Here and below I've replaced Mendelsohn's sequential numeration of these sentences.
[58] [Mendelsohn 2005, 74].

Mendelsohn rightly suggests that (A) is straightforwardly false: 'is red' is indeed a grammatical predicate, and (A) falsely states that it is not. There seems to be no problem at all in holding that 'is red' is indeed a grammatical predicate. Now, (B) is, according to Mendelsohn, equally straightforwardly true: the words "The grammatical predicate 'is red'" do not form a grammatical predicate at all. As Frege points out in his footnote, grammatically, these words form a subject not a predicate. And, again, we have no trouble at all seeing this. The problem in the footnote is that Frege seems to hold that (B) is helpfully and illuminatingly analogous to the troublesome claim that 'The concept *horse* is not a concept'. But, given the straightforward falsity of (A), this analogy gets us nowhere.

Mendelsohn writes:

> But, in the first place, it is not clear why we should regard (B) rather than (A) as the proper analogue to ['The concept *horse* is not a concept']. In the second place, there is no awkwardness about (A), and whatever awkwardness accrues to (B) is readily explained. In the third place, since (A) is false, we are able to speak about grammatical predicates, whereas on Frege's view, we do not seem to be able to say anything intelligible about concepts.[59]

So, Frege does not explain why the subject/predicate structure of natural languages *necessarily* leads us astray when speaking about either concepts or predicates. After this attempt at explanation, then, we are no closer to an understanding of the exact nature of the problem. Unfortunately, Frege himself does not give us much else to go on, and we have to do a bit of digging to see exactly which aspects of Frege's philosophy are at issue in the problem.

Mendelsohn, for his part, does not try to root the concept *horse* problem in the infelicities of language. He holds, rather, that "Frege's claim that the problems with ['the concept *horse* is not a concept'] are the result of a mere awkwardness of language just does not seem to be able to capture the depths of his difficulties."[60] So, where are we to look for the real root of the problem? Clearly the problem first arises when Frege commits himself to the truth of the thought expressed by the sentence:

(Horse) The concept *horse* is not a concept.

So, if we are to understand the source of the problem, we must first understand what sort of reasons Frege could possibly have had for holding that (Horse) is true. Frege himself does not give us any compelling argument for his support of this claim. However, Mendelsohn provides a reasonable reconstruction of an argument that one might make on the basis of Frege's statements in "On Concept and Object".

Let's take a look at Mendelsohn's reconstruction. He begins with the claim that (Horse) is a consequence of the principle:

---

[59] [Mendelsohn 2005, 74].
[60] [Mendelsohn 2005, 74].

(C) $\eta$ is a predicate if, and only if, $\eta$ stands for a concept.[61]

For Frege, this is a fundamental principle of his philosophy of language. A predicate is just an expression which stands for a concept, and something stands for a concept only if it is a predicate. This seems a fairly straightforward claim, and uncontroversial given Frege's other views. So where do we get into trouble and confusion?

Mendelsohn takes it to be the case that Frege fairly obviously holds that

(D) 'The concept horse' is not a predicate

If he did not hold this, then the problem of 'the concept *horse*' would not have arisen in the first place. So, by (C), this also commits him to the claim that

(E) 'The concept horse' does not stand for a concept

Since only predicates are able to stand for concepts, and 'the concept *horse*' is not a predicate, it cannot stand for a concept. Clear enough. Then we also have

(F) That which 'the concept *horse*' stands for is not a concept

But we also have the obvious identity:

(G) That which 'the concept *horse*' stands for = the concept *horse*

And then finally we get (Horse) by substituting 'the concept *horse*' for 'that which 'the concept *horse*' stands for" in (F). But, following this reasoning, we can easily show that there are in fact *no* predicates at all. Nothing here depends upon the peculiarity of 'the concept *horse*,' so we can simply reiterate the argument for any ostensible predicate we like, with the same result. Nothing, then, can be said to stand for a concept. This result is one with which Frege – who speaks of predicates and concepts quite happily throughout his work – ought to have been deeply unsatisfied.

So, what exactly is wrong with this picture? According to Mendelsohn, something is wrong with the very first step in the above argument, with (C).

First, Mendelsohn suggests that throughout "On Concept and Object" everything Frege says shows that he treats '$\eta$ is a concept' as the name of a first-level function. This means that Frege assumes that '$\eta$ is a concept' (or, after Kerry's example, '$\eta$ is a concept easily attained') refers to a function which accepts only objects as values. This assumption is evident when, for instance, he simultaneously suggests that the sentence 'The concept *horse* is a concept' is well-formed (i.e., has a truth value) and that 'the concept *horse*' is a proper name. If 'the concept *horse*' is a proper name, then it can only sensibly be combined with first-level function names (or perhaps

[61] [Mendelsohn 2005, 74].

the names of functions of mixed level) to form a proposition which expresses a complete thought bearing a truth value. Since '$\eta$ is a concept' is obviously not a name of a mixed level function, it must, therefore, be a first-level function name.

The result of all this is that the function denoted by '$\eta$ denotes a concept' can only accept objects as arguments and, therefore, must always be false, as no object is a concept. So, we can never truthfully say of anything that it denotes a concept.

Let's look again at (C). In order for the biconditional to be true, either both of the following conditions must be met or both must fail to be met:

1. $\eta$ must be a predicate
2. $\eta$ must stand for a concept

Frege would certainly like to hold that there are predicates, as he's spent quite a bit of time discussing them elsewhere in his work (hence his plea for a pinch of salt). But, by the argument we've just seen, the second condition can never be met. So, despite the fact that there are predicates (or so Frege's body of work seems to assume), (C) tells us that there can be no predicates because nothing can really stand for a concept. Since all of this is intolerable, we must throw away one or more of the assumptions. Given Mendelsohn's reconstruction, the most obvious solution is to reject the claim that '$\eta$ is a concept' is the name of a first-level function.

We have, then, to look at the reasons why Frege might have held this view in the first place. There is, again, a paucity of evidence from Frege himself. Mendelsohn notes that "The only reason he offers, as far as we can tell, is that '$\eta$ is a concept' can be sensefully completed by an *Eigenname* [a proper name]."[62] But, as Mendelsohn notes, this isn't terribly convincing. For, elsewhere, Frege is more than happy to construe sentences of very similar form as representing completely different logical structures.[63] More important than specific examples, however, is the fact that Frege's lifelong project of crafting a logically perspicuous language is predicated on the belief that natural language is rife with ambiguity, polysemy, and vagueness of all kinds. In many other places in Frege's work, we find him rejecting the typical analysis of the logical structure of a particular sentence in surprising ways.[64] So, as Mendelsohn writes,

> it is rather surprising that [in "On Concept and Object" Frege] should so staunchly maintain that the singular definite article invariably signals an *Eigenname*, for, as we see, he had adopted a more flexible attitude elsewhere, and moreover, his general rule of thumb was that the superficial grammar of a sentence was not always an accurate reflection of its logical structure. In order to establish that '$\eta$ is a concept' is a first-level *Funktionsname*,

[62] [Mendelsohn 2005, 78].

[63] [Mendelsohn 2005, 77-80] presents a clear comparison between Frege's account of the logical structure of (H) and his account (in the *Grundlagen* and elsewhere) of the logical structure of sentences like 'The number 2 exists'.

[64] See, for instance, [Frege 1884b, §91], or the preface to the *Begriffsschrift*, where Frege states the view about as clearly as possible: "If it is a task of philosophy to break the power of words over the human mind, by uncovering illusions that through the use of language often almost unavoidably arise concerning the relations of concepts, by freeing thought from the taint of ordinary linguistic means of expression, then my *Begriffsschrift*, further developed for these purposes, can become a useful tool for philosophers" [Frege 1879, v-vi; translated in Beaney 1997, 50-51].

> then, it is not enough merely to point to the fact that it can be sensefully completed by an *Eigenname* – especially when the interpretation creates such enormous logical difficulties. One must also show that the expression filling the blank space is operating as an *Eigenname* in that context.[65]

But, of course, Frege does not show this, and so there is no real reason to grant him his paradoxical grain of salt. So, taking into account Frege's wider philosophical views, his explanation of linguistic necessity as the source of the concept *horse* problem simply makes no sense. So now we have to ask: can we solve or avoid the problem in a way which is consonant with Frege's views, or is Kerry's objection disastrous for Frege's ontological distinction?

<u>2.2 Frege's proxy-object solution</u>

Frege himself offers an ontologically grounded account of the semantic situation presented by Kerry. This account involves the introduction of two new types of entity into Frege's ontology. He writes

> In logical discussions one quite often needs to say something about a concept, and to express this in the form usual for such predications – viz. to make what is said about the concept into the content of the grammatical predicate. Consequently, one would expect that the reference [*Bedeutung*] of the grammatical subject would be the concept; but the concept as such cannot play this part, in view of its predicative nature; it must first be converted into an object, or, more precisely, an object must go proxy for it. We designate this object by prefixing the words 'the concept', e.g.:
>
> > 'The concept *man* is not empty'.
>
> Here the first three words are to be regarded as a proper name, which can no more be used predicatively than 'Berlin' or 'Vesuvius'.[66]

The solution here involves describing the linguistic situation in ontological terms. If we take the surface grammar here as an indicator of the actual logical structure, then, Frege suggests, we are forced to treat 'the concept *man*' as a phrase which designates an object. But we desperately want this phrase to stand in some relation to a concept. Since Frege's linguistic views seemed to him to forbid from allowing the phrase of the form 'the concept *x*' to *refer* to a concept, he suggests a different relation entirely. The suggestion is that phrases of the form 'the concept x' refer to a new type of entity, not found elsewhere in Frege's works, a special sort of proxy object. To fully explicate the semantic situation here also requires the introduction of a new relation, the relation of 'representing' [*vertreten*], which differs from the relation of reference. Thus Frege has considerably expanded his ontology in order to explain the nature of this linguistic confusion. But this also commits him (or so the reader would hope) to a more explicit account of the nature of the entities (proxy objects, the relation of representing) which he has just

---

[65] [Mendelsohn 2005, 78].

[66] [Frege 1892a, 196-197; in Beaney 1997, 184-185]. Here Geach translates the German '*vertreten*' using the phrase 'go proxy for'. This term would more usually be translated 'represents'.

introduced. Given the difficulties which Frege has already encountered in relation to his ontological hierarchy, it seems that a solution which could avoid introducing new and poorly explained entities would be preferable.

2.3 A salt-free solution to the problem

How might we arrive at such a solution? Well, the first step will be to reject the surface grammar as a useful indicator of the logical structure of the thought expressed here. As Frege elsewhere has shown no hesitation to rehabilitate natural language, it seems entirely reasonable and consistent with his practice to suggest that Kerry is simply pointing out an ambiguity or vagueness in natural language which needs to be corrected. The question thus arises: how are we to construe a sentence like '$\eta$ is a concept'?

Ultimately, Frege's best response seems to be not to grant Kerry his claim in the first place. Without further information about what exactly is meant, sentences of the forms '$\eta$ is a concept' or '$\eta$ is a concept easily attained' are ambiguous. We need to know more specifically which level of concept is intended. For, as we've seen, Frege is keen to insist that concepts of different level and number of argument places are as ontologically distinct from one another as they are from objects. The English word 'concept' and the German word '*Begriff*' are both insufficiently precise to pin down exactly what is meant and, so, without further information or context, we have no way of deciding which level of concept is meant. This is a merely linguistic problem, akin in many ways to specifying the surname of each member of a group of men named 'Gregory' for clarity's sake. This sort of confusion might be obviated with subscripts or more explicit wording, e.g.,

1$^{st}$ level concept: '$\eta$ is a first-level concept' or '$\eta$ is a $\text{concept}_1$'
2$^{nd}$ level concept: '$N$ is a second-level concept' or '$N$ is a $\text{concept}_2$'
Etc.

By use of similar devices we can avoid vagueness when confronted with seemingly identical uses of words like 'exist' in relation to entities of different ontological level (e.g., 'The Moon $\text{exists}_0$,' 'The concept *horse* $\text{exists}_1$,' etc.). On this reading, the concept *horse* problem arises largely because our natural languages fail to make the appropriate logical distinctions between an object's falling under a first-level concept and a first-level concept's falling within a second-level concept. On the linguistic level, both are captured by way of the same subject-predicate form, and by use of the catch-all term 'concept'. This coarse-grained separation of entities simply cannot do justice to the logical subtleties of the realm of thoughts. Again, though, simply adding appropriate terminology to the language obviates this form of the problem. Even without adding to the natural language, Frege's *Begriffsschrift* is fully capable of stating things like 'First-level concepts exist'.

If we employed this more logically perspicuous approach, it would become immediately apparent that Kerry has provided no evidence whatsoever that one and the same entity (viz. the concept *horse*) is simultaneously treated as both an object and a concept. As Mendelsohn notes, Kerry simply cannot provide this kind of evidence in a way which ought to convince (a more self-consistent version of) Frege:

> For somewhere along the line Kerry would have to claim that there is an object $x$ and a function $f$ such that $x = f$; and, on Frege's view, placing the identity sign thus between an object variable and a function variable is incoherent. Actually, it would seem best in this situation simply to deny that ['The concept *horse* is an object'] is meaningful at all. It is entirely consistent with our interpretation that we adopt a suggestion of Geach's that 'the concept *horse*' is never serving as an *Eigenname* in the sense that 'the president' does. If a sentence containing such an expression is genuinely about a concept, then this would be captured in the symbolization by our use of *Funktionsnamen* and second-order quantifiers, and the singular term would thus be eliminated. Alternatively, if the term is ineliminable, the sentence containing it would be deemed to be nonsense.[67]

In short, Frege himself developed the logical tools to make syntactic distinctions which seem to have no previous corollary in natural languages like German and English. In responding to Kerry, rather than simply granting him the meaningfulness of sentences which treat objects as concepts, he should have referred him to precisely these syntactic distinctions. His unwillingness to alter his views of the definite article is both confusing (given his general attitude toward language) and detrimental to his own cause (as it renders his chief ontological distinction ineffable). As Mendelsohn notes, the superior response would simply have been to reject the meaningfulness of claims which identify objects with concepts in the first place and, instead, to present a more perspicuous account of the logical structure of the claim that 'The concept *horse* is a concept easily attained'.

Having shown, then, that Frege's own philosophy contains the resources to address the concept *horse* problem, let's move on to a much more difficult problem with his attempts to fix the references of mathematical terms.

## §3. The Julius Caesar objection

This second problem pertaining to Frege's ontological approach to reference-fixing is the so-called Julius Caesar objection. This objection directly concerns Frege's attempt to construe the natural numbers in terms of his ontological framework. More generally, it is relevant to his understanding of mathematical structures (like Hilbert's ostensibly uninterpreted axiomatization of geometry) as non-relational, fixed entities embedded within that framework. The first instance of this objection is formulated by Frege in §56 of his *Grundlagen der Arithmetik*. In the previous section, Frege gives the following tentative definitions of the numbers 0 and 1, as well as the successor function:

> It is tempting to define 0 by saying that the number 0 belongs to a concept if no object falls under it. But this seems to amount to replacing 0 by "no", which means the same. The following formulation is therefore preferable: the number 0 belongs to a concept, if the proposition that $a$ does not fall under that concept is true universally, whatever $a$ may be. Similarly we could say: the number 1 belongs to a concept $F$, if the proposition that $a$ does not fall under $F$ is not true universally, whatever $a$ may be, and if from the propositions "$a$

[67] [Mendelsohn 2005, 80].

> falls under *F*" and "*b* falls under *F*" it follows universally that *a* and *b* are the same. It remains still to give a general definition of the step from any given number to the next. Let us try the following formulation: the number ($n$ + 1) belongs to a concept *F*, if there is an object *a* falling under *F* and such that the number *n* belongs to the concept "falling under *F*, but not *a*".[68]

The chief aim of Frege's *Grundlagen* is to illustrate that arithmetic is a part of pure logic, and that its basic objects, concepts, and modes of reasoning are all purely logical in character. A crucial component of these logicist theses is Frege's attempt to define the natural numbers in logical terms. After having argued against a variety of rival conceptions of the natural numbers, Frege has begun in §55 to construct his own position. One of the key results of his analysis of the grammatical characteristics of statements involving number terms was the view that numbers can only be understood to be particular objects. Hence, given his understanding of the ontological difference between functions and objects, numbers cannot be construed as functions or concepts. Nevertheless, Frege admits that they do have a peculiar relation to concepts. This becomes clear in the tentative definitions above: there numbers are construed as 'belonging' to concepts. The number 0 'belongs' to all those concepts under which no objects fall, i.e., concepts whose *extensions* are empty. The number 1 and the successor function are defined in similar fashion by reference to concepts with extensions of a particular type.

Now, before we examine the Julius Caesar objection, which Frege raises immediately after stating these tentative definitions, let us recall a few fundamental facts about his understanding of definitions. Recall that Fregean definitions are, ostensibly, mere logical shorthand: they are stipulations that the sense of a previously senseless sign is to be identical with the sense of an already given sign. Essentially, definitions are identity statements whose truth is based upon stipulation. Such a view requires that the previously given sign in a definition (the *definiens*) must possess a determinate meaning in order for the definition to function properly. This is a somewhat obvious point: if a new sign is to possess the same sense as a previously given sign, then that previously given sign must have some sense to give. Now, with this in mind, let us turn to the Caesar problem.

In §56 of *Grundlagen*, Frege writes that

> The most likely [of the three definitions] to cause misgivings is the last; for strictly speaking we do not know the sense of the expression "the number *n* belongs to the concept *G*" any more than we do that of the expression "the number ($n$ + 1) belongs to the concept *F*". We can, of course, by using the last two definitions together, say what is meant by
>
> "the number 1 + 1 belongs to the concept *F*"
>
> and then, using this, give the sense of the expression
>
> "the number 1 + 1 + 1 belongs to the concept *F*"

[68] [Frege 1884b, §55].

> And so on; but we can never—to take a crude example—decide by means of our definitions whether any concept has the number JULIUS CAESAR belonging to it, or whether that same familiar conqueror of Gaul is a number or is not. Moreover we cannot by the aid of our suggested definitions prove that, if the number $a$ belongs to the concept $F$ and the number $b$ belongs to the same concept, then necessarily $a = b$. Thus we should be unable to justify the expression "*the* number which belongs to the concept $F$", and therefore should find it impossible in general to prove a numerical identity, since we should be quite unable to achieve a determinate number.[69]

Here Frege rejects these tentative definitions for the reason that they do not determine, for each object in the universe, whether or not it falls within the scope of the definition. More specifically, the last definition (that of the phrase 'the number ($n$ + 1) belongs to a concept $F$') is unsatisfactory because it contains a phrase whose meaning is as unknown as the phrase it is intended to define. At best, Frege argues, we can employ the last two definitions together to construct a system of subsequent definitions, none of which will possess a determinate sense until we clarify what it means for a number $n$ to belong to a concept $F$ in general.

This problem arises for Frege because he believes that proper definitions should determine generally whether or not an object falls within the scope of the definition. This is entailed by his view that properly meaningful statements pertain always to a single universe of discourse, namely the universe, which is characterized by Frege as an ontological hierarchy of objects and functions of various levels. Since definitions are intended to fix the references of new terms, and reference-fixing involves connecting symbols and specific elements of the ontological hierarchy, any definition which fails to determine exactly which entities fall within its scope is no definition at all. Hence, any use of the terms supposedly defined in this manner will, ultimately, be meaningless or rely upon prior assumptions for their meaning. As Frege's goal is to define the concepts of arithmetic using only logical concepts, he cannot rely upon any prior assumptions; the whole premise of his work is that an analysis of the logical structure of arithmetic vitiates the need for any special non-logical assumptions. So the Julius Caesar objection cannot be answered simply with a shrug of the shoulders.

While it might seem ludicrous to worry that we cannot tell by means of a definition that Julius Caesar is not the number 1, or that he is not a number at all, the objection is a serious one for Frege. Frege of course does not doubt that Caesar is not a number; his objection is not the epistemological one that, without a proper definition, we do not *know* that Caesar is not a number. Instead, his objection is semantic (or ontological) in nature: without a proper definition, the words we ostensibly use to refer to numbers do not clearly refer to any particular object, and hence statements employing those words will fail to be properly meaningful.[70] The objection, understood in these terms, is directly relevant to Frege's logicist project, a crucial component of which hinges upon his ability to fix the reference of number terms in an unexceptionable way.

---

[69] [Frege 1884b, §56].

[70] This semantic objection has an epistemological correlate, however. Frege's interest in the *Grundlagen* and elsewhere is to show that arithmetic is reducible to logic by showing, contra Kant, that we can gain arithmetical *knowledge* without resorting to any form of intuition. He argues in favour of this view by showing that the meanings of arithmetical statements are actually purely logical in character, and, thus, that we ought to be able to grasp those meanings by purely logical means.

More to the point, for our purposes, many of Frege's objections to the free-floating, relational view of geometric structures proposed by Hilbert are based upon his contention that Hilbert fails to properly fix the references of his terms. So, it is imperative that Frege avoid problems related to reference-fixing if he is interested in the success of his own work, as well as that of his critical attack on Hilbert-style mathematics.

Essentially, then, the problem amounts to giving a determinate sense to propositions which contain number words. As number-words can pop up in a huge variety of contexts, this is a rather large problem. Frege's approach at the outset is to make the problem more manageable by beginning with a particular class of propositions, ones which he knows *must* have a sense in order for his project to get off the ground. Given that he has already determined that numbers are objects, he believes that propositions which "express our recognition of a number as the same again,"[71] i.e., statements of identity, must be given a determinate sense. Whether or not it is practicable in every case,[72] if numbers are objects then it must always be possible (at least in principle) to determine if a given number is identical to or different from any other object. In other words, if number terms have fixed reference we should—if we properly grasp that reference in a particular case—be able to determine whether or not the reference of a sign possessing a different *sense* nevertheless has the same reference as the number term in question. This won't necessarily solve the problem of determining the sense of *all* propositions containing number terms, but it is at least a necessary first step in the process. Frege states the whole predicament as follows:

> In our present case, we have to define the sense of the proposition 'the number which belongs to the concept *F* is the same as that which belongs to the concept *G*'; that is to say, we must reproduce the content of this proposition in other terms, avoiding the use of the expression 'the number which belongs to the concept *F*'. In doing this, we shall be giving a general criterion for the identity of numbers. When we have thus acquired a means of arriving at a determinate number and of recognizing it again as the same, we can assign it a number word as its proper name.[73]

Thus, to solve his problem, Frege must figure out some method of explaining the general notion of a number belonging to a particular concept without explaining that notion in terms of the truth-value of identity statements which already employ the (as yet referentially unfixed) phrase 'the number which belongs to the concept'. Until this last condition is met, however, the Julius Caesar problem remains, and Frege's attempt to characterize arithmetic in terms of his ontological hierarchy is in danger. What we need, then, is a non-circular way of recognizing an

---

[71] [Frege 1884b, §62].

[72] Even if Frege's view that we apprehend logical truths in a purely rational manner ends up being true, it is presumably the case that, whatever our rational faculty of apprehension happens to be, it has certain empirical limits related to, e.g., the structure of the human brain. While in principle we ought to be able to determine the truth value of identity statements pertaining to extremely large numbers which are defined by repeated application of the successor function, we might in fact be physically incapable of such determinations. This, though, would be no evidence against Frege's logicist programme, any more than would be the fact that we tend to learn arithmetic through its simplistic physical applications.

[73] [Frege 1884b, §62].

object (in our case, a number) as the same again, even if the way in which the object is given (i.e., the sense of the term used to refer to the object) is quite different in separate cases.

After assigning himself this task, in §62, Frege attempts to fulfill it in §§64-68 of *Die Grundlagen der Arithmetik*. He begins by constructing a geometric example which is exactly analogous to the numerical examples we have examined above. In his examination of this example—where the direction of a line is analogous to the number belonging to a concept—Frege elaborates on the nature of his problem. He writes that:

> In the proposition 'the direction of *a* is identical with the direction of *b*' the direction of *a* plays the part of an object, and our definition affords us a means of recognizing this object as the same again, in case it should happen to crop up in some other guise, say as the direction of *b*. But this means does not provide for all cases. It will not, for instance, decide for us whether England is the same as the direction of the Earth's axis—if I may be forgiven an example which looks nonsensical. Naturally no one is going to confuse England with the direction of the Earth's axis; but that is no thanks to our definition of direction. That says nothing as to whether the proposition "the direction of *a* is identical with *q*" should be affirmed or denied, except for the one case where *q* is given in the form of "the direction of *b*". What we lack is the concept of direction; for if we had that, then we could lay it down that, if *q* is not a direction, our proposition is to be denied, while if it is a direction, our original definition will decide whether it is to be denied or affirmed. So the temptation is to give as our definition: *q* is a direction, if there is a line *b* whose direction is *q*. But then we have obviously come round in a circle. For in order to make use of this definition, we should have to know already in every case whether the proposition "*q* is identical with the direction of *b*" was to be affirmed or denied.[74]

This reformulation of the problem cuts to the heart of the matter. Here Frege insists that the way in which an object is given to us (e.g., the *sense* of the particular term we employ to refer to the object) is not a property of the object itself. The problem with defining the direction of a line by reference to the direction of another line is that, without an explicit understanding of the notion of direction, our definition requires that the object in question be given to us in a particular form in order for us to recognize it once again as the same. So, if our definition of 'the direction of *a*' is given by the proposition 'The direction of *a* is identical with the direction of *b*', then, unless the direction of *a* is given as the direction of *b,* we simply have no means of re-identifying the object in question. Our definition, then, is only useful within contexts where the relevant object is given as a direction and proves no help in grasping the general concept of the direction of a line. The situation when it comes to Frege's attempted contextual definitions of 0, 1, and the successor function is exactly parallel: using Frege's definitions we can, e.g., grasp what it means to say that the number 1 + 1 belongs to a concept *if* we already understand what it means generally for a number to belong to a concept. But, using these sorts of definitions, we get no closer to this goal insofar as they require the objects in question to be given in a particular way (i.e., already as numbers which belong to particular concepts). When confronted with objects not given in the proper way, these definitions—and, Frege suggests, contextual definitions in general—simply fail to fulfill their role. If we ask whether England is the same as the direction

[74] [Frege 1884b, §66].

of a particular line, or if Julius Caesar is identical to the natural number 1, these contextual definitions are of no help.

And, again, the problem here is not epistemological for Frege: he readily admits that, presumably, we already know that Caesar isn't a number, and that England isn't the direction of a line. The problem is a semantic one. The semantic function of a definition is to lay it down that a given sign is to possess the same sense as another already-understood sign whose reference is clear. The problem with contextual definitions like those in Frege's examples is that they fail to completely fix the reference of the new sign because there are cases where it is unclear if the definition applies or does not apply. This is because an epistemological relationship between ourselves and an object (i.e., the way in which a sign is given to us) goes proxy for direct reference to the object itself. Because the sense of a term is not a feature of the object to which that term refers, there are many cases in which the definition will be inapplicable because we use a term with a different sense to refer to the same object. In cases like this—not all of which will be as absurd as the Caesar or England examples—these purported definitions cannot help us recognize an object as the same again and, more importantly, from a semantic point of view they simply do not determine the truth conditions for (most) identity statements involving the term in question. Frege's goal, then, is to provide definitions whose strict application does not require that objects be given in a particular way.

In §68 Frege begins to outline his response to this problem. After admitting that he cannot get "any concept of direction with sharp limits to its application, nor therefore […] any satisfactory concept of Number"[75] using the methods of contextual definition, he suggests the following:

> If line *a* is parallel to line *b*, then the extension of the concept "line parallel to line *a*" is identical with the extension of the concept "line parallel to line *b*"; and conversely, if the extensions of the two concepts just named are identical, then *a* is parallel to *b*. Let us try, therefore, the following type of definition:
>
> > the direction of line *a* is the extension of the concept "parallel to line *a*";
> > the shape of triangle *t* is the extension of the concept "similar to triangle *t*".
>
> To apply this to our own case of Number, we must substitute for lines or triangles concepts, and for parallelism or similarity the possibility of correlating one to one the objects which fall under the one concept with those which fall under the other. For brevity, I shall, when this condition is satisfied speak of the concept *F* being *equinumerous* [*gleichzahlig*] to the concept *G*; but I must ask that this word be treated as an arbitrarily selected symbol, whose meaning is to be gathered, not from its etymology, but from what is here laid down. My definition is therefore as follows: the Number which belongs to the concept *F* is the extension of the concept "[equinumerous] to the concept *F*".[76]

[75] [Frege 1884b, §68].

[76] [Frege 1884b, §68]. Austin's translation has 'equal' for '*gleichzahlig*'. Despite Frege's insistence that we ought not pay attention to the etymology of his neologism, it does seem slightly misleading to render it 'equal', ignoring entirely the occurrence of '*-zahlig*' in Frege's term, as well as the difference between '*gleichheit*' (general equality) and Frege's '*gleichzahlig'*. Accordingly, where it seems appropriate I have employed the term 'equinumerous'.

Frege’s strategy here is to get rid of the context-dependence of his previous attempts at definition and replace it by direct reference to specific objects: namely the extensions of particular concepts. In the case of directions, the relevant extensions are determined by the concept of parallelism (which is, presumably, a clearly fixed concept). In the case of numbers, the relevant extensions are determined by the concept of one-to-one correlation (which, again, ought to be a concept with a clearly defined boundary). By determining explicitly when a number belongs to a particular concept, Frege may now use his previous method of characterizing 0, 1, and the successor function to explicitly define each of the natural numbers.

Ideally, we ought to be able to answer the Caesar objection directly now. Armed with this new definition, are we able to definitively rule out Julius Caesar as a member of the set of natural numbers? Certainly, Frege seems to think so. But the case is not so clear. For what Frege has done here is to require either that we have a clear understanding of the concept of an extension, or else that objects must be *given as* extensions for us to recognize them as numbers. Clearly, the latter option is not ideal, as it simply recapitulates the Caesar problem. But neither is it clear that the former is a viable solution to Frege’s woes. He himself seems not to have been entirely unaware that his proposed solution was not ideal, for in a footnote to §68, he writes: “I believe that for ‘extension of the concept’ we could write simply ‘concept’”.[77] Frege’s earlier definition of the concept of number, namely:

> The number which belongs to the concept *F* is the extension of the concept “[equinumerous] to the concept *F*,”

would then become:

> The number which belongs to the concept *F* is the concept “equinumerous to the concept *F*”.

In light of Frege’s other views, the idea that we can simply switch concepts for objects seems a rather shocking claim. Frege seems aware of the fact and notes that his suggestion

> [...] would be open to the two objections:
>
> 1. that this contradicts my earlier statement that the individual numbers are objects, as is indicated by the use of the definite article in expressions like “the number two” and by the impossibility of speaking of ones, twos, etc. in the plural, as also by the fact that the number constitutes only an element in the predicate of a statement of number;
> 2. that concepts can have identical extensions without themselves coinciding.
>
> I am, as it happens, convinced that both these objections can be met; but to do this would take us too far afield for present purposes. I assume that it is known what the extension of a concept is.[78]

[77] [Frege 1884b, §68 note 1].
[78] [Frege 1884b, §68 note 1].

First, Frege notes the obvious objection that this would contradict the entire drift of his work up to this point: his criticisms of rival views and his analysis of the grammar of statements concerning number terms have both led him to the view that numbers are particular objects and not concepts at all. The entire Caesar problem itself is phrased by Frege as that of figuring out specifically which objects number terms refer to. If he could simply have shrugged off the crucial difference between objects and concepts from the start, then why bother doing the work in the first place, and why formulate the question in terms of identity statements pertaining to objects? It seems that taking this route would suggest that we have at least two competing ways to fix the reference of number terms: one on which these terms refer to objects (i.e., extensions) and another on which they refer to concepts. But, the whole point of Frege's ferretting out a response to the Caesar problem as to definitively pin down references in the first place! This seems like a serious objection one way or another, either to Frege's claim here or to the relevance of his previous work supporting the claim that numbers are objects. In any case, it is not clear how Frege's suggestion here could work given his deep commitment to the division between concepts and objects.

The second objection, that different concepts may possess the same extension, seems to suggest that, by switching extensions for concepts, we will lose the specificity Frege had gained in his response to the Caesar problem. In Frege's original definition, particular numbers are identified with particular extensions. Despite the fact that different concepts can pick out the same extension, the extension itself is a specific object, so the fact that multiple concepts have the same extension does not call into question the determinacy given to number terms by reference to specific extensions. Now, when we begin to treat numbers as concepts, no further indeterminacy seeps in *unless we wish to claim that the two accounts are equivalent*. Concepts are, for Frege, perfectly determinate (if incomplete) entities, so that, once we identify numbers with particular concepts, there should be no difficulty determining the truth-value of statements concerning numbers. If, however, we wish the two proposed accounts (i.e., the first account involving extensions of concepts, and the second involving just concepts) to be equivalent, then a further indeterminacy does in fact arise. And this is what Frege notes in his second objection. Because different concepts may have the *same* extension, there will be many different concepts which could just as well serve for each of the natural numbers, with no particular reason to decide which ones really are those numbers, if any. This is the crucial point. If more than one concept can be identified with a particular number, then it is illegitimate to speak of *the* number *n* on our second account, and hence Frege will have lost the specificity he had been seeking. Presumably it *is* Frege's intention to claim that his extension-based account and his unspecified concept-based account are both equally serviceable, so that he will indeed have to figure out a solution to this second objection as well as the first.

Regardless of Frege's ability to answer these two particular objections (which we will not examine here), the fact remains that he recognizes a problem. This is evident, first, in his willingness to entertain the obviously problematic idea that we might switch extensions for concepts and, second, in his acknowledgement of the assumption that it is known what extensions of concepts are. Frege is generally wary (and rightly so) of the use of extensions throughout his work; in a later section he writes that his "way of getting over the difficulty cannot be expected to meet with universal approval, and many will prefer other methods of removing the doubt in question. I attach no decisive importance even to bringing in the extensions of concepts at all".[79] The assumption that it is known what extensions are (without them being given as such) would return to haunt Frege when a problem very similar to the initial Julius Caesar objection arose in his later attempt, in *Grundgesetze der Arithmetik*, to produce a formally precise version of the logicist programme discussed in the *Grundlagen.* At this point, however, it seems clear that Frege's first attempt at a solution to the Caesar objection is something of a stopgap measure, one which depends upon our recognition of extensions of concepts without them being given as such. It does not take much effort to recognize the Caesar objection lurking in a different guise here; for the problem arose, initially, when Frege could not pick out numbers unless they were given in a particular way. Here, he has simply shifted the burden onto extensions and required of us the capacity to recognize extensions no matter how they may be given.

## <u>3.1 An analogue of the Caesar problem in *Grundgesetze der Arithmetik*</u>

Let us now turn to an examination of the Caesar problem (or, at least, its close analogue) as it arises within §10 of *Grundgesetze der Arithmetik*. Up to this point in the *Grundgesetze*, Frege has attempted to familiarize the reader with his ontological scheme, as well as some of the peculiarities of his notation. As far as the formal system, he has so far only introduced the two truth-values as objects, and the functions $\xi = \zeta$, $—\xi$, and $\top\,\xi$. In §9, Frege introduces the smooth breathing notation for courses-of-values: the course-of-values of the function $\Phi(\varepsilon)$ is represented as $\dot{\varepsilon}\Phi(\varepsilon)$.[80] He ends that section by noting that "The introduction of a notation for courses-of-values seems to me to be one of the most important supplementations that I have made of my Begriffsschrift since my first publication on this subject".[81] So, while Frege was obviously quite unsatisfied with his use of extensions of concepts in his earlier work, here he seems to stand firmly behind the usefulness of the more general notion of a course-of-values. In the section immediately following this introduction of the notation for courses-of-values, Frege raises the analogue of the Caesar problem and attempts to solve it.

This stretch of the *Grundgesetze*, entitled 'The course-of-values of a function more exactly specified,' contains Frege's attempt to answer the nagging questions about extensions left unanswered by his proposed solution to the Caesar problem in the *Grundlagen*. He begins by stating a version of that familiar problem, this time applied to expressions for courses-of-values rather than directions or numbers. In this case, the particular expression is that for the course-of-values of an arbitrary function $\Phi(\varepsilon)$. Frege writes:

> Although we have laid it down that the combination of signs
>
> '$\acute{\varepsilon}\Phi(\varepsilon) = \acute{\alpha}\Psi(\alpha)$'
>
> has the same denotation as

[79] [Frege 1884b, §107].
[80] Note that, due to a seemingly insurmountable typesetting problem, my representation of the smooth breathing notation (here and throughout) does not exactly match Frege's own.
[81] [Frege 1893, §9].

> '$\forall\alpha(\Phi(\alpha) = \Psi(\alpha))$,'[82]
>
> this by no means fixes completely the denotation of a name like '$\acute{\varepsilon}\Phi(\varepsilon)$'. We have only a means of recognizing a course-of-values if it is designated by a name like '$\acute{\varepsilon}\Phi(\varepsilon)$', by which it is already recognizable as a course-of-values. But we can neither decide, so far, whether an object is a course-of-values that is not given us as such, and to what function it may correspond, nor decide in general whether a given course-of-values has a given property unless we know that this property is connected with a property of the corresponding function.[83]

Here, again, we encounter a problem very similar to the Caesar objection: we have the tools to recognize the course-of-values of a particular function, but only if that course-of-values is given in a particular way. In this case, we are licensed to recognize two courses-of-values, $\acute{\varepsilon}\Phi(\varepsilon)$ and $\acute{\alpha}\Psi(\alpha)$, as one and the same only if they are given as courses-of-values (i.e., if they are referred to using Frege's smooth-breathing notation). In cases where these same entities are given differently, Frege's contextual definition affords us no means of recognizing their identity. The problem however is deeper than the epistemological one of our recognizing the entities as the same again: for, if we are given only the contextual definition to characterize courses-of-values, statements like '$\acute{\varepsilon}\Phi(\varepsilon)$ = Julius Caesar' will not possess a truth-value. Note, too, that not all cases of this sort will be as absurd as the Caesar example: for, if numbers are identified with extensions (as Frege suggests), then statements of the form '$\acute{\varepsilon}\Phi(\varepsilon) = 2$' will be of crucial importance for establishing the foundations of arithmetic. Using Frege's contextual definition, such statements are neither true nor false—indeed, they are, by Frege's lights, completely senseless until we determine more carefully the meaning of terms of the form '$\acute{\varepsilon}\Phi(\varepsilon)$'.

Frege's goal, then, is to provide a *general* criterion for determining whether an object is a course-of-values or not. As before, Frege has reduced this problem to one of fixing the reference of particular types of identity statements. Ostensibly, if we can determine the truth or falsity of an identity statement containing terms which refer to courses-of-values *without the condition that those terms be of a particular form*, then we will have provided a general criterion for recognizing courses-of-values. The contextual definition above allows one to determine the truth of identity statements concerning entities which are given to us already as courses-of-values by stipulating that the reference of these statements (i.e., their truth value) is to be the same as a statement whose sense is already grasped (e.g., the statement that the two functions in question always take the same value for the same argument). But, crucially, this contextual definition does not and cannot determine the truth-value of identity-statements concerning entities which are not already given as courses-of-values. We cannot, for instance, determine whether or not Julius Caesar is identical to any course-of-values simply by using the definition above. Frege's overall point in raising these objections is to insist that, without a solution to this sort of problem, the terms at issue here are actually senseless, and sentences containing them will therefore fail to refer to a truth-value. Hence we are not licensed to (and, by Frege's lights—are actually

[82] For ease of typesetting, here and throughout, I use the common modern notation for the universal quantifier (i.e., '$\forall$') in place of Frege's idiosyncratic symbolization.
[83] [Frege 1893, §10].

incapable of) making inferences which employ the purported ‘thoughts’ about courses-of-values as premises. As we’ve seen, when using courses-of-values to solve the Caesar problem in the *Grundlagen*, Frege had simply assumed that it was known what the extension of a concept was (and, perhaps too, what the course-of-values of a function was). His remarks there, and elsewhere, seem to indicate that he was not entirely comfortable with this situation. Now, in what was to be the culmination of his life’s work, Frege appears ready to face the problem head on. Whereas previously he attempted to shrug off the Caesar objection via reference to the extensions of concepts, here he raises the problem for courses-of-values, and hence for extensions as well. His earlier remedy will thus not work here, and so a new tactic is required.

Frege’s method for avoiding the problem of contextual definition when it comes to courses-of-values is to pin them down more precisely by clarifying their relationship to functions. More specifically, his suggestion is that, for each function, when it is introduced, we ought to determine what value it takes for courses-of-values as arguments. Doing so will clarify the truth value of any statement containing a course-of-values term, eliminating any indeterminacy left unsettled by his previous solution to the Caesar problem. Let’s examine how this is supposed to work.

Since, up to this point in the *Grundgesetze*, Frege has only introduced three functions, the required determinations are relatively brief. The three functions Frege has discussed up to this point are: $\xi = \zeta$, $—\xi$, $\top \xi$; i.e., identity, the horizontal, and negation. Essentially, Frege’s concern here is to determine for each of these functions what value it will take when a course-of-values is input as argument.

The negation function can be discounted from consideration immediately, for, on Frege’s dispensation, the function “can be considered always to take a truth-value as argument”.[84] This is because Frege defines negation as follows: “The value of the function $\top \xi$ shall be the False for every argument for which the value of the function $—\xi$ is the True; and shall be the True for all other arguments”.[85] Further, the function $—\xi$ is defined as follows: “$—\Delta$ is the True if $\Delta$ is the True; on the other hand it is the False if $\Delta$ is not the True”.[86] So, since the argument for the function $\top \xi$ is defined in terms of the values taken by the function $—\xi$ (which are, by stipulation, always truth values), any argument taken by the function $\top \xi$, even if it is an object other than a truth value, can nevertheless be treated as a truth-value. The function $—\xi$, too, can be eliminated from consideration here insofar as it can be reduced to the function $\xi = \zeta$. Frege writes that “by our stipulations the function $—\xi$ has for every argument the same value as the function $\xi = (\xi = \xi)$; for the value of the function $\xi = \xi$ is the True for every argument. From this it follows that the value of the function $\xi = (\xi = \xi)$ is the True only for the True as argument, and that it is the False for all other arguments, exactly as with the function $—\xi$.”[87] So, then, we are left only with the function $\xi = \zeta$ to consider. Frege’s problem, again, is to determine what value this function gives for courses-of-values taken as arguments.

[84] [Frege 1893, §10].
[85] [Frege 1893, §6].
[86] [Frege 1893, §5].
[87] [Frege 1893, §10].

3.2 The restriction problem

At this point in his argument, Frege makes some remarks which are rather difficult to interpret, given certain key commitments of his philosophy. He writes:

> Since in this way everything reduces to consideration of the function $\xi = \zeta$, we ask what value this has if a course-of-values occurs as argument. Since up to now we have introduced only the truth-values and courses-of-values as objects, it can only be a question of whether one of the truth-values can perhaps be a course-of-values. If not, then it is thereby also decided that the value of the function $\xi = \zeta$ is always the False if a truth-value is taken as one of its arguments and a course-of-values as the other. If on the other hand the True is at the same time the course-of-values of some function $\Phi(\xi)$, then it is thereby also decided what the value of the function $\xi = \zeta$ is in all cases in which the True is taken as one of the arguments, and likewise if the False is at the same time the course-of-values of a certain function.[88]

Here Frege has reduced his problem to the determination of the value taken by the function $\xi = \zeta$ for courses-of-values. His first concern is to determine whether the only two objects hitherto introduced in his formal system (i.e., the two truth-values) are themselves courses-of-values. His reasoning regarding the values taken by the function $\xi = \zeta$ if either of these objects is or is not a course-of-values is unexceptionable: if neither truth-value is a course-of-values, then the function gives the False as value when a course-of-values is taken as one argument and a truth-value as the other. But if either truth-value is in fact a course-of-values of some function, then we also have a determinate value for the function when one of the truth-values is taken as an argument; this value will be the True when both argument places are filled by names for one of the truth-values (be they referred to by terms of the form '$\acute{\varepsilon}\Phi(\varepsilon)$' or otherwise), and the False in all other case. This of course requires that we have picked out precisely which courses-of-values the two truth-values are. All this is fine, as far as it goes.

The problematic portion of the above quotation is Frege's restriction of his consideration to just these two objects. His suggestion seems to be that, because he has only introduced the two truth-values as objects thus far, he need only decide whether either of these is itself a course-of-values to determine the value taken by the function $\xi = \zeta$ for courses-of-values, and thereby, also, to fix the reference of the term '$\acute{\varepsilon}\Phi(\varepsilon)$' completely. This seems to imply that the functions so far introduced only take as arguments the objects which have been introduced by Frege at this point in the *Grundgesetze*; i.e., that Frege's functions have a highly restricted domain. Taken at face-value, this seems to directly contradict Frege's views that functions have an ontological status which is not dependent upon their being given in a particular way (e.g., in Frege's Begriffsschrift). Here he seems to be restricting the domain of functions in a way which depends upon the status of his formal system. More than this, the restriction also seems to conflict with his demand that functions be defined for every object in the universe—a demand which seems to be one of the sources of the Caesar problem in the first place. While Frege's views allow for the

[88] [Frege 1893, §10].

'restriction' of functions in a certain sense, such restrictions do not involve the exclusion as arguments of objects without names in a particular formal system.[89]

Moreover, it seems to imply that the problem carried over to the *Grundgesetze* (because of Frege's inability to provide a general solution to the Caesar objection in the *Grundlagen*) is in fact a pseudo-problem. For, if a particular formalization determines which objects a function takes as arguments, then we ought not be concerned with the fact that our definitions cannot determine for an arbitrarily given object—not just those which have been introduced into the system or given names—whether or not they fall under the definition. If Caesar isn't part of the system, or if there is no name for Caesar in Frege's formalization of number theory, for instance, then there simply is no problem in determining the value of the function $\xi = 2$ for Caesar as an argument. In the terms of the *Grundgesetze,* we may just as well have avoided the problem of fixing the reference of terms of the form '$\acute{\varepsilon}\Phi(\varepsilon)$' by never introducing such terms in the first place.

So what exactly is going on here? Some commentators[90] have suggested that the best way to understand this restriction is to characterize it in what we would now call meta-linguistic terms. That is, they suggest that the restriction is a meta-linguistic statement concerning which *names* have thus far been introduced into the system.

But it seems clear from what Frege explicitly states that it is not merely a matter of names:

> Since up to now we have introduced only the truth-values and courses-of-values *as objects*, it can only be a question of whether one of the truth-values can perhaps be a course-of-values.[91]

Taken on its own, this fairly clearly shows that Frege is not concerned with which names have been introduced into the system up to this point, but with which objects have been introduced. Moreover, if we keep in mind that this restriction is undertaken within the context of the Caesar problem, then we have good reason to suspect that the meta-linguistic reading misses the mark.

If Frege was merely interested in the linguistic features of his formal language, the Caesar problem could never have arisen in the first place, as it is fundamentally not a linguistic problem concerned with which sorts of statements are formulable within the confines of *Begriffsschrift*. The problem is rather concerned with our relation to the *references* of such statements. Or, as Richard Heck puts it,

---

[89] One way in which a Fregean function might be 'restricted' to, e.g., the natural numbers is by defining the function in the intended manner for the natural numbers as arguments (presuming we have a rigorously circumscribed concept of natural number to work with), while the function takes on the value the False (or some dummy value) for objects which are not natural numbers. The key here is that the function remains defined for all objects in the universe, and that the so-called restriction requires a referentially fixed concept of natural number (or whatever) to get off the ground.

[90] Cf., e.g., [Moore and Rein 1986].

[91] [Frege 1893, §10]. The introduction of objects (rather than names) is equally clear in the German: "Da wir bisher nur die Wahrheitswerthe und Werthverläufe *als Gegenstände* eingeführt haben, so kann es sich nur darum handeln, ob einer der Wahrheitswerthe etwa ein Werthverlauf sei" [Frege 1893, §10, emphasis mine].

> What the Caesar problem suggests is that we have no understanding of what numbers are that supports the idea that they are independent of how they are given to us, that is independent of our capacity for thought about them.[92]

As Heck's remarks here help to show, the Julius Caesar problem is not merely a bug arising from Frege's particular choice of formal system. It is instead an essential problem which cuts to the core of his attempt to understand the nature of mathematics. If he is avoid the kind of free-floating, purely formal understanding of mathematics which he sought to combat in his engagements with Hilbert, Frege must be able to show that mathematical terms refer to something in particular. The Caesar problem arises precisely because Frege is unwilling to allow contextual definitions (which do rely on the peculiarities of a particular context, language, or formalism) as a means of fixing the references of mathematical terms. Thus, it would be incredibly odd for Frege, right at the point when he begins to solve the problem, to restrict his solution to the merely linguistic peculiarities of a particular stage of his formal system.

But, even if we reject the meta-linguistic reading, we are still presented with a problem. For, while Frege does not restrict his solution to the linguistic realm, he does nevertheless restrict himself to considering only the objects which are introduced to his formal system at a given stage. Given Frege's methodological scruples, such a restriction seems fundamentally at odds with his goal of solving the Caesar problem once and for all. How, then, are we to understand Frege's move here?

### 3.3 Heck's Argument Against Frege's 'Universalism'

Let's try to reconstruct the situation we seem to be presented with here. If Caesar is in the domain of the functions Frege considers here, then of course these functions have to be defined for Caesar as an argument. On the reading of Frege I have presented thus far in this dissertation, he holds that a function must be defined for *every* object in the universe. So, straightforwardly, Caesar must be in the domain of the functions Frege is about to consider. But the restriction of his attention to the few objects he has already introduced seems to indicate that, for Frege, Caesar is not in fact in the domain of these functions. This situation leaves us with at least two options.

1. Something is wrong in the above reconstruction of the situation.
2. Frege has (knowingly or otherwise) committed himself to an inconsistent position.

Before we leap to either conclusion, let's try to re-evaluate our reconstruction of the situation. Richard Heck, in a paper examining §10 of the *Grundgesetze* offers an alternative account of the restriction which involves re-evaluating the 'universalist' component of our reconstruction. That is to say, Heck wants to call into question the view that Frege holds that a function always has to be defined for every object in the universe. If Heck's argument is correct, then Frege's restriction is an unproblematic one. How, then, does Heck think that we can deny Frege's universalism? Rather straightforwardly, he states:

[92] [Heck 1999, 266, note 16].

> When [Frege] says that we have 'introduced only the truth-values and value-ranges as objects', he means to tell us that the quantifiers in his formal language are, at this point, to be taken as ranging only over such objects.[93]

For a modern logician, mathematician, or model theorist this sort of restriction of the domain of quantification is a matter of course. For Frege, however, this seems difficult, as every 'model' must ultimately be embedded within a single ontological hierarchy. So in what sense can the domain of quantification be restricted for Frege?

Well, let's look at what Heck has to say about the motivations for Frege's 'universalism'. He writes:

> Frege understood some of the logicians of his time to be proposing to deal with restricted quantification by making use of what the Booleans called 'universes of discourse'.[94] Frege objected vehemently to this practice. He insisted, firstly, that sentences involving restricted quantification—e.g., 'Every real number is *F*'—can be interpreted as generalized conditionals, in now familiar fashion: That is, he argues that there is no *need* to invoke a restricted domain, containing only the reals, for this purpose. He argues, secondly, that such sentences *should* be interpreted as generalized conditionals. Treating restricted quantification in terms of universes of discourse makes it difficult to represent the validity of certain sorts of inferences: For example, 'Every real number has a non-negative square' implies 'Any number which has a negative square is non-real'. At the very least, then, some way needs to be provided for the 'universe of discourse' to be represented in the logical forms of such sentences.[95]

As the italics indicate here, Heck's suggestion is that Frege's opposition to restricted quantification is not fundamental but, instead, relative to particular examples where this sort of quantification seems to obscure logical relations rather than clarify them. Heck's argument for this reading, it seems to me, involves imputing a significant amount of modern, model-theoretically sophisticated logic into Frege's work. While Heck himself admits that this reading requires "only a dollop of anachronism"[96], the dollop seems rather a large one. Let's look a bit closer.

---

[93] [Heck 1999, 272-273].

[94] Boole himself introduces the notion of a universe of discourse in *The Laws of Thought*. There he writes "In every discourse, whether of the mind conversing with its own thoughts, or of the individual in his intercourse with others, there is an assumed or expressed limit within which the subjects of its operation are confined. The most unfettered discourse is that in which the words we use are understood in the widest possible application, and for them the limits of discourse are co-extensive with those of the universe itself. But more usually we confine ourselves to a less spacious field. […] Now, whatever may be the extent of the field within which all the objects of our discourse are found, that field may properly be termed the universe of discourse" [Boole 1854, Chapter III, §4]. As will be made clearer below, Frege's point is that the restriction of quantification is only undertaken within *the* universe (i.e., within his ontological hierarchy), and, since logic is concerned with the most general laws of truth, such restrictions, too, have to be cemented clearly within that universe.

[95] [Heck 1999, 273].

[96] [Heck 1999, 274].

The view is that Frege does not object to all uses of restricted quantification, but only to those which involve the confusion of domains of quantification with universes of discourse. Let's look at what Heck says about universes of discourse.

> The logical form of 'Every number is *F*' is therefore much more nearly: '$\forall x(Nx;Fx)$. The first argument of the quantifier specifies the 'universe of discourse', in the sense relevant to the analysis of restricted quantification.[97]

His suggestion here is that the logical role of '*Nx*' is to specify a universe of discourse. On this view, presumably, the naked quantifier '$\forall x$' indicates the bare universe, while *Nx*' is meant to whittle this immense totality down to those entities which fall under the concept *N*, whatever that might be. Thus the universe of discourse, on this view, is always a subset of the universe properly speaking. In mathematics and logic, of course, the universe of discourse understood in this way is almost never indicated explicitly. Universes of discourse, understood in this way, are superfluous on Heck's reading of Frege. For, in our example, if the concept *N* is well-defined, then any statement about the predicate-restricted universe of discourse can simply be rephrased in terms of a generalized conditional (in a manner similar to that of Frege's proposed emendations of Hilbert's views).

So, Heck admits that Frege has good reasons to eschew the use of 'universes of discourse' understood in the sense of predicate-restrictions placed on the universal quantifier. We are left, then, to consider the availability of variable 'domains of quantification' to Frege. Heck writes that

> Obviously, though, even if we make it explicit, in the syntax, that the quantifier can be […] 'restricted' by some predicate, that does not make domains of quantification, in the sense relevant to model-theory, superfluous: One can—indeed, *must*—still allow for varying domains of quantification when defining validity, implication, and the like, for the usual sorts of reasons.[98]

This, it seems to me, is where Heck's argument requires more than his suggested dollop of anachronism, for he has not attempted to show that variation in the domain of quantification (in this sense) is required for *Frege's* understanding of validity, implication, and the like. This is the crucial point in the argument, and it rests on the reader's accepting that Frege's notions of 'validity, implication, and the like' are sufficiently similar to their modern model-theoretic cousins that they will require varying the domain of quantification "for the usual sorts of reasons". But, as I hope to have shown at several points in this thesis, there are good reasons not to assume that Frege's logical notions are primitive versions of contemporary notions. On the face of it, then, it seems suspect to simply import the modern model-theoretic conceptions of these notions into Frege's view, which elsewhere (e.g., in his correspondence with Hilbert) seems highly critical of the philosophical underpinnings of an enterprise quite like model theory in important respects.

[97] [Heck 1999, 274].
[98] [Heck 1999, 274].

So, it seems, neither the meta-linguistic reading of the restriction nor the anti-universalist reading suggested by Heck offer a clear explanation of our problem. Without good reasons for thinking otherwise, then, we are left with the suspicion that Frege's restriction here is simply in conflict with his established views.

3.4 Frege's attempted solution to the Caesar problem in the *Grundgesetze*

Regardless of the status of this restriction relative to the rest of Frege's philosophy, he does employ it in his attempted solution to the Caesar problem. Let's return to our examination of the rest of this solution, with the caveat that it may hinge in part on this inconsistency.

After further reducing his problem via the restriction, Frege sets about determining whether or not either of the truth-values is a course-of-values. He notes that he cannot decide whether or not either of the truth-values is a course-of-values by appeal to his stipulation that "'$\acute{\varepsilon}\Phi(\varepsilon) = \acute{\alpha}\Psi(\alpha)$' is to have the same denotation as '$\forall\alpha(\Phi(\alpha) = \Psi(\alpha))$'.[99]

Frege then engages in a technical argument intended to illustrate that "it is always possible to stipulate that an arbitrary course-of-values is to be the True and another the False"[100] and, hence, that the answer to his question is that the truth-values themselves are (or, rather, can be taken to be) courses-of-values. Let's examine his argument supporting this claim. He writes that:

> it is possible to stipulate generally that '$\tilde{\eta}\Phi(\eta) = \tilde{\alpha}\Psi(\alpha)$' shall denote the same thing as '$\forall\alpha(\Phi(\alpha) = \Psi(\alpha))$' without the identity of $\acute{\varepsilon}\Phi(\varepsilon)$ and $\tilde{\eta}\Phi(\eta)$ being derivable from this. We should then have a class of objects with names of the form '$\tilde{\eta}\Phi(\eta)$', and for whose differentiation and recognition the same distinguishing mark held good as for courses-of-values.[101]

This first step is simply a generalization of our earlier problem: instead of considering the singular case, Frege moves to consider an entire class of objects with names of the form '$\tilde{\eta}\Phi(\eta)$'. After having determined such a class, he defines a function $X(\xi)$ as follows:

| | |
|---|---|
| $X(\tilde{\eta}\Lambda(\eta))$ | = the True |
| $X$(the True) | $= \tilde{\eta}\Lambda(\eta)$ |
| $X(\tilde{\eta}M(\eta))$ | = the False |
| $X$(the False) | $= \tilde{\eta}M(\eta)$ |

For all other arguments, the value of the function is simply the argument itself.

Now, the functions $\Lambda(\xi)$ and $M(\xi)$ can either take the same value for every argument or they can differ in value for some arguments. In the case where they differ in value for at least one object, then the

[99] [Frege 1893, §10].
[100] [Frege 1893, §10].
[101] [Frege 1893, §10].

> [...] function $X(\xi)$ never has the same value for different arguments, hence '$X(\tilde{\eta}\Phi(\eta)) = X(\tilde{\alpha}\Psi(\alpha))$' also always has the same denotation as '$\forall\alpha(\Phi(\alpha) = \Psi(\alpha))$'. The objects whose names were of the form '$X(\tilde{\eta}\Phi(\eta))$' would then be recognized by the same means as the courses-of-values, and $X(\tilde{\eta}\Lambda(\eta))$ would be the True and $X(\tilde{\eta}M(\eta))$ the False.[102]

Essentially, what Frege has done here is to construct a function which licenses him to give the truth-values names of the same form as those for courses-of-values, and thus to treat the truth-values as courses-of-values. But this license does not uniquely pick out which course-of-values the truth-values are, for we could very well have defined a function $X'(\xi)$ such that the roles of $\tilde{\eta}\Lambda(\eta)$ and $\tilde{\eta}M(\eta)$ are reversed, thereby identifying the truth-values with different courses-of-values. In other words, Frege's argument shows only that we *can* identify the truth-values with particular courses-of-values, it doesn't allow us to take the further step of picking out precisely which courses-of-values these might be, other than arbitrarily. Accordingly, he writes that:

> Thus without contradicting our setting '$\acute{\varepsilon}\Phi(\varepsilon) = \acute{\varepsilon}\Psi(\varepsilon)$' equal to '$\forall\alpha(\Phi(\alpha) = \Psi(\alpha))$' it is always possible to stipulate that an arbitrary course-of-values is to be the True and another the False. Accordingly let us lay it down that $\acute{\varepsilon}(\text{—}\varepsilon)$ is to be the True and $\acute{\varepsilon}(\varepsilon = (\top\forall\alpha(\alpha = \alpha))$ is to be the False. $\acute{\varepsilon}(\text{—}\varepsilon)$ is the course-of-values of the function $\text{—}\xi$, whose value is the True only if the argument is the True, and whose value for all other arguments is the False. All functions for which this holds, have the same course-of-values, and this is by our stipulation the True. Accordingly $\text{—}\acute{\varepsilon}\Phi(\varepsilon)$ is the True only if the function $\Phi(\xi)$ is a concept under which falls only the True; in all other cases $\text{—}\acute{\varepsilon}\Phi(\varepsilon)$ is the False. Further, $\acute{\varepsilon}(\varepsilon = \top\forall\alpha\,(\alpha = \alpha))$ is the course-of-values of the function $\xi = \top\forall\alpha\,(\alpha = \alpha)$, whose value is the True only if the argument is the False, and whose value for all other arguments is the False. All functions for which this holds have the same course-of-values, and this is by our stipulation the False. Thus every concept under which falls the False and only the False, has as its extension the False.[103]

Here Frege stipulates precisely which courses-of-values he wishes to identify with the True and the False. His earlier argument was designed to illustrate that it is always possible to identify the True and the False with an arbitrary course-of-values, so that is precisely what he does. Ostensibly, then, the *Grundgesetze* version of the Caesar problem is solved, and Frege writes that "With this we have determined the courses-of-values so far as is here possible".[104]

Let's review Frege's reasoning thus far. In order to precisely fix the reference of terms of the form '$\acute{\varepsilon}\Phi(\varepsilon)$', Frege argues that we must determine the value taken by all functions for such terms taken as arguments. Now, up to this point Frege has introduced only three functions, and consideration of these can, in fact be reduced to the consideration of a single function, namely the identity function $\xi = \zeta$. Frege's problem, then, is to figure out what value this function takes for courses-of-values as arguments. Just as he has only introduced a few functions thus far in the *Grundgesetze*, so too he has only introduced a few objects: the True and the False. Frege's problem, if we allow him his restriction, thus ultimately reduces to figuring out two things: first,

---

[102] [Frege 1893, §10].
[103] [Frege 1893, §10].
[104] [Frege 1893, §10].

whether or not the truth-values are themselves courses-of-values and, second, what values the function $\xi = \zeta$ takes for the truth-values as arguments if they *are* courses-of-values. To solve his problem, then, Frege constructs a function designed to licence him to identify the truth-values with arbitrary courses-of-values. With this license, he then stipulates precisely which courses-of-values they are, thus allowing us to decide the values the function $\xi = \zeta$ takes for courses-of-values. At the end of this chain of reasoning, Frege believes he has fixed the reference of terms of the form '$\acute{\varepsilon}\Phi(\varepsilon)$' as precisely as is possible, and thus that he has solved the Caesar problem as it applies to courses-of-values for his system as elaborated up to this point. By stipulating the values taken by functions for courses-of-values as arguments whenever new functions are introduced, the solution should hold for further elaborations of the system as well.

## 3.5 An insoluble problem

But, with historical hindsight, we now know that there is trouble yet lurking in Frege's project in the form of Russell's paradox. It is something of a commonplace to suggest that the ultimate failure of Frege's project rests on the inconsistency which results from this paradox. And, to be sure, the discovery of a contradiction in his formal system was a significant blow to Frege's life's work. But, I will argue (in substantial agreement with Richard Heck)[105] that the contradiction is merely an effect of the real failure, which is Frege's inability to satisfactorily solve the Caesar problem and, hence, his inability to fix the references of mathematical terms.

The proximal source of Russell's paradox in Frege's work is his use of Basic Law V. This axiom can be stated as follows:

$$\acute{\varepsilon}F(\varepsilon) = \acute{\varepsilon}G(\varepsilon) \equiv \forall x(Fx \equiv Gx)$$

Or, in plain English, the course-of-values of a function $F$ is identical to the course of values of a function $G$ if, and only if, the functions $F$ and $G$ take the same values for every object. Frege employs this axiom in a number of places throughout the *Grundgesetze*, but there are only two which are ineliminable, and these are the source of the problem.

The two uses of Basic Law V which cannot be eliminated, i.e., which are not merely uses of convenience, are connected to Frege's attempted derivation of Hume's principle. Thus it is only in his attempt to prove this principle that Frege's system is threatened by Russell's paradox.

The term 'Hume's principle' was coined by George Boolos,[106] but the principle itself does have its source in Hume. In his 1739 *Treatise*, Hume gives the following definition of equality between numbers:

> When two numbers are so combin'd, as that the one has always an unite answering to every unite of the other, we pronounce them equal.[107]

---

[105] Cf. [Heck 2005]; [Boolos 1993, 230] also clearly establishes the importance of the Caesar problem for understanding the collapse of Frege's work.
[106] [Boolos 1987a, 137].
[107] [Hume 1896, Book I, Part III, §1].

Frege in fact cites this passage from Hume in §63 of the *Grundlagen*. Hume's formulation of the principle is less relevant to us than its analogue in the philosophy of Frege. Put in Frege's language of concepts and objects, this principle holds that the number of objects falling under a concept *F* is the same as the number of objects falling under a concept *G* if, and only if, the *F*s can be put in a one-to-one correspondence with the *G*s. As noted by Charles Parsons and others, the bulk of Frege's system of arithmetic is derived from Hume's principle and the axioms other than Basic Law V.[108] Moreover, if we drop Basic Law V, there are a number of different *consistent* fragments to be found in Frege's system.[109] So, if the source of Frege's problems is Basic Law V, and this axiom is only used essentially in a very few cases, why not simply drop it and replace it with Hume's seemingly very similar principle?[110]

The answer is that, by using only Hume's principle in his definition of number, Frege simply cannot solve the Caesar problem. Let's see why not.

In Frege's work, Hume's principle first arises in connection with his attempts to fix the meaning of identity-statements between numerical terms. As we've seen, fixing the meaning of such statements is a crucial step in Frege's larger attempt to fix the references of mathematical terms. And, at first glance at least, Hume's principle seems like an ideal candidate for a solution to this problem. "But," Frege writes, "it raises at once certain logical doubts and difficulties."[111] The chief difficulty, as he sees it, is that Hume's principle seeks a special definition of numerical identity, or 'equinumerosity'.

In some well-known passages from the *Grundlagen*, Frege attempts to define the notion of the direction of a line using the sort of limited, contextual definition suggested by Hume's principle. He writes of his proposed definition that

> If […] we were to adopt this way out, we should have to be presupposing that an object can only be given one way [i.e., *as a direction* for directions, or *as a number* for numbers]; for otherwise it would not follow, from the fact that *q* was not introduced by means of our definition, that it *could* not have been introduced by means of it. All identities would then amount to this, that whatever is given to us in the same way is to be reckoned as the same. This, however, is a principle so obvious and so sterile as not to be worth stating. We could not, in fact, draw from it any conclusion which was not the same as one of our premisses.[112]

Such definitions, then, require objects to be given in a particular way, thus making them relative to a given system, language, or formalization. But this is exactly what Frege is seeking to avoid (as is evident from his rejection of the formalism suggested by Hilbert and Korselt). In order to

---

[108] Cf. [C. Parsons 1995, 196ff]; [Boolos and Heck 1998] gives a bit more detail.
[109] Cf., e.g., [Boolos 1987a], [Boolos1987b], [T. Parsons 1987], [Heck 1996] and [Wehmeier 1999].
[110] [Heck 1995] argues convincingly that Frege himself knew that almost all of his uses of Basic Law V were matters of mere convenience.
[111] [Frege 1884b, §63].
[112] [Frege 1884b, §67].

truly solve the Caesar problem, then, Hume's principle is insufficient and an explicit definition must be given.

And so it is that Frege is put in an impossible position. He must either employ an explicit definition at some point in his derivation of Hume's principle or he must abandon his search for a context-independent way of fixing the references of mathematical terms.

He chose the first route which famously leads to Russell's paradox and the inconsistency of his system. By rejecting Basic Law V in favour of Hume's principle, however, we are no closer to solving the Caesar problem. Thus, without Basic Law V, Frege saw no way of actually fixing the references of mathematical terms, other than relative to a given system. We are, therefore, left again at an impasse in our attempt to understand what mathematics is about. The free-floating, structure-focussed road taken by Hilbert seems to lead to the threat of meaninglessness (without a background ontology) and the relativisation of truth; the fixed-reference path taken by Frege leads either to contradiction and incoherence or back once more to the meaninglessness of Hilbert's position.

## §4. Deadlock

Now, how does the ultimate failure of Frege's attempt to solve the Julius Caesar problem affect the outcome of his debate with Hilbert over the nature of axiomatics and mathematics more generally?

As we've seen above, many of Frege's criticisms of Hilbert focus on the fact that Hilbert is unable (and unwilling) to pin down the references of the terms of an axiomatic system, prior to the process of interpretation. Hilbert's view, in fact, was based on the (now widely-held) belief that the schematic or uninterpreted character of formal mathematical systems was one of the chief benefits of axiomatization. Given the fact that Frege's own best efforts were unable to fix the references of the key terms of even his modest fragment of mathematics, his criticisms of Hilbert might seem to lose much of their persuasiveness. If Frege himself is unable to do what he demands of Hilbert, and Hilbert's method appeared incredibly productive to the mathematical community at large, what is there left to recommend Frege's view? Why shouldn't we simply adopt Hilbert's view?

But I believe we are left with a more profound difficulty in the wake of Frege's inability to solve the Caesar problem. One lesson to be learned from a detailed examination of the Frege-Hilbert correspondence is that the difficulties faced by Frege here are much more obviously present in Hilbert's work. The fact that *neither* Frege nor Hilbert seem capable of fixing the references of the terms of their axiomatic systems does not mean that the uninterpreted, purely formal point of view simply wins the day—quite the contrary! We are left, instead, with a much more carefully specified problem lying at the heart of both approaches to mathematics.

What Frege's Caesar problem shows us, then, is a problem which remains open in the philosophy of mathematics. Namely, the problem of specifying exactly what mathematics is *about*. Frege's approach to this problem was to suggest that, ultimately, mathematics (or, at least a crucial and potentially foundational component of mathematics) was reducible to logic. Since his view of logic was an ontological one, mathematics, too, had a clear ontological basis: it was about some of the more general features of the world. But, as we have just seen, this view is not tenable; the Caesar problem shows us that Frege, too, is incapable of pointing out exactly which entities mathematics deals with, and it remains unclear what exactly mathematics is about.

Hilbert, by contrast, held the view that (at least in its more mature forms) mathematics is ultimately concerned with certain formal structures and their inter-relations. Through our examination of the Frege-Hilbert correspondence, however, we have seen that this view is plagued with difficulties quite similar to those faced by Frege in his own work. When trying to become clear about exactly what a formal structure is, Frege noted several dangerous equivocations that he attempted to solve by situating such structures within his logical/ontological hierarchy. But, as his inability to solve the Caesar problem shows, this route is simply incapable of shoring up Hilbert’s account of mathematics. Hilbert’s work throughout the 1920s was similarly focussed on the attempt to provide a clear referential basis for higher forms of mathematics, though his formalist/proof-theoretic programme, too, faced insurmountable difficulties. Just as Frege’s best efforts led to what he himself saw as a dead-end, so too Hilbert’s own approach came to seem mathematically and philosophically unsatisfying.

In the following chapter, we will examine some of the ways in which more recent work in mathematics and the philosophy of mathematics has tended to repeat (rather than solve) this dichotomy.

# Chapter 4: Benacerraf, abstract structures, and structuralism

## §0. Introduction

In previous chapters, we have examined two distinct views of mathematics offered by Hilbert and Frege in their correspondence and other works. Frege's view entails that mathematics deals with a specific content, and that the meaningfulness of mathematical statements relies in part on reference to an ontological hierarchy of functions and objects. Hilbert's view, by contrast, resists Frege's attempt to fix the references of mathematical terms and treats mathematics as the study of the properties of and relations between formal axiomatic structures which cannot (or should not) be referentially fixed. As we saw in these chapters, both Hilbert's relational view and Frege's ontologically-grounded view face profound difficulties related to meaning and reference.

The primary difficulty facing Hilbert's view was the threat of meaninglessness and mathematics as a contentless science. As Frege pointed out in their correspondence, without references for their terms, statements within an axiomatic theory would appear to be meaningless and incapable of possessing an absolute truth value. While both Hilbert and his defender Korselt argued in favour of a view of axiomatic systems as uninterpreted formal structures, Frege countered by demanding that even statements within such formal structures require a determinate meaning if they are to become relevant for scientific inquiry. On Frege's view, the applicability and scientific pedigree of mathematics required that it be more than a complex game of symbols. Frege's suggestion was to embed the referentially unfixed statements of Hilbert's uninterpreted systems within his ontological hierarchy. But, as we saw, Hilbert resisted this move, viewing it as contrary to the very point of axiomatizing a theory in the first place (namely, providing a flexible structure applicable in multiple distinct situations).[1] Thus Hilbert's view is faced with a dilemma. Axiomatic theories are either purely formal systems and, hence, meaningless prior to being given an interpretation, or they possess some specific content and require at least some of their terms to have fixed references, thus eliminating their flexibility. Frege saw no way out of the first horn of the dilemma, and instead dedicated his attention to the second.

It is worth noting here that, despite his initial hesitance to accept Frege's demand for fixed reference, Hilbert's later work attempted to achieve something quite similar. Throughout the 1920s and early 1930s Hilbert and his collaborator Paul Bernays sought to embed higher mathematics (particularly any mathematics which required the use of infinites) in a 'contentual' (*inhaltlich*) system of stroke-symbols which, it was presumed, were epistemologically and semantically accessible.[2] The guiding idea of Hilbert's Program (as it has come to be called) was that, even if the statements of higher mathematics are meaningless or unverifiable, given a properly constructed syntactic theory of proof we can ensure that the allowable moves in the formal system always lead us from contentual truths, through a formal game, and back again to contentual truths. By grounding mathematics within his system of stroke symbols, Hilbert sought to ensure that any detour through the infinite would always loop back to some verifiable,

---

[1] Cf. [Frege 1980, 40-41].

[2] The most important papers by Hilbert and Paul Bernays which develop proof theory and the formalist programme are collected (in English translation) in [Mancosu 1998], with insightful commentary by Mancosu. See [Hilbert 1919-1920, chapter 9] as well.

meaningful claim in the realm of the finite. Just as Frege sought to give meaning to Hilbert's purely formal conception of axiomatics by embedding his supposedly uninterpreted statements within an ontological hierarchy, so Hilbert's later work in proof theory attempted to ground all higher mathematics on a fixed foundation of straightforward reference.[3]

Unfortunately, this interesting programme suffered a fate rather similar to that of Frege's own work on the foundations of mathematics. Whereas the inability to solve the Caesar problem had destroyed Frege's hope of embedding mathematics within his ontological hierarchy, Hilbert's program was undermined in fundamental ways when Kurt Gödel—who, somewhat ironically, had been a great contributor to proof theory—announced his incompleteness results in the autumn of 1930. What Gödel had shown (with some help from John von Neumann)[4] was that a consistency proof of the system of finitary reasoning (i.e., the system of 'contentual' mathematics) was impossible without already appealing to transfinite, 'non-contentual' modes of reasoning.[5] This effectively reduced Hilbert's Program to another dead end in the search for mathematical foundations.[6]

But, just as Hilbert's later program faces serious difficulties, so too Frege's view (which grounded his suggested emendations of Hilbert's axiomatization of geometry) is subject to deep difficulties. Frege encountered seemingly intractable problems in his effort to fix the references of the terms employed in his logicist reduction of basic arithmetic. As we saw in the previous chapter, his attempts to solve various forms of the Julius Caesar problem eventually required the use of his Basic Law V, leading directly to the inconsistency of his logical system. As a result, Frege ultimately abandoned his life's work, and published very little in his remaining years. In post-paradox work, there were many attempts to avoid the paradoxes implied by the 'naïve' understanding of set theory and logic. This is one of the chief aims of Russell and Whitehead's *Principia Mathematica*[7] and an important part of Hilbert's program, too. But as we've just seen, these projects, too, ran against the insoluble difficulty of Gödel's incompleteness results, which of course relate directly to *Principia*.

---

[3] Note that my gloss of Hilbert's Program here ignores the important epistemological aspect of the project. In addition to seeking consistency and completeness proofs, Hilbert also claimed that his system of stroke symbols was "surveyable" [cf. Hilbert 1922], so that the age-old question of epistemological access to the abstract entities of mathematics disappeared. [Kitcher 1976] presents a fair account of the epistemological side of the project, as well as an account of some of its failings.

[4] See [Gödel 1931] for the initial publication of these results. Von Neumann's help first appears in letters written to Gödel between November of 1930 and January of 1931. These are available in [Von Neumann 2005, 123-126].

[5] Gerhard Gentzen's celebrated 1936 proof of the consistency of first-order Peano arithmetic, for instance, employs transfinite induction.

[6] Hilbert's collaborator Bernays almost immediately recognized the destructive significance of the results. For a thorough examination of Gödel's effects on Hilbert and Bernays' programme, see [Tait 2006, §4 and §8]; for a broader look at Bernays and Gödel's correspondence regarding finitary mathematics and the formalist programme, see [Feferman 2008]. Despite the widespread acceptance of the failure of Hilbert and Bernays' particular program, there have been a number of interesting attempts to develop the program in a more satisfactory form (e.g., [Feferman 1992] and [Simpson 1988]). Of course the success of the broader idea of a formalized proof theory within logic, mathematics, and computer science can hardly be overstated.

[7] "A very large part of the labour involved in writing the present work has been expended on the contradictions and paradoxes which have infected logic and the theory of aggregates" [Russell and Whitehead 1963, vii].

We are left, then, at something of an impasse. It turns out that *both* the relational view of Hilbert and the ontologically grounded view of Frege face tremendous difficulties, particularly in relation to meaning and reference-fixing. We've traced the emergence of a structure-focussed view of mathematics and examined two of the most influential ways of explaining what mathematical structures might be, but both views face great problems. Without significant alteration, neither Frege's fixed ontological view nor Hilbert's uninterpreted view are capable of satisfactorily explaining what much of modern mathematics is about.

Throughout the 1920s and 1930s there was a pervasive feeling that the foundations of mathematics had been seriously undermined. This deepening sense of crisis was brought about by a variety of factors: the emergence of non-Euclidean geometries, the prominence of the ill-understood technique of axiomatization, the set-theoretic paradoxes, and, finally, Gödel's incompleteness results of the 1930s.[8] But, even despite the widespread belief in a foundational crisis, mathematics continued to develop at a rapid pace. Though problems with set theory continued to be raised by both philosophers and mathematicians, it was, for the most part, employed as an acceptable, if philosophically troublesome, foundation for mathematics.[9] Set theory had been axiomatized by Ernst Zermelo between 1905 and 1908,[10] and subsequently altered by Abraham Fraenkel and Thoralf Skolem in 1922.[11] The resulting system, Zermelo-Fraenkel set theory (ZF) acted as the de facto basis for mathematics for three quarters of a century, and still figures prominently in foundational work and philosophical debates today.[12] Between the 1930s and the 1960s general foundational questions about mathematics were, for the most part, formulated in set-theoretic terms. Most mathematicians and philosophers interested in foundational questions during this period tended to concern themselves with the specific problems of set theory and its basic concepts. All of this set-theoretic work was undertaken despite the many unresolved conceptual problems with set theory (many of which were quite similar to those faced by Frege's logical system). Solutions to these problems moved away from the so-called naïve set theory of the late 19th century and towards more explicit axiomatic treatments like ZF, or the less widely used Von Neumann-Bernays-Gödel axiomatization (NBG) developed between the mid-1920s and the 1950s, though even in these systems many conceptual problems remained. The general view among those concerned was that 'higher' mathematics was ultimately reducible to the basic concepts of set theory, and that

---

[8] While it is a historical commonplace to note that there was a foundational crisis in mathematics in the 1920s (see, e.g., [Weyl 1921] for one of the earlier and best-known expressions of this view), relatively fewer accounts focus on the continuity between the factors mentioned above. An interesting exception is [Ferreirós 2008]. As earlier chapters in this thesis attest, I hold that the conceptual crisis in mathematics was in full swing already by the time of the Frege-Hilbert correspondence, and well before the public debate surrounding the set-theoretic paradoxes.

[9] This sort of attitude ("We'll figure out why it works later") is remarkably similar to the rapid expansion in the application of the calculus prior to its rigorization in the 1800s. Most mathematicians in both periods remained unconcerned with what we might call the 'philosophical' underpinnings of their mathematical tools and were instead content to extend their successes as far as possible. In contrast to the calculus, however, it seems to me that we have yet to achieve a satisfactory philosophical foundation for modern axiomatics.

[10] The chief reference here is [Zermelo 1908]; almost all of the key papers surrounding the early axiomatization of set theory are available in [Van Heijenoort 1967].

[11] For an illuminating history of the development of set theory, see [Ferreirós 1999].

[12] For example, Stewart Shapiro's theory of structures is self-consciously modelled on ZF (cf., [Shapiro 1997, 93-97]), though he departs from the usual philosophical understanding of ZF as a foundation.

problems with this reduction were simply a matter of finding the correct axiomatization of set theory or the proper understanding of the relevant concepts.[13]

Throughout the first half of the twentieth century (and perhaps a few decades beyond), the philosophy of mathematics was dominated by an intense interest in set-theoretic reductionism as a means of putting mathematics on a solid foundation. When people (or, at least, philosophers) asked "what is mathematics *really* about?" the answer was, invariably, "Sets!".[14]

## 0.1 Set-theoretic reductionism and set-theoretic foundationalism

But what does this mean? What exactly is set-theoretic reductionism and how does it relate to the desire to properly found mathematics? We'll need to answer a few sets of general questions about the nature of reduction and foundations in mathematics to understand the role of set theory in the post-Fregean world.

First question: what does it mean to reduce one part of mathematics to another? To answer this question, let's look at the most often discussed reduction: the reduction of the natural numbers and their arithmetic to set theory. How is this reduction accomplished? The basic idea is that for every arithmetical object, property, or relation of the natural number system, there is a corresponding object, property, or relation in the universe of whichever set theory we are working with. Thus, the entire sequence of numbers itself has to be replaced by or reduced to some collection of sets which bears properties and relations that (somehow) mirror those of the natural numbers. In essence, a set-theoretic reduction of the natural numbers and their basic arithmetic involves modelling the natural number structure within some particular universe of sets. This of course poses a number of questions we've been examining throughout this thesis (i.e., what is the natural number structure, how do we figure out the 'somehow' in the previous sentence, what are models, etc.). We'll return to these concerns below.

---

[13] Of course there were exceptions to this general tendency. [Putnam 1967] attempts to eschew the search for foundations in mathematics entirely. An earlier and more focussed rejection of specifically set-theoretic foundations can be found in [Skolem 1922], where Skolem first introduces his eponymous 'paradox'. Not really a paradox at all, Skolem's theorem illustrates the startling relativity of set-theoretic notions. For him, this relativity introduced a profound and confusing complexity into set theory which was directly at odds with most set theorists' dreams of reducing all of mathematics to the notions of set theory. Of foundations, Skolem writes that "Our only concern, then, should be that the initial foundations be something immediately clear, natural, and not open to question. This condition is satisfied by the notion of integer and by inductive inferences, but it is decidedly not satisfied by set-theoretic axioms of the type of Zermelo's *or anything else of that kind*; if we were to accept the reduction of the former notions to the latter, the set-theoretic notion would have to be simpler than mathematical induction, and reasoning with them less open to question, but this runs entirely counter to the actual state of affairs" [Skolem, in Van Heijenoort 1967, 299, my emphasis]. Unsurprisingly, Zermelo was not thrilled with Skolem's work, initially believing it to contain a mistake [cf., Van Dalen and Ebbinghaus 200, 145-146].

[14] Wilson makes a similar claim which includes at least the 20th century: "set theory has become the canonical backdrop to which questions of structural existence are referred. So when we go beyond napkins and set up a structure […] formally, we need both: (a) to mark the 'intrinsic' elements proper to the internal workings of the structure, without extraneous fat; (b) to provide sufficient extrinsic links to set theory so that the existence of our structure can be established" [Wilson 1993, 209]. Shapiro seems to extend the dominance of set theory even to the 21st century: "Within the community of professional mathematicians, if not philosophers, a set-theoretic proof of satisfiability resolves any legitimate questions of existence" [Shapiro 2005, 72].

Second question: what is the point of reducing one part of mathematics to another? There are at least three related answers to this question which are relevant here. We might call them the epistemological, the pragmatic/heuristic, and the ontological answers, respectively. I'll isolate them here for clarity, but it should be noted that they are really rather difficult to separate (as Frege's work very clearly shows).

The epistemological answer to the question suggests that reducing one area of mathematics to another can help to clarify, simplify, or solve epistemological problems. Frege's work is probably the best example of reductionism undertaken for the sake of epistemology; we might also think of parts of Hilbert's Program in this way. One of the central aims of Frege's work prior to the emergence of Russell's paradox was precisely to prove an epistemological point via the reduction of arithmetic to logic. By showing that arithmetic can be understood in purely logical terms, Frege hoped to show that arithmetic is not a synthetic *a priori* science, as Kant famously argued, but that it is instead a purely analytic discipline. To use the Kantian vocabulary, his reduction of arithmetic to logic would have shown that arithmetic had no intrinsic connection to the pure intuition of time. So, at least part of his goal was to overturn an error he saw in Kant's influential understanding of the epistemology of mathematics, with all the implications this might have for the philosophy of mathematics.

We might also reduce one part of mathematics to another for pragmatic or heuristic reasons. There are numerous instances in the history of mathematics where a concept or technique from one branch or sub-discipline has proven essential to the advancement of seemingly unrelated work in another branch or sub-discipline of mathematics.[15] Such cross-pollination is crucial for mathematics. But, at least since the early 1800s, mathematics has become a field of increasingly isolated specialization, such that someone working in one branch of mathematics might be unable even to understand (let alone use) results in another – despite the fact that these results may be of dramatic importance for their particular problem. This poses a pragmatic problem: how can mathematicians from disparate fields speak to one another? One solution to this problem is to recast these disparate fields in a common language which could provide the common ground required for fertile cross-disciplinary collaboration. As the category-theorist and philosopher F. William Lawvere puts it, "Unification and simplification are necessary not only for the dissemination of results, but also for the coherent advance of research in the diverse branches of mathematics."[16] Thus, if mathematicians can cash out their higher-level mathematical discussions in terms of a single theory or way of speaking, then the recognition of similarities between seemingly disparate parts of mathematics will be made considerably easier. By reducing all of mathematics to a single theory (e.g., set theory or category theory), the pragmatic reductionist gives all mathematicians a basic vocabulary to communicate with one another. A related benefit of this form of reductionism is that, when a surprising connection between fields is in fact discovered, it ought to be explicable in the uniform and universal language developed by the reduction (e.g., one might explain the powerful connections between

[15] The application of techniques for solving equations to geometric problems via Descartes' coordinate system is a particularly well-known example of this phenomenon; another fruitful example is the deep connection between permutation groups and the roots of polynomial equations which forms the basis of Galois theory.

[16] [Lawvere 2000, 715].

algebra and geometry of analytic geometry by showing how both equations and spaces can be understood in set-theoretic terms). So pragmatic/heuristic reductionism can help to explain connections, and, the better explained such connections are, the better exploited, and the more productive the cross-pollination becomes.

Reductionism can also be ontological in character. This sort of reduction aims to show that our discussions of higher-level mathematical entities are really and ultimately about something simpler (things like sets or categories or numbers). So when we do mathematics what we are *really* doing is engaging with and speaking about this most fundamental ontological layer. Ontological reductions in this sense aim to show us that everything other than set theory or category theory or (in Frege's case) logic is merely a *façon de parler*, a historically entrenched, partially misleading linguistic fiction. Only the final level of the reduction touches reality. One clear benefit of this sort of view is ontological parsimony: in place of the bewildering array of abstract objects and structures developed by mathematics (Abelian categories, fuzzy sets, Hausdorff spaces, Fermat primes, Grothendieck universes, Blotto games, etc., etc.) we have one or two sorts of things (e.g., sets, objects and arrows, objects and functions, numbers, etc.). This parsimony itself has the added benefit of simplifying semantic and epistemological concerns: instead of working things out on a case by case basis, we can solve these sorts of problems in a uniform way. Perhaps the chief philosophical benefit, though, is that this kind of ontological reduction promises to tell us what mathematics is *really* about.

Third set of questions: what do we mean when we speak of the foundations of mathematics? What are such foundations meant to provide? How do the various forms of reductionism act as foundations? In some ways we have given answers to these questions already.

Epistemological reductions are meant to provide (unsurprisingly) epistemological foundations for mathematics. Let's again take Frege as an example. Part of Frege's worry was that mathematicians actually have no idea what they are talking about when they engage in proofs. Indeed, much of the first half of Frege's *Grundlagen der Arithmetik* is taken up with showing that extant views of the concept of number are nonsensical. As a result, proofs based upon these views don't really prove anything because they are meaningless concatenations of symbols. If we can show that important parts of mathematics are reducible to logic, and that logic is epistemologically accessible, then we can show that it is possible to make mathematics more epistemologically secure. If we know what numbers are, and follow rigorous and gapless proof procedures, then we can ensure that mathematicians aren't engaging in flights of fancy. We can really prove things now. This is how Frege's reduction was meant provide an epistemological foundation for mathematics.

Pragmatic reductions can provide pragmatic foundation for mathematics. By showing that we can, if we so choose, cash out our higher level talk in terms of set- or category-theoretic talk, when we collaborate with others we have a firm basis for thinking that this collaboration is adding to the stock of mathematical knowledge in general, rather than just the stock of our particular nook or cranny of mathematics. This pragmatic reduction guarantees that mathematics is not, in fact, splintering apart into a thousand dissonant voices like an abstract Tower of Babel. Many of the proponents of category theory as a foundation for mathematics have emphasized its pragmatic use as a kind of *lingua franca* for the mathematics.[17] A pragmatic foundation of this sort is meant to give a foundation for the working unity of the science. This won't necessarily guarantee that the tower is stable, nor that it touches reality, but it can help guarantee that we are engaged, in some sense, in the same global project.

Finally, ontological reductions can give us ontological foundations. When we show that a new mathematical concept or object can be reduced to the single ongoing ontology of our particular reduction, we know that our utterances about these novelties actually refer to something real and that they are therefore meaningful and capable of truth and falsity. Ontological reduction aims to show that the abstract claims of mathematics have a foundation in reality, in much the same way that the claims of physics and zoology have. By showing it is possible to reduce higher-level mathematical entities to an acceptable ontology, we provide a means for connecting mathematics to the world that we inhabit. Such reductions help to answer the question "What is mathematics about?" by answering "Reality".[18]

In this simplified view of the nature of mathematical foundations, we already have three distinct approaches. Part of the problem in the philosophy of mathematics over at least the last century and a half has been that these approaches often intermingle to form a confusing tangle, so that ontological questions are posed in pragmatic terms, epistemological questions have ontological implications, and so on. In this dissertation thus far, I have focussed on the question: what is mathematics really about? I have meant this question to be taken ontologically, i.e., I have meant to ask which features of reality (if any) does mathematics concern itself with? If we think of the debate between Hilbert and Frege as chiefly an ontological one—and I suggest that we should—then we have already seen a few answers to this question. Frege's answer is that the ontology of mathematics (or at least that part of mathematics derivable from basic arithmetic) is just the ontological hierarchy of objects and functions. Hilbert's suggestion is that mathematics, taken on its own, *has* no ontology—indeed, that this lack of ontology is its chief feature. We've seen, too, that both of these answers are deeply problematic. So, again, we've returned to our impasse.

## §1. Benacerraf on the natural numbers

In the next few sections I would like to see whether the more recent history of the philosophy of mathematics can offer us plausible new ways of answering the ontological question. I've chosen as a starting point an influential paper by the American philosopher of mathematics Paul Benacerraf. The 1965 paper, "What Numbers Could Not Be" has been cited over eight hundred

[17] Mac Lane's well-known textbook "Categories for the Working Mathematician" is perhaps the best known application of this view. There Mac Lane "aims to present those ideas and methods that can now be effectively used by mathematicians working in a variety of other fields of mathematical research. […] [C]ategories provide a convenient conceptual language, based on the notions of category, functor, natural transformation, contravariance, and functor category" [Mac Lane 1998, vii]. In an appendix to the text, Mac Lane also attempts to show that categories "can be used as a possible foundation for all of mathematics, thus replacing the use in such a foundation of the usual Zermelo-Fraenkel axioms for set theory" [Mac Lane 1998, 290]. The view of category theory as a conceptual language or organizational tool for large parts of mathematics is also developed in a more philosophical vocabulary in [Landry 1999] and [Mac Lane 1992].

[18] Of course such reductions often entail a fairly expansive notion of 'reality' – thus, for instance, an ontological account of the traditional set-theoretic reduction will include as part of 'reality' the entire set-theoretic universe *V*.

times (something of a feat in the rather cloistered world of the philosophy of mathematics) and it seems to have heralded a new way of understanding the problem of foundations in mathematics.[19]

At the height of the popularity of set-theoretic reductionism, Benacerraf raised a significant conceptual problem with the usual set-theoretic reduction of the natural numbers. He noted what was already common knowledge: that, even within Zermelo-Fraenkel set theory, there can exist multiple different, equally valid ways in which to define the natural numbers as sets. Benacerraf proceeded to argue that the existence of multiple different reductions poses a problem for the naïve view of set-theoretic foundationalism, namely the view that there is some unique correlate in the set-theoretic hierarchy corresponding to each higher level mathematical domain, in this case the natural numbers. If natural numbers are sets, Benacerraf asked, well, just which sets are they?

In many ways his is a very charitable approach to the problem. At the outset Benacerraf concedes almost everything to reductionism. Sets are real? Fine. Numbers are sets? Fine. All that he wishes to know is which particular sets they are. If there is no principled reason for picking out particular sets as the real correlates for particular numbers, then there seems to be no reason to accept set theory as the final ontological foundation for mathematics in the first place. This, in broad strokes, is what has come to be called the non-uniqueness problem. Its effect is to call into question the post-Fregean, post-paradox attempts by set theorists and philosophers to explain just what it is that mathematics is about. In the following section, I will investigate Benacerraf's account of this problem in more detail, and also look at his admittedly tentative 'structuralist' suggestions toward its solution. The point of looking so closely at Benacerraf's paper is to see whether his work offers us any clear alternative to the impasse we encountered in the Frege-Hilbert debate.

## 1.1 Benacerraf's non-uniqueness problem

The focus of Benacerraf's paper is a single, relatively straightforward question: if numbers are reducible to sets, then which sets are they?

In one sense, this question is as easily answered as it is asked, since in any reduction of the natural numbers to the universe of set theory, it is a simple matter to make explicit which numbers are identified with which sets. This is indeed the chief point of such reductions. The real problem for Benacerraf begins when multiple adequate though extensionally distinct set-theoretic reductions are available. Benacerraf examines two fairly common approaches, one attributed to Ernst Zermelo, the other to John von Neumann.

On the Zermelo reduction, we have the following identifications:

$$\varnothing = 0$$
$$\{\varnothing\} = 1$$

[19] One of the chief contemporary proponents of structuralism, Stewart Shapiro, suggests that Benacerraf's paper "provides the standard motivation for structuralism" [Shapiro 2005, 61].

{{Ø}} = 2
{{{Ø}}} = 3

Whereas on the von Neumann account, we have:

Ø = 0
{Ø} = 1
{Ø, {Ø}} = 2
{Ø, {Ø}, {Ø, {Ø}}} = 3

We needn't get into great detail about the particulars in each reduction, but it is worth noting two general facts here. First, each reduction is entirely adequate in the sense that all the (set-theoretic correlates of the) relations and properties of the natural numbers can be deduced in either reduction. And, second, many of the 'same' elements in each reduction possess quite different properties. For instance, on the von Neumann account, the number 3 has three members, one of which is the set {Ø, {Ø}}. On the Zermelo account, the number 3 has a single member, {{Ø}}, which is not a member of the von Neumann reduction's 3. In short, then, we possess supposedly equivalent arithmetic reductions which are both extensionally and intensionally different in many respects. Beyond these two reductions, which Benacerraf employs mainly for illustrative purposes, there are of course an infinite number of other equally valid set-theoretic reductions of the natural numbers. Moreover, we needn't restrict ourselves to set theory; we might employ category theory, lattice theory,[20] or even a physical structure of spacetime points to explain just what the natural numbers are.[21]

The question, then, is this: if there are multiple equally valid but radically different reductions of the natural numbers, which of these is *the* correct one? Which sets are the natural numbers really? Benacerraf writes

> We are left in a quandary. We have two (infinitely many, really) accounts of the meaning of certain words ("number," "one," "seventeen," and so forth) each of which satisfies what appear to be necessary and sufficient conditions for a correct account. Although there are differences between the two accounts, it appears that both are correct in virtue of satisfying common conditions. If so, then the differences are incidental and do not affect correctness. Furthermore, in Fregean terminology, each account fixes the *sense* of the words whose analysis it provides. Each account must also, therefore, fix the *reference* of these

[20] There are several ways to define the natural numbers in terms of many general algebraic structures, not just set theory. Since category theory, for example, can model set theory (cf. [Osius 1974]), one could simply employ the category-theoretic correlate of, say, the von Neumann reduction in a definition of the natural numbers. Alternatively, one can define the natural numbers directly in terms of a category containing a terminal object, a global element, and arrows with the requisite properties. Similarly, one can define the natural numbers (including zero) in lattice-theoretic terms as a semi-lattice under the operations 'min' and 'max', with 0 as the 'bottom'.

[21] Cf. [Field 1980]. Field likely would not treat his use of this spacetime structure as an attempt to explain what the natural numbers are (given that the title of the book is *Science Without Numbers*), though this structure *is* intended to fulfill the duties the (real) numbers generally perform within the physical sciences. Someone with less interest in nominalizing science could certainly employ Field's structure in ways very similar to the set-theoretic reductions we refer to here.

> expressions. Yet, as we have seen, one way in which these accounts differ is in the referents assigned to each of the terms under analysis.[22]

At this stage Benacerraf points out that, on the assumption that numbers just are sets (or that they are ultimately reducible to sets), we are left with only two possible options. Either both (or all) of these accounts are correct or it is not the case that both (or all) are correct. At the outset he rejects the idea that multiple referentially-fixed accounts of the natural numbers can be correct in the required sense. This is because the point of the ontological reduction is to pick out the exact objects which the numbers are, not a group of objects which share certain similarities—unless of course that group is a specific set which uniquely fulfills the relevant criteria, but Benacerraf's point is that this is not in fact possible in the first place. If the numbers *are* sets, then they must be particular sets. Benacerraf's position here is straightforwardly ontological in character. Since no two distinct entities can also be identical—at least by a plausible and fairly common metaphysical assumption[23]—it simply cannot be the case that the number 3 is both the set {{{Ø}}} and the set {Ø, {Ø}, {Ø, {Ø}}},

> [f]or if the number 3 is in fact some particular set *b*, it cannot be that two correct accounts of the meaning of "3"—and therefore also its reference—assign two different sets to 3. For if it is true that for some set *b*, 3 = *b*, then it cannot be true that for some set *c*, different from *b*, 3 = *c*.[24]

So, bypassing the first option of rejecting the indiscernibility of identicals, Benacerraf examines the second. If numbers just are sets, and it is not the case that there are multiple, equally correct accounts of which sets exactly they are, then there must be a single correct account. If this is the case, Benacerraf suggests that there ought to be some reason for preferring this account over all the others, which all suffice equally well for the purposes of arithmetic. He writes

> We are now faced with a crucial problem: if there exists such a "correct" account, do there also exist arguments which will show it to be the correct one? Or does there exist a particular set of sets *b*, which is *really* the numbers, but such that there exists no argument one can give to establish that it, and not, say, [the von Neumann reduction], is really the numbers? It seems altogether too obvious that this latter possibility borders on the absurd. If the numbers constitute one particular set of sets, and not another, then there must be arguments to indicate which.[25]

---

[22] [Benacerraf 1965, 55-56].

[23] Though, in the philosophy of mathematics, not a universally held assumption. Cf., e.g., [MacBride 2006], [Ketland 2006], [Leitgeb and Ladyman 2008], and [Shapiro 2008].

[24] [Benacerraf 1965, 56].

[25] [Benacerraf 1965, 57-58]. It seems to me that a more reasonable claim (or at least one which is more agnostic about the relation between human capacity and the world of mathematics) would suggest that if we are to *believe* that the numbers constitute one particular set of sets, and not another, then there must be arguments or at least evidence to indicate which. In a later appraisal of the argument of the 1965 paper, Benacerraf himself notes "no red-blooded realist […] should accept the bald statement that if there isn't some *a priori* proof that some particular candidate reduction is the correct one, there can't be a 'correct' one" [Benacerraf 1996b, 25].

As this remark shows, Benacerraf's dilemma is remarkably similar to Frege's Caesar problem. Frege, in his attempt to embed the natural numbers within his logical/ontological hierarchy, was compelled to provide completely determinate identity statements for each number, such that it was clear for every entity in his hierarchy whether or not it was identical with the number in question. His problem was to ensure that he picked out exactly which entities he wished to talk about, not simply a cluster of entities having some selection of shared properties sufficient to satisfy a collection of axioms. He wanted to show, for instance, which exact entity the number 2 was, not give us a list of different possible entities which sufficed to play the role of the number 2 up to isomorphism. Given his logical scruples, this desire for strictly fixed reference required that his definitions of the numbers exhaustively determined not only which entities they were, but also which entities they were not.

Benacerraf's aim here is quite similar. He wants to be able to show that, among the many arithmetically equivalent reductions of the natural numbers, one and only one is the correct reduction—the one which actually tells us which sets the natural numbers are (or, if we prefer, what we really refer to when we use number words). Benacerraf, again following Frege, is not interested in privileging a particular reduction for purely aesthetic or idiosyncratic reasons, or out of convenience; he demands principled argumentation illustrating the superiority of one reduction over all others. Since the set-theoretic differences between, e.g., the von Neumann and Zermelo reductions are irrelevant from a purely arithmetical point of view they cannot be used to militate in favour of either reduction as the superior account. How then can we distinguish between the two in anything other than an *ad hoc* or idiosyncratic fashion? As we saw in the last chapter, Frege's attempts to solve his version of this problem led him into difficulties which he eventually viewed to be insurmountable. Though Benacerraf is not bound to the particularities of Frege's ontological view of logic, nor to Frege's logicist program, his problem here is nevertheless almost identical in structure to Frege's. So what is Benacerraf's strategy for avoiding the dead end of Frege's approach without falling into the meaninglessness of a purely formal approach like Hilbert's? How can we determine precisely which sets the natural numbers are?

### <u>1.2 Restricted identity</u>

Benacerraf outlines a few ways we might deal with the non-uniqueness problem. He begins, again in a familiar Fregean mode, to investigate exactly which question it is that we are asking here. An important part of what we really want to know is the truth value of certain identity statements of the form '$n = \ldots$' where $n$ is some number. There are, Benacerraf suggests, at least three different types of identity statement we might wish to investigate here. First are the more commonplace identities between obviously arithmetical entities, i.e., statements like '2 + 2 = 4,' where each component of the statement is presented in recognizably arithmetical form. Second, we have more circuitous but ultimately similar claims, like '2 = the number of shoes currently on my feet,' where we have obvious reference to a numerical entity without the use of a numeral or strictly arithmetic term. Neither of these statement-types is particular troubling; in a given reduction, the (relative) truth value of these statements will be uncontroversial. The third sort of identity statement is rather more problematic. These statements, e.g., '2 = Julius Caesar', or '5 + 8 = {Ø, {Ø, {Ø}}}' are familiar from Frege's Caesar problem. Indeed, Benacerraf explicitly refers to Frege's problem in his discussion of these identities:

> To speak from Frege's standpoint, there is a world of objects—that is, the designata or referents of names, descriptions, and so forth—in which the identity relation had free reign. It made sense for Frege to ask of *any* two names (or descriptions) whether they named the same object or different ones. Hence the complaint at one point in his argument that, thus far, one could not tell from his definitions whether Julius Caesar was a number.[26]

Benacerraf's first move in responding to the non-uniqueness problem is to distance himself precisely from Frege's requirement that every identity statement is meaningful (and, relatedly, that each such statement has a determinate truth value). Benacerraf suggests that identity statements of the third type are 'unsemantical' for the reason that there is no context in which there are determinate truth-conditions for such statements.[27] He writes that

> If an expression of the form "$x = y$" is to have a sense, it can be only in contexts where it is clear that both $x$ and $y$ are of some kind or category $C$, and that it is the conditions which individuate things *as the same C* which are operative and determine its truth value.[28]

And, further, that

> What determines that something is a *particular lamppost* could not individuate it as a *particular number*. I am arguing that questions of the identity of a particular "entity" do not make sense. "Entity" is too broad. For such questions to make sense, there must be a well-entrenched predicate $C$, in terms of which one then asks about the identity of a *particular C*, and the conditions associated with identifying $C$'s as *the same C* will be the deciding ones.[29]

Here again we see Benacerraf attempting to break free of Frege's constraint that every identity statement must be meaningful. He does so by suggesting that, in order for there to be an identity relation between two entities, there must be some context in which it is made clear that there is a semantically acceptable category under which both entities fall. In the case where there is no such category available, we cannot make sense of the identity relation, and identity statements can't be given a truth value. Presumably, a statement like '2 = Julius Caesar' will be neither true nor false on this view, but, instead, 'unsemantical'. This demand for a uniform semantic category within which to subsume the terms of a meaningful identity statement is broadly similar to Frege's insistence that functions have specific types of argument places which cannot be satisfied by the wrong sort of entity. Thus, for instance, we saw Frege engaged in a great deal of fancy footwork to avoid the implication (of the concept *horse* problem) that his view allowed meaningful identity statements between objects and concepts. Frege, for his part, attempted to explicate as best he could the ontological reasons for these formal conditions. Benacerraf, likewise, is presented with the difficulty of specifying (in a uniform way) exactly which

---

[26] [Benacerraf 1965, 64].
[27] Note that by 'unsemantical' Benacerraf does not mean strictly nonsensical, for, as he notes, "we grasp enough of [the] sense [of unsemantical claims] to explain why they are senseless" [Benacerraf 1965, 64]. He doesn't further define this notion, unfortunately. We will return to this difficulty below.
[28] [Benacerraf 1965, 64-65].
[29] [Benacerraf 1965, 65].

categories are semantically acceptable and which are ‘unsemantical’. Unlike Frege, Benacerraf does not give us any detailed story about how this might be accomplished, though his use of the phrase ‘well-entrenched predicate’ is suggestive that his account would differ significantly from Frege’s.

To avoid this problem entirely, one might simply define a disjunctive category or predicate (e.g., ‘*x* is an *F* just in case *x* is either a Roman general or a number’) which includes whichever entities we wish to make identity claims about. Benacerraf, however, wishes to reject this possibility as well, despite his admitted inability to give a precise account of what he means by ‘uniform conditions’ sufficient to prevent such *ad hoc* counterexamples.[30]

1.2.1 Problems with Benacerraf’s restriction of identity

The restriction of statements of identity to properly semantical contexts promises to solve the non-uniqueness problem, at least in part, by allowing us to ignore thornier questions about mathematical objects like Frege’s Caesar problem. What the Caesar problem and others like it show is that we cannot straightforwardly pick out mathematical entities like numbers precisely enough to pin them down in an absolute ontological sense. Benacerraf’s suggestion that we restrict the relation of identity to semantical contexts allows us to say, ‘There really isn’t an answer to the question “Is Julius Caesar a number?” because we are dealing here with an ill-defined, unsemantical context within which the (or *an*) identity relation can gain no purchase’. The problem, then, simply dissolves and we are free to discuss the truth or falsity of identity claims *within* various well-defined contexts like the Zermelo and Von Neumann reductions.

But there are clear difficulties in this line of thinking. Let’s take a look at a few of these before we proceed to Benacerraf’s next attempt to address the non-uniqueness problem.

As we’ve seen, Benacerraf’s restriction of identity leans heavily on ‘uniform semantic conditions’ and ‘unsemanticality’. The notion of uniform semantic conditions for the meaningfulness of identity statements is meant to protect his account from the introduction of *ad hoc* solutions to the problem. Such solutions would simply dictate by fiat which sorts of identity claim are allowed and which are not. But, on this view, anything which was deemed to be unsemantical (e.g., ‘Julius Caesar = 2’) could simply be made meaningful by the introduction of an appropriately defined context.[31] To avoid all this, Benacerraf demands semantic conditions which would be applied across the board. But how are we to know which contexts are meaningful?

Certainly we would not want to reject all disjunctive categories. One might think of a perfectly acceptable category like ‘resident of Canada or resident of the United States’. What, then, would be the basis for rejecting some disjunctive categories over others? Something like ‘naturalness’ might be such a basis: we find it natural to think of ‘residency in the United States’ and

[30] [Benacerraf 1965, 65 note 12].

[31] A purely postulational view of mathematics, like that of Hausdorff, would presumably have no trouble with this sort of view. But such a view, of course, would have little trouble with any questions pertaining to the uniqueness of mathematical entities. Its difficulties lie elsewhere.

'residency in Canada' as falling under the more general notion of 'residency'—after all, we already have a word for this category and could likely tell a convincing story about why and how this concept emerged in the first place. But perhaps a category under which Roman generals and natural numbers both fall is an unnatural one. This, of course, simply pushes the question back further by requiring some account of naturalness which, it seems to me, would be equally difficult to produce.

Benacerraf's use of the phrase 'well-entrenched predicate' might very well indicate that his definition of 'uniform conditions' (and, relatedly, of 'naturalness') would be largely historical in character: we find it natural to think of categories like 'mammal', 'residency', 'cat', or 'temperature' but unnatural to think of 'general or number', 'hairpiece or firing range', etc.[32] This may be the case simply because we have found it prudent over the course of our history to think and speak in terms of various separate entities as mammals, residents, cats, or temperatures but have never been called upon or found it helpful to think or speak about hairpieces and firing ranges at one time. On this suggested view, the uniform conditions for determining which predicates and categories are acceptable or natural or well-entrenched would be given relative to a personal, historical, or cultural trajectory. Perhaps, then, the category which would allow us to decide whether '2 = Julius Caesar' is true or false is one which has not yet become (or will never become) relevant enough to become well-entrenched. Though at this point, after a century of arguing about Caesar's numerical status, it seems likely that we have dug a deep enough trench. Benacerraf, for his part, doesn't elaborate on this aspect of his view here, so we can only offer conjecture as to how we might capture just the categories he has in mind. Clearly, though, there will be significant difficulties in such a project.

### 1.2.2 Empty generality

As we've seen, Frege's own solution to this problem (or his version of it) was to restrict meaningful statements to those whose terms referred to the elements of his ontological hierarchy and which obeyed the combinatoric logic of that hierarchy. This required him to employ very general notions like 'object' and 'function'. Benacerraf finds the extremely general notions at the basis of Frege's ontological hierarchy so broad as to be useless—for him they simply do not form categories capable of providing the context for individuation between such radically disparate entities as, e.g., the number 27 and a particular radish. Such notions are not *ad hoc* in the way that custom-built predicates like 'is a number or a general' are, and they seem to be rather 'well-entrenched' on any charitable account of that phrase. So what, then, is the problem with these notions? Benacerraf is fairly brief in this objection, writing that

> To say that they are both "entities" is to make no presuppositions at all—for everything purports to be at least that. "Entity," "thing," "object" are words having a role in the language; they are place fillers whose function is analogous to that of pronouns (and, in more formalized contexts, to variables of quantification).[33]

---

[32] In some ways, the requirement of historical naturalness would mirror Hilbert's suggestion (cf. [Corry 2004a, 20]) that we ought only to axiomatize those parts of mathematics which have suitable historical pedigree.

[33] [Benacerraf 1965, 65-66]. Interestingly, a key component of Benacerraf's own structuralist view of the natural numbers is that they are *not* objects.

Benacerraf suggests here that we cannot employ general notions like entity and object as appropriate categories enabling us to determine the truth value of identity claims because they are, in a sense, without real content. It is not that they are not well-entrenched notions—surely few notions have the historical pedigree of 'object' and 'entity'—but, instead, that they offer no meaningful contrast through which a clear meaning can be assigned to an identity claim. They are 'presuppositionless' insofar as we must already speak of entities or things if we are to speak logically about anything whatsoever. Part of Frege's problem, on this view, is that his ontological hierarchy was too general to capture Benacerraf's more restricted notion of semanticality. Indeed, for Benacerraf, it is precisely this generality which leads us into absurdities like dividing by Caesar.

But here, too, we seem to run into significant difficulties with Benacerraf's attempt to side-step Frege's point of view and its related troubles.

Is it true that there is no real content to notions like entity, object, and thing? In the brief passage just quoted, it seems that Benacerraf equates the roles of the words 'entity' and 'object,' while also holding that everything purports to be an entity. If the equation is to hold, then, whatever we mean when we use the word 'entity' is what we mean when we use the word 'object'. On this view, neither of these terms is terribly useful because they don't pick out any particular category. Anything at all is an object or an entity or a thing. Indeed, the use of the term 'anything' in the previous sentence seems to presuppose this fact. So, if we are trying to develop a category which delimits the semantical from the unsemantical, these notions will be powerless in their generality. They simply do not provide a meaningful contrast.

But is this in fact the case? There are, I think, reasonable grounds for thinking otherwise. For one, it does seem that many mathematicians and philosophers assume some particular content for the notion of 'object', at the very least, and probably 'entity' as well. Even without the technical ontological notion of object possessed by Frege, it seems clear that most mathematicians (excluding, perhaps, set theorists)[34] would differentiate between the objects of a given system and, say, the relations between the objects of that same system—this is particularly true of most branches of abstract algebra. On what basis is this distinction made if not on some presupposition about the nature of object-hood?

Benacerraf's response to this query would likely be: indeed the category- or number-theorist *does* make a distinction between objects and relations, but this distinction is only made on the basis of a narrower, theory-specific notion of object-hood. On this view, orthographic devices could be developed to help maintain logical perspicuity (i.e., differentiation between objects of various types by use of subscripts: '$\text{object}_{\text{category-theoretic}}$,' '$\text{object}_{\text{set-theoretic}}$,' '$\text{object}_{\text{Frege}}$,' etc.). But

[34] The fact that ZF set theory a) was widely accepted as the best candidate for a foundation for mathematics when Benacerraf's paper was written and b) is purely extensional, containing no real differentiation between, e.g., objects and relations or concepts, might help to explain the origin of Benacerraf's view here. Category theory, by contrast, operates on a basic distinction between the objects and the arrows (or morphisms) of a given category. We might also compare various axiomatizations of set theory to give some sense of the content of the notion of object; while ZF set theory deals exclusively with sets, [Von Neumann 1925] contains an interesting axiomatized 'set' theory in which functions are logically prior to sets.

to attribute the content of the notion of object within various branches of mathematics exclusively to narrower theory-dependent decisions is a move which seems to me to avoid the fact that, *prima facie*, objects seem to play the same sort of role in many different mathematical domains. Indeed, Frege's distinction between objects and functions was meant to capture exactly this general feature of mathematical thinking. Without further argumentation, there is no obvious reason to reject the reasonable view that what can count as an object for a particular theory in the first place is dictated by a more general notion of object-hood, a notion which is often left tacit.

### 1.2.3 The content of logic

There also seems to be a conflict here between Benacerraf's claim that these notions are presuppositionless and other aspects of his own view. Note that Benacerraf shares with Frege and many others the belief that logic is a science of extraordinary generality, one which is applicable anywhere. Even with his demand that meaningful identity statements require a well-defined category to gain purchase, he writes that

> Logic can then still be seen as the most general of disciplines, applicable in the same way to and within any given theory. It remains the tool applicable to all disciplines and theories, the difference being only that it is left to the discipline or theory to determine what shall count as an "object" or "individual."[35]

Now, we've seen in previous chapters why Frege was able to hold that logic was pan-theoretically applicable. For him, it was the most general science of truth. But this did not at all mean that logic itself was in any way presuppositionless. One might even reasonably say that a large part of Frege's life's work was to bring the tacit presuppositions of logic into the light of day. This generality rested on logic's capacity to deal directly with those entities which were capable of being true or false, i.e., Fregean thoughts. Crucial to Frege's understanding of thoughts was his ontological hierarchy of objects and functions. Though Frege was quite aware that he could not define 'function' and 'object' in a purely precise manner, he nevertheless attempted to provide elucidations which were meant to guide the reader toward the content of these basic notions.[36] The reason, then, why Frege believed logic to be universally applicable to any endeavour concerned with truth was because it dealt with the most general laws of truth, and truth itself is (partially) understood in terms of the ontological connections between entities in Frege's hierarchy. So, while Frege might have been willing to allow for narrower, theory-specific notions of 'object', 'entity,' or 'individual' he would likely have suggested that if these notions are going to be scientifically useful at all (i.e., if they are going to be employed in the search for truths), then they must already have built into them the more general assumptions of logic. For an individual theory to treat some class of entity as an object would, for Frege, already involve the assumption that such entities were self-standing, could act as arguments for first-level functions, etc. While others might not share Frege's own assumptions about the nature of

---

[35] [Benacerraf 1965, 66]. This statement seems a particularly difficult one to maintain without argumentation in 1965, well after the advent of intuitionism in the work of Brouwer and Heyting. For an interesting argument that at least some intuitionistic mathematics simply cannot share the same internal logic as classical mathematics, see [Shapiro 1989].

[36] Indeed, one of the three main guidelines Frege gives himself in the *Grundlagen* is "never to lose sight of the distinction between concept and object" [Frege 1884b, xxii].

objects or logic, it does seem clear that when the category or number theorist suggests that a given entity is an object rather than a function, they are tacitly employing a general notion of object which is not strictly determined by the definitions of category or number theory.

One very basic assumption made by most forms of contemporary logic, for example, might be that the world is (at least partially) structured into entities at all, i.e., things which are particular, individuatable, and capable of possessing properties, et cetera. Even if this assumption turns out to be true—assuming that we can think of truth without already employing these notions—it is certainly not the case that everyone shares or has always shared this view. In philosophy and theology, for example, other ways of structuring the world have been developed extensively.[37] So, while the presuppositions implied by accepting a 'world of objects' (or, using Rudolf Carnap's phrase, a 'thing language')[38] are extremely general in character, they are indeed presuppositions, and ones which not everyone might share. Certainly they seem necessary for many forms of contemporary scientific thinking, but this, it seems to me, is just another way of saying 'We must employ the notions of logic to think logically'. The language of logic and science is certainly not the only one possible.

1.2.4 Which logic?

But part of Benacerraf's goal in this paper is to distance himself from the rigidity of Frege's ontological hierarchy and its concomitant problems with fixed reference. So it would seem that his account of the pan-theoretic applicability of logic must resist this answer. So what is its basis for Benacerraf?

I think part of the problem in answering this question for Benacerraf is that it is rather difficult to establish what is meant by 'logic' here. What exactly is it that we are applying, what is this 'most general of disciplines'? Again, accepting Frege's detailed answer to this question would commit us to exactly the sort of profound difficulties with reference that Benacerraf is trying to avoid. But, while Frege's position is off-limits, it doesn't seem that there is a clearly superior alternative.

Even more, while a few modern philosophers have argued in favour of the primacy of one form of logic over another—Frege and Quine being notable examples[39]—this has become an increasingly untenable position. The multiplicity of mutually incompatible logics developed throughout the 20th and 21st centuries is fairly good evidence that a conception of logic as the elaboration of a single intuitive notion of logical consequence might be outmoded or at least

---

[37] A fairly obvious example from twentieth-century philosophy can be found in the works of Martin Heidegger (1889-1976). In "The Question Concerning Technology" (i.e., [Heidegger 1977]), for example, Heidegger suggests that the division of the world into objects occurs in the course of historical time, that the world need not be divided into objects, and that this division in fact obscures crucially important features of the world. A more ancient, though related rejection of the division of the world into objects occurs in the mysticism of the Neo-Platonist Plotinus (c.204/205 – 270), who writes "Our way takes us beyond knowing; there may be no wandering from unity; knowing and knowable must all be left aside, every object of thought, even the highest, we must pass by, for all that is good is later than This and derives from This as from the sun all the light of the day" [Plotinus 2007, VI.9.4].

[38] [Carnap 1950].

[39] See, for example [Quine 1970], particularly chapter 6.

incapable of capturing much of what most of us would now happily call 'logic'.[40] Examined in light of these developments, the history of logic from Frege to the present is similar in many respects to the confusion in geometry after the emergence of non-Euclidean geometry in the 19th century which we examined in the first chapter. We have moved from the informal investigation of a single consequence relation, to tinkering with its formal definition, to the realization that there are many (perhaps infinitely many) such relations. Instead of a single logic, we now have a huge variety of logics, all describing different consequence relations and with different possible applications. So, just as we have moved from geometry to geometries, we have (with less fanfare, perhaps) moved from logic to logics.[41] Given the multiplicity of logics, then, the monolithic claim that logic is always and everywhere applicable must be slightly misleading or at least imprecise. Just as physical space might be describable in terms of a variety of quite different geometries, so too we can imagine a particular theory exhibiting many different logical structures, each with different consequence relations and (perhaps too) different assumptions about the nature of objects.[42] Given these considerations, then, it seems that the uncritical acceptance of the pan-theoretic applicability of logic obscures the use of particular logical assumptions behind a veil of contentless generality which no seems longer tenable.

### 1.2.5 Mathematical truth

Another fairly serious difficulty arising from Benacerraf's attempt to restrict the meaningfulness of identity statements to specific contents lies in the area of mathematical truth. Assuming we have somehow overcome the foregoing difficulties in determining, in a uniform way, which contexts are semantically acceptable and which are not, it will become immediately obvious, given such a context, which identity statements are true, which are false, and which are unsemantical. It will be clearly true, for example, that within the context of ZF and the Zermelo reduction (assuming these can be made semantically acceptable), the statement '$2 = \{\emptyset, \{\emptyset\}\}$' is false. By contrast, if our context is the von Neumann account of the natural numbers, then this statement will be true. By changing the context we change the category which determines the content of the identity relation and the truth value of the relevant statements. For Benacerraf, there simply is no theory-independent account of the meaning or truth of these sorts of identity statements.

As we saw in the second chapter of this thesis, Frege's emendations of Hilbert's view were similar in spirit to Benacerraf's demand for context, though they were intended to avoid precisely the sort of relativization of truth Benacerraf seems to suggest here. Frege's idea was to

[40] Note that this conception of logic as the elaboration of an intuitive notion of consequence, while common, is not shared by all philosophers. As Thomas Ricketts makes clear (in both [Ricketts 1997] and [Ricketts 2010]), it is rather difficult to characterize Frege in this manner, for instance. (For a contrasting opinion, see [Dummett 1973, 81], which suggests that a consequence-centered view of logic is implicit even in Frege's work).

[41] The *locus classicus* for logical pluralism is the work Rudolf Carnap (cf., [Carnap 1950] for one of the more provocative formulations of this view). More recently there has been renewed interest in the profound philosophical implications of the multiplicity of logics (see, e.g., [Beall and Restall 2000] and [Beall and Restall 2006]).

[42] Indeed, some very interesting work has been done on the possibility of different sorts of objects within mathematics and logic. An early example can be found in the works of the Austrian philosopher and psychologist Alexius Meinong (1853-1920) (e.g., [Meinong 1904]). In more recent years, a number of philosophers of mathematics have entertained the possibility of ontologically 'strange' objects, (e.g., [Akiba 2000] and [Balaguer 1998]).

make explicit (what he thought to be) the implicitly conditional nature of Hilbert's uninterpreted statements. Thus, in place of stating that a claim like '2 = {{Ø}}' is true only within or relative to the Zermelo reduction of the natural numbers to the universe of ZF, we instead say something like: "If the Zermelo reduction is true, and if the ZF axioms are true, then 2 = {{Ø}}". But, for Frege's conditionalization approach to work, some rigidly fixed ontological content has to be given to the background theory (in this case ZF and the Zermelo reduction)—something which Benacerraf is keen to avoid.

Thus, we are again presented with a dilemma. On the one hand we can have fixed reference and a non-relative account of truth. With this choice, however, we also run into the problems associated with absolutely fixed reference (e.g., the Caesar problem and its modern cousin the non-uniqueness problem). On the other hand, we can relativize reference to a particular theoretical framework. This leads us again to two formidable problems. First: how are we to understand the relation of scientific truth to the world if that truth is relative to particular frameworks which are not always mutually compatible? Second: do these frameworks themselves have any connection to ontology or must they be free-floating in Hilbert's sense, mere scaffoldings of concepts? If the latter is the case, then we are back at the problems of meaninglessness faced by Hilbert in his correspondence with Frege. If these frameworks *do* have ontological implications, however, we need something more than Hilbert gave us.

## 1.3 Structures

For Benacerraf, as for Hilbert,[43] the attempt to isolate some fixed collection of particular objects which are meant to serve as the natural numbers (or, in Hilbert's case, the elements of Euclidean geometry) is a fool's errand based on a serious philosophical misunderstanding of the nature of mathematical structures. This misunderstanding hinges, for Benacerraf, on the view that mathematics deals with various collections of objects (numbers, sets, points, groups, etc.) which can be picked out and which bear properties in the same way that objects like tables and chairs do. Benacerraf's view entails that mathematics isn't really about such objects at all, or, slightly more cautiously, that it only deals with isolable objects insofar as they are related to a given structure. The mathematically relevant aspect of an object is not one which can be discussed with indifference to the particular structure of which it is a part. This is a large part of the reason why questions like 'Is Pompey Magnus identical to the number 2?' just don't make any sense. On this structure-focussed view, the typical set-theoretic reductionist's claim that the natural numbers just are sets is a mistaken one.

In order to make all of this a bit more precise, Benacerraf makes a distinction between the ontological activity of (unrestricted) identification ('*x* is *y*') and the ontologically neutral activities of reduction and explication ('within context *C*, *x* can be treated as *y*'). If we are to identify the natural numbers with some collection of sets, in the strong ontological sense, we mean quite literally that, when we speak about the natural numbers, we are *really referring to*

[43] Recall Hilbert's view that asking for referentially fixed terms prior to the application or interpretation of an axiomatic theory involves " One […] looking for something that can never be found because there is nothing there, and everything gets lost, becomes confused and vague, and degenerates into a game of hide-and-seek" [Hilbert to Frege, in Frege 1971, 11-12].

this particular collection of sets. Thus, if the set we identify with the number 3 has 847 members, then the number 3 has 847 members, regardless of the context within which we are speaking. But, as the non-uniqueness problem is meant to point out, we have been unable to provide any good reason for choosing one such reduction over another, as all of the distinguishing properties of individual reductions (set-theoretic or otherwise) will be arithmetically irrelevant. And, if the failure of Frege's best efforts is any indication, we are not likely to stumble on such reasons in the future.

The activities of reduction and explication, by contrast, carry no such ontological commitment according to Benacerraf. If on a given reduction of the natural numbers to a collection of sets the number 3 has 847 members, we are not committed to the view that every single time we speak about the number 3 we are speaking about this particular 847-membered object. Rather, we are meant to understand that the number 3 is identified with an 847-membered set *on this reduction* and might very well be reduced to or explicated by a set possessing completely different non-arithmetic properties on another reduction, or even with an object that isn't a set at all on some other reduction. Identification, in the strong sense, is, for Benacerraf, theory-independent, while reduction and explication are only possible within the context of some particular theory.

Keeping these distinctions in mind, Benacerraf tries to overcome the dangers of Hilbert's relational view of structures by giving a more robust account of how abstract structures differ from the object-oriented accounts like those of Frege and the set-theoretic reductionists.

The most obvious difference between Benacerraf's view and Frege's is that Benacerraf explicitly denies Frege's central claim that numbers are objects. Indeed, the central lesson that Benacerraf draws from the non-uniqueness problem is that, contra Frege, numbers simply cannot be objects (as the title of the paper is meant to suggest). The advantage he gains is fairly obvious at this point: he no longer has to address the Julius Caesar problem and others like it. By rejecting the broad claim that numbers are objects, Benacerraf can now look at "the question of whether a particular 'object' […] would do as a replacement for the number 3" as "pointless in the extreme".[44]

In place of the object-oriented picture of (at least) the natural numbers, Benacerraf suggests a structure-oriented view. The explanatory burden, then, is shifted. He no longer needs to explain how we pick out which particular objects the natural numbers are, but he must explain what the natural number structure is, and how its elements (the numbers) differ from objects. In Frege's attempted emendations of Hilbert's free-floating view, he sought to characterize Hilbert's uninterpreted formal systems by pinning them down within his ontological hierarchy of objects and functions. For Frege, then, the direction of best mathematical explanation was from the concrete world of objects and functions to the world of abstract formal structures built out of these entities. Benacerraf's central idea is to reverse this order of explanation, and suggest that mathematics is in fact about abstract structures, and that it only secondarily studies their concrete representations (e.g., representations within Frege's ontological hierarchy, or more familiar representations within the universe of set theory).

[44] [Benacerraf 1965, 69].

Frege's objection, when Hilbert tried to make a similar explanatory shift, was that this seemed to leave mathematics contentless. In order to address this objection, and therefore to move beyond the Frege-Hilbert impasse, Benacerraf needs a more robust account of the relationship between ontology and the abstract structures which he supposes are the subject matter of mathematics.

### 1.3.1 Concrete vs. Abstract Structures

So, what exactly does Benacerraf mean when he speaks of structures?

Given the tentative character of the paper's final few pages, it is perhaps not surprising that Benacerraf does not give us a thoroughgoing definition of the term 'structure'. Nevertheless, we can surmise that, for Benacerraf, at the very least a structure is composed of elements arranged under a certain relation (or relations).[45]

So far as this goes, this account of structure is compatible with Frege's views or those of the set-theoretic reductionist. On Frege's account, both the elements of and the relations which define a structure are entities with a fixed position in his ontological hierarchy, similarly for the set-theorist and the universe of set theory, where everything (relations included) is a set. But Benacerraf wants to make a sharp distinction between concrete structures and abstract structures which will allow more careful differentiation between the 'structuralism' which he wishes to espouse and the referentially-fixed reductionism of Frege and the set theorists. Let's look first at what this distinction says about the elements of each sort of structure.

A concrete structure's elements will be entities[46] which possess properties or stand in relations which are not solely determined simply by their being elements of that structure. For Benacerraf,

> That a system of *objects* exhibits the structure of the integers implies that the elements of that system have some properties not dependent on structure. It must be possible to individuate those objects independently of the role they play in that structure.[47]

---

[45] Cf. [Benacerraf 1965, 69-73], where he speaks repeatedly of the elements of the natural number structure, and also of structures as systems of relations. Some later structuralists, e.g., Stewart Shapiro, countenance structures which are composed of elements but no relations (cf., [Shapiro 1997, 115] and [Shapiro 2008] which discuss cardinal structures which have elements but no relations). Even further, Shapiro considers the degenerate case of the "0 pattern" which has neither elements nor relations [Shapiro 1997, 115].

[46] Benacerraf seems to assume in this paper that the elements of a concrete structure will be objects, though it is not exactly clear what he means by 'object' or even, as noted above, whether he thinks the term is a meaningful one at all outside of a well-defined theoretical context.

[47] [Benacerraf 1965, 70; emphasis mine]. Here Benacerraf seems to tacitly view objects as independently existing entities. It should be noted here that a number of subsequent authors have suggested that there may indeed be objects (not merely 'elements' or 'roles') whose strange nature allows them to be elements of *abstract* structures in Benacerraf's sense. See, for instance, [Akiba 2000], which defends the position "that mathematical objects, along with some others, neither definitely possess nor definitely lack certain properties" [Akiba 2000, 26]; [McLarty 1993] suggests that "the structuralist program [of Benacerraf] is already fulfilled or obviated, depending on how you look at it, by categorical set theory […]. Sets and functions in this theory have only structural properties. There is no need and no further place for a further theory of abstract structures" [McLarty 1993, 487]. Depending upon how we read the situation, these objections may amount to quibbles regarding the meaning of the term 'object' as Benacerraf uses it.

Thus, for instance, even if Julius Caesar is to be an element of the natural number structure[48] (i.e., if he is arranged with other elements under an appropriate relation of recursion) he will have many properties which are not determined simply by his being an element of this structure. Simply being an element of such a structure will not tell us anything about the content of his *Commentarii de Bello Gallico* or which sort of olive he may have preferred, for instance. These latter properties are structure-independent (in relation to the natural number structure, at least). But, if we are to hold that he is an element of the natural number structure, then he will also have structure-dependent properties, like being prime or being the square root of 4 and so on. So, the elements of concrete structures will have both structure-dependent and structure-independent properties.

It is specifically the possibility of extra-structural individuation (of the kind which might have solved the Caesar problem) and the possession of structure-independent properties which characterize concrete structures. The structure-independence of their elements is also the feature which, according to Benacerraf, disqualifies them as the main focus of mathematical interest. For Benacerraf, the chief reason why mathematicians and philosophers have gotten into muddles like the Caesar problem or the non-uniqueness problem is because they mistake concrete structures for abstract ones, and they ignore the fact that what they are actually most interested in is abstract structures. The trouble is in understanding exactly what these abstract structures are—or, more pointedly, in distinguishing abstract structures from empty formal games.

Benacerraf's abstract structures, like their concrete cousins, are composed of elements arranged under a relation (or relations). In contrast to concrete structures, however, the elements of abstract structures do not possess any structure-independent properties. Given what appears to be Benacerraf's tacitly Fregean notion of object, this means that the elements of abstract structures cannot be objects at all.

What, then, are they? There are at least two ways to read this question.

On the first, it asks us: what sorts of entities (other than objects) are abstract structures composed of? On Benacerraf's view, reading the question this way would simply be to recapitulate the object-oriented misunderstanding of mathematics. His point is that looking for the building blocks of structures, those isolable individuals which we might put together and eventually obtain a structure, is only ever going to get us concrete structures.

There is, however, another way of reading this question which might help to shed some light on Benacerraf's view. Instead of asking 'out of which sorts of things are structures composed' we can ask 'what is the relationship between an abstract structure and its elements'? And to this question we might give a very different answer. Here is how Benacerraf puts it:

> So what matters, really, is not any condition on the *objects* […] but rather a condition on the relation under which they form a progression. To put the point differently—and this is

[48] As we will see in what follows, for Benacerraf an object like Julius Caesar cannot be a member of *the* natural number structure (which is, for Benacerraf, an abstract structure) but only of a concrete structure which represents the natural number structure.

> the crux of the matter—that any recursive sequence would do suggests that what is important is not the individuality of each element but the structure which they jointly exhibit. This is an extremely striking feature. One would be led to expect from this fact alone that the question of whether a particular "object" […] would do as a replacement for the number 3 would be pointless in the extreme, as indeed it is. "Objects" do not do the job of numbers singly; the whole system performs the job or nothing does. I therefore argue, extending the argument that led to the conclusion that numbers could not be sets, that numbers could not be objects at all; for there is no more reason to identify any individual number with any one particular object than with any other (not already known to be a member).[49]

The relationship between the elements (which, again, are not objects for Benacerraf) of an abstract structure cannot be made sense of on a case by case basis. An entity is an element of a particular structure not because it possesses unique properties which are ideally suited to playing a peculiar role. No—Benacerraf's point is that while "*Any* object can *play the role of* 3" no object *is* the role of 3.[50] And it is the *role* played by 3 which is the element of the abstract structure, or, perhaps, the third *place* in the abstract natural number structure. Certainly this role can be represented by a particular object, or this place occupied by a particular object—indeed, any particular object, according to Benacerraf. But such representation can only occur within a *concrete* structure, i.e., a structure whose elements will have structure-independent properties like those of the set-theoretic representation of the natural numbers.

So, for Benacerraf, the number 3 is not an object with a kind of structure-independent reality (as Frege suggested), out of which a structure is built, but is instead a role which can only be defined relative to the abstract structure of which it is an element. He writes that "What is peculiar to 3 is that it defines that role – not by being a paradigm of any object which plays it, but by representing the relation that any third member of a progression bears to the rest of the progression."[51] On this view, the elements of a structure are not prior to the structure. What is essential is the abstract structure itself, and this is defined by a particular relation or relations on elements. The relations which these elements stand in, and the properties which they might possess, are only those determined by the simple fact of their being elements of the structure. Thus, their *nature* is structure-dependent, or as Stewart Shapiro puts it, "the essence of a natural number is its *relations* to other natural numbers."[52]

This view, suggestive as it is, poses a number of difficult questions related to our central aim of clarifying what mathematics is about. If mathematics is chiefly concerned with these abstract structures,[53] we'll want to know how (and perhaps which) relations define or otherwise relate to

[49] [Benacerraf 1965, 69].
[50] [Benacerraf 1965, 70].
[51] [Benacerraf 1965, 70].
[52] [Shapiro 1997, 72].
[53] Benacerraf himself is fairly limited in the scope of his claims in [Benacerraf 1965], sticking for the most part to a description of the arithmetic of the natural numbers. As we will see in the following section, it is later thinkers who sought to extend Benacerraf's 'structuralist' suggestions about the natural numbers to the wider world of mathematics as a whole. Nevertheless, it does seem clear that, even for Benacerraf in 1965, his limited suggestions

abstract structures. We'll also want to know what ontological status—if any—such structures are meant to possess. In our pursuit of these questions, it is best to depart from the specificities of Benacerraf's account and begin an examination of more recent structuralist literature in the philosophy of mathematics. As Benacerraf himself notes, the positive conclusions of his paper are exploratory in character. He does not pretend to develop a fully fleshed-out account of the nature of abstract structures, but, instead, draws our attention away from the fool's errand of naïve set-theoretic reductionism toward the structural meat of mathematics.

## §2. Contemporary structuralism

### 2.1 From sets to structures

Historically, Benacerraf's paper was part of the impetus for a substantial reorientation in the philosophy of mathematics. As I've mentioned at the beginning of this chapter, the philosophy of mathematics in the decades preceding Benacerraf's paper was dominated (with few exceptions) by the problems of set-theoretic reductionism. While set theory still figures prominently in the philosophy of mathematics even today, Benacerraf's paper helped to popularize a structuralist approach to all forms of reductionism, even set-theoretic ones. By the mid-1990s, to be sure, the philosophy of mathematics was dominated with questions concerning the viability of structuralism as a solution to some of the important foundational questions in mathematics.

This philosophical movement toward structuralism was preceded by an increased focus on structure within mathematics itself. This focus is evident already in the work of Hilbert and others around the turn of the nineteenth century, but became increasingly more explicit and well-defined in the work of the group of French mathematicians calling themselves Nicholas Bourbaki and in the (related) rise of abstract algebra from the 1930s onward.[54] Bourbaki wrote and published an incredibly influential series of texts called the *Éléments de Mathematique*. The series involved the hierarchical and rigorous axiomatic development of mathematics from the ground up, ostensibly under the notion of structure. While Bourbaki's technical definition of the term 'structure' was not terribly influential,[55] their axiomatic-structural approach was immensely so. As usual, much of the philosophy of mathematics had lagged behind the actual development of mathematics for a few decades by the time Benacerraf's paper appeared. But, as is also often the case, very few mathematicians had a deep enough interest in the philosophical aspects of their discipline to develop a thorough-going account of these new structural developments and the nature of contemporary mathematics.[56]

---

implicitly contain much broader implications—for very little in the paper hinges on the peculiarities of the natural numbers or their set-theoretic representations.

[54] [Corry 2004b] is the finest available historical account of these developments; [Wussing 1984] is an illuminating account of the 'pre-history' of modern abstract algebra as found in the development of group theory; [Dieudonné 1972] presents an insider's look at the somewhat more specific field of algebraic geometry.

[55] Their technical definition is first advanced in [Bourbaki 1939]. Its lack of popularity is evident in the fact that it is rarely employed, neither by the members of Bourbaki nor those they influenced. In a survey of the notion of structure in mathematics, Mac Lane writes that "as best I can determine [Bourbaki] never really made actual use of his definition, and I will not make any use here of my variant" [Mac Lane 1996, 179].

[56] This is not to say that mathematicians involved in these developments were disinterested in the philosophy of mathematics, only that there were (and continue to be) very few mathematicians whose interest extends beyond the publication of one or two brief papers on the subject; Mac Lane is a notable exception (cf. [Mac Lane 1986]).

## 2.11 Benacerraf's legacy

In the 1965 paper, Benacerraf confines himself for the most part to a discussion of the arithmetic of the natural numbers. Since this paper, a more general position in the philosophy of mathematics called 'structuralism' has emerged which extends these claims to include a much larger portion of mathematics. As Benacerraf himself puts it in a later paper, "the structuralist's *cri de guerre*" is "that mathematics is the study of (abstract) structures".[57] My aim here is not to give a complete account of contemporary forms of structuralism which follow in the wake of Benacerraf's work, but to see whether they are capable of addressing the problems present in Benacerraf's 1965 account.

As we saw, Benacerraf attempted to develop structuralism as a means of addressing both the Julius Caesar problem and its more contemporary manifestation, the non-uniqueness problem. His way of doing so involved denying Frege's claim that all identity statements must be meaningful. In effect, he sought to restrict meaningful identity claims to well-defined contexts. One way of reading his view is that, in mathematics at least, abstract structures serve to furnish these well-defined contexts. Thus, when we speak of identity, we speak of identity between the entities of a given structure, say the natural number structure. The benefit of this view is that we can eliminate the Julius Caesar problem by suggesting that, without further contextualization, there is no obvious structure which contains both Caesar and the natural number 2, so the question 'Is Caesar identical to the number 2?' is rendered meaningless. When such a context is provided (say in a structure where components of the physical universe are employed as representatives of the elements of the natural number structure) the statement gains a truth-value.

But there are clear difficulties with his attempt.

First, Benacerraf has what appear to be inconsistent views of the role that objects play (or do not play) in this whole story. He famously denies that the natural numbers are objects – and, by extension, that the elements of any abstract mathematical structure are objects. He does so, it seems, because he takes 'object' to mean 'theory- or structure-independent entity', which is exactly what the natural numbers are *not* on his account. But he also maintains, in the same paper, that we can speak of theory-dependent objects *and* that the notion of 'object' is too general to have any content whatsoever. Intermingled with these concerns are questions about the exact nature of the elements of abstract structures. As quite a bit seems to rely on how we understand his claims about objects and elements, these equivocations need to be addressed.

Second, Benacerraf does not give us a thorough-going account of the exact nature of the abstract structures which he believes are the subject matter of mathematics. This is not terribly surprising, as his paper is intended to be (for the most part) diagnostic in character. If we are to accept his claim that arithmetic and mathematics concern these abstract structures, though, we need to develop a more precise understanding of these entities. Particularly pressing, from our point of view, are questions pertaining to the ontological status of such structures: do they exist or not? If they exist, are they to be understood as permanent, independent features of reality, or are they,

[57] [Benacerraf 1996a, 184].

like their elements, dependent upon a particular context or theory for their existence? If they do not exist, does this leave us to treat mathematics as a contentless enterprise?

Third, a great deal in Benacerraf's account seems to hinge on the status of logic within mathematics. Recall that a chief aim of his paper is to undermine the hold that set-theoretic reductionism has on the philosophy of mathematics. He does this by showing that what we are really interested in is *not* the ontological universe of set theory, but instead the abstract structures, most or all of which can be modelled or exemplified in this universe. His paper is meant to show that our representations of these structures within a single ontological universe are of secondary concern. But he also claims that logic (singular) is pan-theoretically applicable, and that it will operate the same in relation to any given abstract structure we might wish to study. The question we have to ask, especially given the almost universal understanding of logic in terms of the background ontology of model and set theory, is whether or not this ends up being circular. Does Benacerraf's view of the universality of logic commit him to the background ontology of set theory that the paper seeks to avoid?

In order to further address these concerns let's turn to an examination of some more contemporary forms of structuralism.

## 2.2 Two types of structuralism

There are quite a number of distinct philosophical positions which can be placed under the heading 'mathematical structuralism'. Indeed, there are even a number of quite distinct *typologies* of such views.[58] In what follows, I would like to provide a very general typology of my own, one which contains precisely two forms of structuralism. These I will call (loosely following a similar division made by Shapiro and Hellman)[59] 'assertoric structuralism' and 'algebraic structuralism'. Very schematically, assertoric forms of structuralism hold that mathematics contains assertions about a particular content. Usually, but not always, this content is fleshed out in terms of some theory of abstract structures. Algebraic structuralism, by contrast, contends that mathematics does not have a single well-defined content, and that, to the extent that it contains assertions at all, these are always relative to a particular theory, framework, language, or structure. In the first, assertoric, case we might speak of a global content for mathematics, while in the latter, algebraic, case if mathematics is said to have content at all, it is only in a local sense. My intention is to investigate whether either of these forms of structuralism is capable of satisfactorily addressing the problems with Benacerraf's view and, thus, helping us to explain what it is that mathematics is about.

---

[58] [Hellman 1996] presents a typology containing four versions of structuralism (the model- or set-theoretic account, the category-theoretic account, the *sui generis* account, and his own modal account). Five years later, in [Hellman 2001], the typology is reduced to three versions (set-theoretic, *ante rem*, modal). There are others besides (e.g., [Dummett 1991, chapter 23], [Horstein 2014]).

[59] Cf., [Shapiro 2005]. For Shapiro, the division concerns the status of sentences within a given account – an account is assertory if its sentences "are meant to express propositions with fixed truth values. Algebraic sentences are schematic, applying to any system of objects that meets certain conditions" [Shapiro 2005, 66]. While Shapiro's division is my inspiration here, my own may diverge from his in what follows.

Both of these forms of structuralism insist that mathematics is chiefly concerned with the study of abstract structures. They differ, however, in their understanding of the ontological status of these structures. For the assertoric structuralist, there is (or ought to be) what we might call a ‘final theory’ of structures, one in which the intricate network of relationships between structures (e.g., the sort of hierarchy of algebraic structures we find in Bourbaki’s *Éléments*) is ultimately cashed out in terms of a theory which makes assertions about the world directly. The difference between this form of structuralism and the more straightforward, naïve version of set-theoretic reductionism, is that the assertoric structuralist is always concerned with *abstract* structures, in Benacerraf’s sense, whereas the set-theoretic reductionist (or Frege) thinks in terms of a single concrete structure (the set-theoretic universe *V*, Frege’s ontological hierarchy, etc.).

The category ‘assertoric structuralism’ is a very broad one, and it includes the work of structuralists who are, in many other respects, opposed. Thus, the (putatively) nominalist structuralism of Geoffrey Hellman[60] and the realist structuralism of Stewart Shapiro[61] both count as forms of assertoric structuralism. They differ markedly in what exactly they assert, of course. For Shapiro, structures are *ante rem* universals—about as far as one can get from nominalism. For Hellman, structures are ‘eliminated’ “in favor of sentences with modal operators”.[62] These in turn are cashed out in terms of “a modal second-order logic with a restricted (extensional) comprehension scheme,”[63] with the end result being that Hellman makes assertions about possibilities, rather than actual universals. Aside from Shapiro and Hellman, there are a number of assertoric forms of structuralism which employ unique background ontologies to account for abstract structures for different reasons.[64] What ties them together is that they claim that mathematics is best understood as making assertions whose truth value is ‘absolute’ in the sense that there is a final theory of structures (or whatever) within which the relative truths of specific mathematical theories are embedded.

Algebraic structuralism is so named because of its connections (conceptual and historical) to abstract algebra. A key component of abstract algebra involves detailed work on relations between various well-defined algebraic structures (e.g., rings, groups, monoids, etc.). Prior to the development of category theory in the 1940s and 1950s, these algebraic structures were ultimately cashed out in the background ontology of set-theory (usually ZF). But, in many ways, algebraists found—and still find—this connection to set theory and its ontology an encumbrance. Algebra has, historically, been concerned with operations at least as much as it has been concerned with particular objects (in most cases, much more so). To try to ‘pin down’ algebraic work in the purely object-based ontology of set theory is, in many ways, foreign to the operation-focussed character of abstract algebra. If, for example, your interest is in discussing certain homomorphisms between two rings, doing everything in set theory is both time-consuming and, for the most part, beside the point. In set theory, both morphisms and structures (as well as the elements of structures) will be understood in terms of sets.[65] But, as many algebraists will note,

[60] Cf., e.g., [Hellman 1989], [Hellman 1996], and [Hellman 2001].
[61] Cf., e.g., [Shapiro 1997] and [Shapiro 2008].
[62] [Hellman 1996, 102].
[63] [Hellman 1996, 102].
[64] One influential such view is that of Michael Resnik, who has developed an account of structures based on the notion of patterns (cf. [Resnik 1980], [Resnik 1982], and [Resnik 1997]).
[65] Or, possibly, urelements.

morphisms behave nothing like sets—so why understand them in these terms in the first place? Since its inception, category theory has been treated by many algebraists as a more 'natural' mathematical language (or setting) for speaking about these structural interrelations. This is perhaps not surprising, given that the earliest incarnations of category theory, in collaborations between Saunders Mac Lane and Samuel Eilenberg, involved research into the surprising and profound morphisms between algebraic and topological structures, which were, at that time, pursued more or less independently.[66] The impetus of category theory, then, was to make work on morphisms between structures easier, more flexible. To this end, category theory—which *can* be modelled in the universe of sets,[67] but which need not be[68]—begins with a class of structures (i.e., categories) which are defined in terms of morphisms *and* objects.

The emergence of category theory in the 1940s and 1950s thus pointed to a new direction for abstract algebra, where the idea of set theory having the last word no longer seemed necessary or even plausible. As John Bell eloquently puts it,

> Because the practice of mathematics has, for the past century, been officially founded on set theory, the objects of a category […] are normally constructed as *sets* of a certain kind, synthesized, as it were, from pure discreteness. As sets, these objects manifest set-theoretic relationships—memberships, inclusions, etc. However, these relationships are irrelevant—and in many cases are actually *undetectable*—when the objects are considered as embodiments of a form, i.e., viewed through the lens of category theory. […] This fact constitutes one of the 'philosophical' reasons why certain category theorists have felt set theory to be an unsatisfactory basis on which to build category theory—and mathematics generally. For categorists, set theory provides a kind of ladder leading from pure discreteness to the category-theoretic depiction of the real mathematical landscape. Categorists are no different from artists in finding the landscape (or its depiction, at least) more interesting than the ladder, which should, following Wittgenstein's advice, be jettisoned after ascent.[69]

This is not to say that algebraists abandoned set theory, nor that category theory itself has not often been treated as a new 'last word' in cashing out claims about relations between structures. But, for at least some algebraists, and some philosophers of mathematics as well, the emergence of category theory as a means of describing morphisms between structures has illustrated not only that set theory is not the privileged resting place of all structures, but also that we do not need such a privileged resting place and that, anyway, *there can be none*.[70] This is the central claim of algebraic forms of structuralism: mathematical assertions can only be made relative to a particular background, and there can be only local reasons (reasons pertaining to one's particular problem, to a particular class of morphisms, etc.) for preferring one background or framework to another. Thus, for example, one might study the natural number structure in relation to the background ontology of set theory or category theory, or with indifference to both. The dream of

---

[66] Cf., [Mac Lane and Alexanderson 1989, 20] for a discussion of Mac Lane's initial collaboration with Eilenberg.
[67] Cf., e.g., [Mac Lane 1998], though here Mac Lane is keen to point out that categories need not rely on set theory.
[68] Cf., e.g., [Lawvere 1964], .
[69] [Bell 2001, 152].
[70] [Awodey 2004] contains one of the most convincing defenses of this position.

a single universe within which all mathematical claims operate is, to the algebraic structuralist, just that: a dream. It is neither followed in practice (especially not in most algebraic practice), nor does it make sense philosophically (as Benacerraf's non-uniqueness problem can be said to show, at least partially). To make the move of the assertoric structuralist is, for the algebraic structuralist, to slip back into the game of hide-and-seek that Hilbert already saw to be impossible at the turn of the nineteenth century.

In contrast to assertoric structuralism, the algebraic form of structuralism holds that there is not (and, often, there *could not possibly be*) any final background theory which rules over all of mathematics. For the algebraic structuralist, claims about relations between structures are themselves best thought of as embedded within a particular structure, or point of view. For different purposes, we might switch between a background theory of set theory or category theory or something else besides. The assertoric structuralist might suggest that switching between these theories implies that there is some sort of equivalence between the two, and, thus, that there is some broader, final abstract structure which encompasses both. The algebraic structuralist will hold that there may be such a structure, but that it, too, is subject to revision as a momentary background ontology, and not at all a final arbiter of what is. Though there are important differences, I hope the reader is struck by the similarity between the assertoric/algebraic divide and the debate between Frege and Hilbert. Frege, like the assertoric structuralist, sought to secure a determinate content for all of mathematics. In contrast to the assertoric structuralist, however, Frege's work endeavoured to embed mathematics within a single concrete structure. Similarly, the algebraic structuralist's unwillingness to pin mathematical claims down within a single ontological frame is strikingly similar to Hilbert's free-floating view. Neither the algebraic structuralist nor Hilbert believed that mathematical claims required extra-theoretical or extra-structural grounding. The question we must ask ourselves now is whether the divide between assertoric and algebraic structuralists holds more promise than the impasse of the Frege-Hilbert debate.

## 2.3 Assertoric Structuralism

Let's begin our investigation of contemporary structuralism by examining the capacity of assertoric structuralism to address the problems left by Benacerraf's account. In this section I've chosen to focus chiefly on the work of Stewart Shapiro. Though the choice is partially arbitrary, there are a few motivating factors. First and foremost is the fact that Shapiro often frames his structuralism in direct response to Benacerraf, Frege, and Hilbert. This makes the task of situating his work in relation to our driving concerns much easier. Shapiro's work is also more directly and frequently engaged in the ontological aspects of structuralism with which I am chiefly interested here.[71] Shapiro himself also provides a compelling argument which suggests that some of the key forms of assertoric structuralism are intertranslatable.[72] So while the choice

[71] This is not to suggest that other structuralists are not interested in ontological questions, only that Shapiro's ontological focus is more often front-and-center. Michael Resnik's work on a pattern-theoretic account of structures strikes me as focussed on epistemological concerns to the same degree that Shapiro focuses on ontological ones.
[72] Cf., [Shapiro 1997, 90-97]. There he writes that "Anything that can be rendered in the modal structural system […] can be rendered in either the set language or [Shapiro's own] structure language" [Shapiro 1997, 96] and, further, that "when all is said and done, the different accounts are equivalent" [Shapiro 1997, 96]. The modal structural system he refers to here is Hellman's, as developed in [Hellman 1989] and [Hellman 1996].

is at least partially idiosyncratic, there are (I think) plausible reasons for engaging chiefly with Shapiro's form of structuralism.

2.3.1 Shapiro on the nature of mathematical objects

For Benacerraf, the elements of at least abstract structures (the chief focus of mathematics) cannot be objects. As we saw, his reasons for holding this belief are rather puzzling. On the one hand, he suggests that there is no real content to be given to any theory-independent notion of object, and, on the other hand, he presumes precisely that to be an object is to possess properties and stand in relations independently of any given theory, language, or structure. In any case, his view leads him to claim that whatever numbers might be, they simply cannot be objects (at least in the sense of entities possessing properties or standing in relations independent of theory or structure). As a result, the unrestricted form of identity which Frege employed, one which ranges over all objects, must be replaced. Thus, Benacerraf's strange account of objects is at least partially intertwined with his rejection of unrestricted identity.

Shapiro finds this view as puzzling as we have. He writes that

> In mathematics, at least, the notions of 'object' and 'identity' are unequivocal but thoroughly relative. Objects are tied to the structures that contain them. It is thus strange that Benacerraf should eventually conclude that natural numbers are not objects. Arithmetic is surely a coherent theory, 'natural number' is surely a legitimate category, and numbers are its objects.[73]

Here I think we have a slightly more careful statement of what Benacerraf himself was trying to get at in his 1965 paper. The strangeness of Benacerraf's claim stems from the fact that there are two competing conceptions of 'object' at work in his paper. The first is Frege's: objects—mathematical or otherwise—are self-standing, independently existing entities. They exist prior to and irrespective of the relations they might stand in. This notion, both Benacerraf and Shapiro suggest, is not at all helpful when discussing the elements of abstract mathematical structures. To construe the elements of structures in terms of Frege-style objects is to become ensnared once again by the Caesar and non-uniqueness problems which structuralism seeks to avoid in the first place.

But, in contrast with Benacerraf, Shapiro is not ready to abandon mathematical objects *tout court*. There is a second notion of object at work here, one which is dimly present in Benacerraf's paper, but which gets lost or muddled in his later equivocations. This notion gets rid of the theory-independence and self-standing character of the Fregean object and, instead, suggests that mathematical objects, at least, are essentially related to the structures of which they are components. Or, more simply put, Shapiro writes that "mathematical objects are tied to the structures that constitute them".[74] This can seem somewhat arcane if we limit ourselves to mathematics, but Shapiro marshals several down-to-earth examples to help explicate the position. Most of us are, presumably, quite happy saying that the game of chess exists, in some

[73] [Shapiro 1997, 81].
[74] [Shapiro 1997, 80].

sense. Now, whatever metaphysical spin we want to give to this claim, it seems reasonable to admit that, if the game of chess did not exist, then the white queen's bishop could not exist either. Equally, most of us would probably agree that it *doesn't make any sense* to say that the white queen's bishop exists unless we already accept that the game of chess exists. None of this, according to Shapiro, ought to convince us that the white queen's bishop is not an object. For him, it is indeed an object, but one whose existence depends upon the existence of a structure we call 'the game of chess'. So, in contrast to Benacerraf, Shapiro is quite happy to hold that the elements of even abstract structures can be understood as objects—only they are objects whose existence is dependent upon the structures which they constitute.

But how exactly are objects related to the structures of which they are the elements? What does Shapiro mean when he uses the term 'constitute'? Shapiro's own position on this question has been refined in the last two decades. In his best-known work—his 1997 book *Philosophy of Mathematics: Structure and Ontology*—Shapiro gives some of the metaphysics behind his use of the language of constitution to describe the relationship between mathematical objects and abstract structures. There he writes that

> Each mathematical object is a place in a particular structure. There is thus a certain priority in the status of mathematical objects. The structure is prior to the mathematical objects it contains, just as any organization is prior to the offices that constitute it. The natural-number structure is prior to the offices that constitute it [i.e., the natural numbers].[75]

Here we see the core of his view. Mathematical objects are *places* in structures, and they depend for their existence on the prior existence of the structure. It is in this sense that they 'constitute' the structure, while also being dependent upon it for their existence. The chess analogy is again helpful: while the chess-independent existence of the white queen's bishop makes no sense, similarly, the game of chess is constituted in part by the place of the white queen's bishop. Without this place, we have a significantly different structure (or, perhaps, no structure at all).

Here we are reminded of Frege's criticisms of Hilbert's independence and consistency proofs. For Hilbert, these proofs involved the provision of models of the 'same' content. For Frege, such proofs involved the equivocation between distinct contents. Here we see that this feature of Frege's critique has an analogue in Shapiro's system. Structures are constituted by their places and relations. If we change either of these, we do not alter the structure but begin speaking about a different structure. This seems to contradict the everyday intuition that we can still speak of *the* game of chess or baseball despite the significant alterations in their rules over time. To what extent this intuition can be recaptured by Shapiro via descriptions of interrelations among distinct structures is an open question.

In a paper from 2008, Shapiro refines his understanding of the relation of constitution which holds between structures and their objects. He writes:

> Places are *components* of universals. Each *ante rem* structure consists of some places and some relations. If this makes metaphysical sense […] then the dependence relation in the

[75] [Shapiro 1997, 78].

> slogans for *ante rem* structuralism is that of constitution. A structure is constituted by its places and its relations, in the same way that any organization is constituted by its offices and the relations between them. The constitution is not that of mereology. It is not the case that a structure is just the sum of its places, since in general the places have to be related to each other via the relations of the structure. I think of an *ante rem* structure as a whole consisting of, or constituted by, its places and its relations.[76]

And, further, that

> the metaphysical view is that a structure is constituted by its places *and* its relations. Neither the places nor the relations are prior to the other. Again the slogan that the objects depend on the structure can be understood in that spirit. The dependence is constitution.[77]

Here we get a fuller sense of the view. A structure as a whole is constituted by both relations and places. The relations are not metaphysically prior to the places, nor are the places metaphysically prior to the relations (in fact we can't really understand either in isolation: what is the white queen's bishop without its allowable moves, what are the moves without the object?). It is only both together which constitute a complete structure, and, according to the last-quoted sentence, both are *dependent* for their existence on the structure which they constitute.[78]

So, abstract and not just concrete structures contain objects as their elements. We must now ask: for Shapiro, what is the relationship between the objects of an abstract structure *A* and the objects of a system *B* which exemplifies *A*? This is a fairly difficult question and requires an examination of some of Shapiro's terminology.

2.3.2 Places, positions, offices, and objects

In the above quotations, we've seen Shapiro refer to the elements of abstract structures by a number of terms: 'places,' 'offices,' and 'objects'. Elsewhere he also uses the term 'position'. For Shapiro, all these terms are, in a sense, inter-translatable. Nevertheless, there are some subtleties which will help us understand the relationship between the objects of abstract and concrete structures. Shapiro uses 'place' and 'position' as catchall terms which refer in general to components of abstract structures which are not themselves relations. He does not give precise definitions of these terms but, instead, provides elucidations through a wide variety of mathematical and non-mathematical examples. Much as Frege employs the primitive notion 'object' Shapiro suggests that "'places' are indeed simple, or atomic, in the sense that they do not themselves have places or other components. 'Place' is a primitive of structure theory."[79]

We've already established that, for Shapiro, places are objects. There are a few other characteristics of places which are important. First, we must understand places both as objects (in

---

76 [Shapiro 2008, 19]
77 [Shapiro 2008, 20].
78 Though, again, note that [Shapiro 1997] countenances degenerate cases of structures which contain no relations (his 'cardinal structures') or no elements (his '0 pattern').
79 [Shapiro 2008, 21].

the above sense) and as 'offices'. What does it mean for a place to be understood as an office? Shapiro writes:

> The structures studied in mathematics are what I call *free-standing*. Anything at all can occupy their places. So perhaps it is not amiss to think of a place in an *ante rem* structure as itself a universal, albeit a rather trivial one. Each place of each structure can be instantiated by anything at all, in a system that exemplifies the structure. Perhaps a place in a structure is just the universal 'thing' or 'object' or 'particular'.[80]

The language here deliberately echoes that of Hilbert: Shapiro begins by highlighting the fact that abstract structures are like mobile bits of scaffolding which can be laid over other parts of reality via exemplification, modelling, or instantiation. Here, too, Shapiro suggests, without using the term, exactly what it means for a place to be an office. If we are discussing the abstract natural number structure directly, then we can treat the number '2' as a structure-dependent object. This is the places-as-objects perspective, and it is the one that Shapiro believes is employed most often in pure mathematics. But the chief characteristic of abstract, versus concrete structures, is that their places possess no extra-structural properties, which allows these abstract structures, in turn, to be instantiated in or exemplified by other structures—Hilbert's 'scaffolding of concepts'. When we note, for example, that Caesar, my right index finger, the set {{Ø}}, or any other object at all can represent the number 2 within a suitably defined concrete structure, we are treating the places of the abstract natural number structure as *offices*. That is to say, we recognize that the very character of the object we call 'the number 2' involves the capacity to be represented or exemplified by any object whatsoever. This is fairly straightforward if we think of offices in the more familiar sense: the office of the registrar is a position with determinate relations to other positions within the university's administrative structure – in this sense it is a structure-dependent object. But it is of the essence of any bureaucratic office that it can be occupied or exemplified or represented by different people at different times.[81]

The chief insight of Shapiro's position here is that places are always *both* offices and objects, though different perspectives may highlight one or the other aspect of their nature. From the

[80] [Shapiro 2008, 23]. In reading this passage, recall from above Shapiro's statement that "In mathematics, at least, the notions of 'object' and 'identity' are unequivocal but thoroughly relative" [Shapiro 1997, 81]. [Shapiro 2008] speaks of places in structures as being capable of exemplification or representation by any object whatsoever (as long as the object is arranged in a suitable concrete structure, i.e., a system). It seems to me that Shapiro's remark here that we might treat places simply as the universal 'object' implies a notion of object which is *not* in fact relative, but which encompasses all objects from whatever system we like. Despite the Hilbert-style approach to structures, this notion of places that accept any objects as occupants is strikingly reminiscent of Frege's understanding of first-level functions. As we will see in what follows, this similarity perhaps runs deeper than is obvious here.

[81] Of course the office of registrar, unlike the office of the natural number 2, cannot be occupied by *any* object whatsoever. An air conditioner, for instance, cannot fulfill the duties of registrar, though it might very well act as a perfectly adequate number 2. Shapiro also attempts (at [Shapiro 1997, 98-99]) to distinguish between abstract and concrete structures on the basis of the sort of relations they contain. He does so using a version of Alfred Tarski's (cf., [Tarski 1986]) interesting distinction between the logical and the nonlogical. For Shapiro, the reason why an air conditioner or 'any old object' is incapable of acting as a registrar is that there are implicit non-logical (or non-formal) constraints on the allowable relations (e.g., the registrar must be a sentient, physical being).

perspective of a mathematician studying an abstract structure directly, the places of the structure appear to be (and are, according to Shapiro) bona fide objects. From the perspective of someone modelling an abstract structure, or applying it, the elements of the structure are best understood as offices, which can be occupied by any object whatsoever. The key thing to keep in mind is that the relationship of occupation or exemplification here is *not* the same as identity, even though a place might be said to exemplify itself.[82] When I occupy or exemplify the number 2, I do so relative to a particular concrete structure. The difference between myself as an occupant of the 2-place in this structure and *the* number 2 of the abstract natural number structure is that I still possess extra-structural properties even when I exemplify the number 2. To be sure, we ignore these properties when I occupy the position, but, in the case of the actual number 2, the 2-place of the *abstract* natural number structure, there are simply no properties to ignore.

### 2.3.3 Cross-structural identity

Part of Benacerraf's reason for rejecting the view that numbers can be objects is that he found claims which sought to identify wildly different types of objects to be 'unsemantical'. Thus he claimed that if we want to identify a particular number with a particular set, we must first provide a well-defined context within which to do so. The general (i.e., non-relative) notion of 'object' does not provide such a context, so we cannot discuss identity relations between entities on this basis alone. If we read Benacerraf as a progenitor of structuralism, we can understand well-defined structures as a means of providing the relevant context for identifying (structure-dependent) objects with one another. On this view, then, Benacerraf's rejection of Caesar-type questions as unsemantical amounts to the claim that we cannot make cross-structural identity claims like those apparently made between the natural number structure and the set-theoretic structure *unless* the objects of these structures are (perhaps tacitly) understood to be elements of a single structure, a single well-defined context. But of course the traditional way of providing this single structure has been to reduce one structure (the natural number structure) to another (the set-theoretic hierarchy).

It seems that Shapiro follows Benacerraf in much of this. He writes, for instance that "a philosophy of mathematics should show why these questions [i.e., 'does Julius Caesar = 2?'] need no answers, even if the questions are intelligible"[83] and that "Identity between natural numbers is determinate; identity between numbers and other sorts of objects is not, and neither is identity between numbers and the positions of other structures".[84] And, as we've already seen, he holds that "In mathematics, at least, the notions of 'object' and 'identity' are unequivocal but thoroughly relative".[85]

Given these views, one would expect that Shapiro would follow Benacerraf's *conclusion* as well, namely that we cannot make cross-structural identity claims. For, without a determinate context

---

[82] "When we switch back to the places-are-objects perspective—and construe the places of a given structure as objects—we see that each structure exemplifies itself. That is, the relations of the structure hold of its places—when the latter are construed as objects" [Shapiro 2008, 24].
[83] [Shapiro 1997, 79].
[84] [Shapiro 1997, 79].
[85] [Shapiro 1997, 81].

(i.e., a well-defined structure), it does not make any sense to speak of objects or an identity relation in the first place. But he does not follow Benacerraf's conclusion. Instead, he writes:

> I do not wish to go as far as Benacerraf in holding that identifying positions in different structures (or positions in a structure with other objects) is always meaningless. On the contrary, mathematicians sometimes find it convenient, and even compelling, to identify the positions of different structures. […] For [example], it is surely wise to identify the positions in the natural-number structure with their counterparts in the integer-, rational-, real-, and complex-number structures. […] The point here is that cross-identifications like these are matters of *decision*, based on convenience, not matters of discovery.[86]

So, Shapiro wants to allow that we can make meaningful claims about identity between objects in distinct structures but that this is, in some sense, a product of a decision. I'm not entirely sure what he can mean here, and what the contrast with Benacerraf's position is. Perhaps the 'decision' is to combine our distinct theory-bound notions of identity and object into another, by defining (at least tacitly) a more inclusive structure. For, if the claim of identity is to be a meaningful one, then it will have to be unequivocal and relative to a particular structure (on Shapiro's view). Whether or not we decide this to be the case seems to be beside the point, for irrespective of the question of decision or discovery, the identity claim will require relativization to a particular structure. Note, too, that Shapiro is committed to the independent existence of structures (i.e., to their existence regardless of our decisions about them).[87] Without further elaboration on Shapiro's part, then, it seems slightly misleading to suggest that the objects of different structures can be identified with one another, as such identification requires a kind of equivocation between distinct notions of object and identity. I think Shapiro's refusal to accept the impossibility of cross-structural identification in fact belies a deeper commitment, on his part, to precisely the kind of absolute, universal background ontology which structuralism is (at least putatively) concerned to avoid. I think we can make this point most clearly if we turn now to an examination of Shapiro's account of abstract structures themselves.

### 2.3.4 Shapiro on structures

So, what are structures for Shapiro, exactly?

In our discussion of objects, we've already gotten a bit of insight into the nature of structures for Shapiro. They are, as we've seen, constituted by their elements, i.e., their places, and relations on those objects. As far as abstract structures go, the nature of these objects does not matter at all, only that they stand in the appropriate relations with one another. Shapiro gives a slightly more formal definition of structure along these lines:

> a *structure* has a collection of *places* and a finite collection of functions and relations on those places. The isomorphism relation among structures and the satisfaction relation

[86] [Shapiro 1997, 81].

[87] He writes, for example, that "Structuralists hold that a nonalgebraic field like arithmetic is about a realm of objects—numbers—that exist independently of the mathematician" [Shapiro 1997, 72] and, in a later paper, that "the structure itself is a chunk of reality" [Shapiro 2005, 67].

> between structures and formulas of an appropriate formal language are defined in the standard way. We could stipulate that the places of different structures are disjoint, but there is no reason to do so.[88]

In relation to our earlier discussion of the possibility of cross-structural identity claims, the last sentence of the above quotation is quite revelatory. He claims that we could, if we so desired, stipulate that the places of each structure are distinct from the places of any other structure. This, of course, would be in keeping with Benacerraf's dictum that there can be no cross-structural identity claims. For *Benacerraf*, then, there is a very good reason to stipulate (or at least recognize) that the places of different structures are disjoint. But, as we see here, Shapiro holds that there is no reason to do so. Why?

Though it is not made explicit here, I suggest (and will argue in what follows) that, for Shapiro, the reason that we need not explicitly state that the places of different structures are disjoint is because *they are not in fact disjoint on Shapiro's account*. He has suggested that the notion of 'object' (and, thus, that of 'place' as well) is relative to a particular structure. Or, more broadly, he argues that ontology is relative to a given structure. But here he seems to implicitly ignore this relativization by suggesting that places can be 'shared' between distinct structures. How is this possible, given his previous account? It is possible, I will argue, because there is, ultimately, only *one* ontology for Shapiro, and the function of his structure theory is to describe this ontology. Indeed, the only point in having a theory of structures at all is to provide this background ontology.

### 2.3.5 The ontology of structures

As we've seen, structuralism in the broadest terms involves the attempt to move beyond the uniquely fixed ontology of set-theoretic foundationalism, and, instead, to suggest that mathematics is the science of structure. Thus, it would be rather surprising if a dyed-in-the-wool structuralist like Shapiro simply recapitulated the traditional set- and model-theoretic foundations of mathematics.

Yet, Shapiro takes his own axiomatic theory of structures to be loosely imitative of second-order Zermelo Fraenkel set theory,[89] while also thinking of it as "an axiomatization of the central framework of model theory".[90] And, indeed, his few axioms mimic very clearly traditional formulations of axiomatic set theory.[91] So, if Shapiro's theory is to differ significantly from the traditional set-theoretic foundationalism he wishes to combat, the difference cannot be solely in the form and content of his axioms. The difference will be in his philosophical account of the

---

[88] [Shapiro 1997, 93]. Presumably the 'standard way' referred to here is that of model theory. In what follows, we will see that this implicit reliance on the model-theoretic approach is widespread in Shapiro's structure theory, as well as problematic for his goals.

[89] [Shapiro 1997, 95].

[90] [Shapiro 1997, 93].

[91] For instance, in place of the usual set-theoretic axiom of infinite, which holds that there exists a set with an infinite number of members,, Shapiro's axiom of infinity holds "There is at least on structure that has an infinite number of places" [Shapiro 1997, 93], in the place of the usual power *set* axiom, he has a power *structure* axiom, etc., etc.

point of these axioms, or of the nature of the items about which the axioms speak (e.g., structures).

Like the traditional set-theoretic foundationalist, Shapiro accepts the necessity of a background ontology. In the traditional case, this ontology was the universe of set theory (and usually that of ZFC – Zermelo-Fraenkel set theory with the axiom of choice). On the traditional picture, this ontology is importantly non-structural: claims made about the ontology of ZFC are understood to be claims about absolute features of the world. This means that the fundamental notions of the ontology (notions like membership, for instance) are not 'relative' but, instead, possess non-structural properties just like an everyday object like my keyboard. This absolute ontology is exactly the sort of thing that structuralism is designed to oppose—as evidenced by the origins (of some forms) of structuralism in Benacerraf's critique of the viability of set-theoretic foundationalism. Benacerraf wanted to show that it was nonsensical to pin down mathematical structures like the natural number structure within the set-theoretic hierarchy. And, moreover, Skolem, Gödel, and Paul Cohen (among others) had by 1965 surely already shown that the very notions of set theory *had* to be relative themselves.[92]

So in thinking about Shapiro's background ontology, we have a few questions to answer.

(i) Why does Shapiro think it is necessary to presuppose a background ontology for structuralism?
(ii) How does his background ontology differ from the absolute ontology of traditional set-theoretic foundationalism?

Let's tackle these in order

2.3.6 The necessity of a background ontology

So, why do we need a background ontology in the first place? Why not eschew the universality so characteristic of set theory and instead allow for multiple distinct, disjoint ontologies? To answer this, we have to examine once more the role that the elements of structures play in Shapiro's account.

Recall that Shapiro outlines two different (though ultimately inseparable) ways of understanding the places of structures: the places-as-objects view and the places-as-offices view. He suggests that, on the places-as-offices view, we must have enough objects to actually fill the offices of the structure. The structure, especially if it is to be 'real' in Shapiro's sense, *must* be constituted by its objects and its relations. So we need objects to have a structure, and (for abstract structures at least) *vice versa*. This is why, in our discussion of any given structure, we *need* an ontology, however far in the background it might be when we actually get down to our mathematical work on the structure.

[92] Cf., [Skolem 1922], where the implications of Skolem's 'paradox' are very clearly drawn. In two papers ([Cohen 1963], [Cohen 1964]) Paul Cohen further dismantled the absolute picture of the set-theoretic hierarchy by producing models of ZFC within which the continuum hypothesis fails, alongside Gödel's models of ZFC wherein the continuum hypothesis is true.

We need a background ontology but, importantly, when we are dealing with mathematical (i.e., abstract) structures, this background ontology can contain any objects we please. Recall that the distinguishing feature of abstract structures is that the nature of their elements is a matter of indifference – any object at all can exemplify an element of the natural number structure (i.e., a number) as long as it stands in the appropriate relations to the other exemplars of the elements of the structure. So, the nature of the contents of the background ontology should not and does not matter for Shapiro's theory. Whether the objects are sets or hats or chimneysweeps or whatever, none of this matters for the background ontology. What does matter is that there are sufficiently many objects to fill all the places. And, since many of the interesting structures within mathematics are not only infinite but *uncountably* infinite, Shapiro's theory requires a background ontology which contains (at least) an uncountable infinity of objects as well.

Now, as we've seen, for Shapiro, the places of an abstract structure are also objects in their own right. This fact is embodied in what he calls the places-as-objects perspective. This means that we can simply populate the requisite background ontology of any given structure by the places of that very structure. The offices of the structure can be occupied by the objects of the structure (though, of course, they could be anything else as well).

So far so good. We should now examine how this (necessary) background ontology differs from that of the traditional set-theoretic account which Shapiro's structuralism is meant to oppose.

On the traditional set-theoretic view, the background ontology is populated by objects whose properties are not solely dependent upon a given structure. These can be either sets or urelements or both. In set theory (as in structure theory) we are interested in examining structures with infinitely many elements. Often, we are interested in structures with an uncountable infinity of elements. Thus, on this approach, our background ontology requires infinitely many structure-independent objects. And so, in set theory, we have axioms which posit such entities.

For Shapiro, since the elements of structures are not 'structure-independent' in the way that sets and urelements are, this set-theoretic approach to the background ontology will not do. But, like the traditional set-theoretic foundationalist, Shapiro still wants to examine structures with infinitely many elements. Again, recall that, for Shapiro, structures are prior to their elements, and these elements are entirely dependent upon the structures they constitute. So, if we want structures with infinitely many elements, we cannot begin by postulating infinitely many objects to act as the elements of the structure. Instead, Shapiro suggests, we begin by postulating *structures with infinitely many places*. The 'background ontology' is thus not filled with structure-independent objects like sets or urelements but with structures themselves. And, since the places of structures can occupy themselves, the existence of infinite structures guarantees the existence of infinite systems which exemplify those structures.

On the traditional dispensation, set theory itself is not a structural theory, insofar at least some of its elements are understood to be 'concrete'. Shapiro's view ostensibly allows for a more thoroughly structural approach by insisting that the elements of his ontology (i.e., structures) are themselves understood in structural terms.

This approach to the background ontology is meant to improve not only on the traditional set-theoretic approach to structures, but also to solve a problem to which eliminative forms of structuralism are prone.[93] The problem, as Shapiro sees it, is one of infinite oscillation or regress between systems and structures. The eliminative structuralist *qua* structuralist wishes to study structures. In order to do so, a background ontology is employed which is itself non-structural. But, given the eliminative structuralist's general structural views, this background ontology itself exhibits a structure. This structure again requires a background ontology, which again will exhibit another structure, and on and on. Shapiro's solution aims to "stop the regress of system and structure at a universe of structures".[94]

Shapiro's putative solution to this regress is to identify structure and system one with the other. But does his solution actually stop the regress? There is good evidence that, in the final analysis, it does not. Let's take a look.

2.3.7 Stop the regress I want to get off

A good way into the problem is to examine Shapiro's criticisms of the views which he believes are subject to this regress. Of a set-theoretic approach to structuralism, he writes that

> The crucial feature of this version of eliminative structuralism is that the background ontology is not understood in structuralist terms. If the iterative hierarchy is the background, then set theory is not, after all, the theory of a particular structure.[95]

In short, this approach is not a structural one, because it ends up being founded upon a theory whose very basis is 'concrete', i.e., one whose elements have non-structural properties. The structural aspect of mathematics is actually founded upon a system. He makes much the same argument about the modal form of structuralism which is developed by Geoffrey Hellman. For Hellman, structures are not the final item on the ontological menu, but are instead eliminated in favour of the consideration of possible systems. The problem, as Shapiro sees it, is that the usual way of dealing with logical possibilities (particularly in the kind of modal language that Hellman employs) is to characterize them in terms of the set-theoretic hierarchy as employed in model theory. Shapiro writes that

> Hellman accepts this straightforward point, and so he demurs from the standard, model-theoretic accounts of the logical modalities. Instead, he takes the logical notions as *primitive*, not to be reduced to set theory.[96]

---

[93]The term 'eliminative structuralism,' as Shapiro characterizes it, covers any form of structuralism which 'eliminates' structures in favour of some background ontology which does not itself contain structures. Thus, for instance, Hellman's modal eliminative structuralism employs a background ontology of possible systems.

[94] [Shapiro 1997, 92]. The details of this regress are not made clear by Shapiro, nor does he give a clear account of why the eliminativist can't simply model the structure of the background ontological system in that system itself.

[95] [Shapiro 1997, 87]. Shapiro calls any version of structuralism which eschews direct quantification over structures in favour of some other ontology 'eliminative'. Both the traditional set-theoretic approach to structures and Geoffrey Hellman's modal account of structures fall under this heading for him.

[96] [Shapiro 1997, 89].

But, Shapiro continues later in the work,

> because Hellman is out to drop the realist perspective, it is not clear why he is entitled to the traditional model-theoretic explications of the modal operators of logical necessity and logical possibility. For example, the usual way of establishing that a sentence is possible is to show that it has a model. For Hellman, presumably, a sentence is possible if it might have a model (or if, possibly, it has a model). It is not clear what this move brings us.[97]

On the readings Shapiro offers here, his chief criticism of 'eliminative' forms of structuralism is that they inevitably end up relying on a non-structural ontology of some kind. In the case of the set-theoretic approach to structuralism, this is very straightforwardly the ontology of set theory. In Hellman's case, rather more circuitously, he ends up relying (says Shapiro) on the very same ontology simply by his unexplained reliance on the usual model-theoretic semantics for modality.

In response to these views, he writes that

> A structuralist might be tempted to step back from this competition of background theories and wonder if there is a structure common to all of them. However, on the ontological option, this temptation needs to be resisted. The structures studied in two theories can be compared only in terms of a more inclusive theory.[98]

The 'ontological option' here is the option of embedding abstract structures within a particular, determinate ontology. But, again, if the goal is to understand structures, and multiple available ontologies seem to 'model' these structures equally well, we have no reason to choose between one or the other. And, moreover, since our interest is in the structures themselves, the use of an ontology whose elements possess non-structural properties is ultimately opposed to the structuralist programme. All of this simply recapitulates, perhaps on a grander scale, the non-uniqueness problem we began with. So Shapiro needs an ontology which stops the regress, but which does not involve the reduction of structures to a concrete or non-structural background ontology.

The obvious solution, and the one which we've already mentioned, is to treat structures themselves as the background ontology by developing a theory of structures. The difficult part here will be in developing an account of structures which does not itself rely on non-structural notions.

### 2.3.8 Shapiro's structure theory

So Shapiro sets out to develop just such a theory. Note that he is quite happy to claim (and he gives compelling arguments supporting this claim) that his structure theory is intertranslatable

[97] [Shapiro 1997, 229].
[98] [Shapiro 1997, 87].

with both the set-theoretic picture and Hellman's modal/eliminative approach. So what is the difference?

Well, what he means by intertranslatable is not that they say the exact same thing (for he doesn't want to make any claims about non-structural entities like the traditional set-theorist's non-structural sets). He writes

> there are several ways to render structuralism in a rigorous, carefully developed background theory, and there is very little to choose among the options. In a sense, they all say the same thing, using different primitives. The situation with structuralism is analogous to that of geometry. Points can be primitive, or lines can be primitive. It does not matter because, in either case, the same structure is delivered. The same goes for structuralism itself. Set theory and structure theory are equivalent in the sense defined above. To speak loosely, the same 'structure of structures' is delivered. Modal structuralism also fits, once the notion of 'equivalence' is modified for the modal language.[99]

So, it isn't that they make exactly the same claims, but that they, ultimately, say the same things *about structures*. The key difference between them, according to Shapiro, is that his ontology never bothers to make the detour through a concrete system of sets or possible systems: it stays with structures the whole time.

2.3.9 Quasi-concreteness

But, near the close of this chapter on ontology, we find Shapiro making some rather remarkable claims about the relationship between his theory of structures and 'concreteness' (or, what amounts to the same thing for our purposes, 'quasi-concreteness'). The term 'quasi-concrete' Shapiro borrows from a 1990 paper by Charles Parsons. In that paper, Parsons presents a sustained argument against the viability of a *purely* structural account of mathematics like the one ostensibly offered here by Shapiro. On the basis of the importance he accords to the role of quasi-concrete objects and concepts within mathematics, Parsons concludes finally that

> if the structuralist view of mathematical objects is taken to mean that all mathematical objects are only structurally determined, it has to rest on legislation about what counts as a mathematical object. The explanatory and justificatory role of more concrete models implies, in my view, that it is not the right legislation even for the interpretation of modern mathematics.[100]

Given this conclusion, it is somewhat surprising (as will be shown presently) that Shapiro seems to endorse much of Parsons' argument here, recognizing, in particular, that his own *ante rem* structuralism relies inescapably on the quasi-concrete. For Parsons, an entity is quasi-concrete if it is

[99] [Shapiro 1997, 97].
[100] [C. Parsons 1990, 338].

> directly 'represented' or 'instantiated' in the concrete. Examples might be geometric figures (as traditionally conceived), symbols whose tokens are physical utterances or inscriptions, and perhaps sets or sequences of concrete objects. The 'concrete objects' that David Hilbert talked about in his accounts of intuitive, finitist mathematics are in my terminology quasi-concrete. A purely structuralist account does not seem appropriate for quasi-concrete objects, because the representation relation is something additional to intrastructural relations. Because they have a claim to be the most elementary mathematical objects, and also for other reasons, quasi-concrete objects are important in the foundation of mathematics.[101]

This notion of the quasi-concrete is perhaps more widespread within mathematics than it might appear at first glance, for any proof or theorem which refers to or quantifies over, e.g., the formulas of a formal language will, on this dispensation, be quasi-concrete. In any case, the precise extension of the notion is not important for us here. What *are* important are the conclusions that Shapiro draws about his own theory, under the influence of Parsons' notion.

After his discussion of Parsons, Shapiro gives three distinct reasons why his own theory of structures cannot but be understood, at least partially, in quasi-concrete terms. These reasons are:

(i) His (crucial) account of the concept of coherence for structures is implicitly motivated/justified by quasi-concrete systems.
(ii) His definition of 'structure' itself entails reference to quasi-concrete entities/procedures.
(iii) His use of a second-order language involves the use of quasi-concrete notions.

Rather than enter into an examination of Shapiro's (admittedly problematic) notion of coherence for structures, I will examine the implications of the quasi-concreteness pertaining to the definition of 'structure' and his use of a second-order language. Regarding his definition of 'structure,' he writes

> Another reason to think that the quasi-concrete cannot be eliminated is that I have appealed to quasi-concrete items in order to define the very notion of a *structure*. Recall that a structure is the form of a system, and a system is a collection of objects under various relations. The notion of 'collection' is an intuitive one. There is something fishy about appealing to the set-theoretic hierarchy, a freestanding *ante rem* structure, in order to explicate the notion of 'collection' in the characterization of 'system' and thus 'structure'. Where did we get on this merry-go-round, and how do we get off?[102]

Here we see that the problem of the quasi-concrete ultimately puts him back into the regress he sought to escape. His way of escaping the regress was to put structures directly into his ontology. But in his definition of structure, he employs the notion of 'collection'. This, according to his own rendering here, is a quasi-concrete notion which relies directly on our intuitive understanding of collecting (concrete) items together. So it seems that his attempt to characterize

[101] [C. Parsons 1990, 304].
[102] [Shapiro 1997, 105].

the set-theoretic hierarchy as a 'freestanding *ante rem* structure' actually involves the more usual, intuitive notion of sets as concrete collections. Thus, the mechanism of his escape from the concrete account of mathematics itself involves an inescapable reference to the concrete. This, on its own, seems deeply troubling for his purely structuralist account.

But he also notes that his use of a second order language involves the quasi-concrete. He writes

> A related point concerns the practice of characterizing specific structures using a second-order language. Such languages make literal use of intuitive notions like 'predication' or 'collection'. A crucial step in the defense of second-order languages is that we have a serviceable, intuitive grasp of notions like 'all subsets'. This notion is also quasi-concrete.[103]

Thus, another deep problem emerges in the very language and logic used to describe his structuralist programme. In the face of the problem of ineliminable reference to quasi-concrete objects and notions, Shapiro writes

> In all cases, then, the conclusion is the same. We can try to hide the quasi-concrete, but there is no running away from it. […] However, the caveat does not undermine the main ontological thesis of *ante rem* structuralism, the idea that the subject matter of a branch of pure mathematics is well construed as a class of freestanding structures with formal relations. The role of concrete and quasi-concrete structures is the motivation of structures and the justification that structures with certain properties exist. The history of mathematics shows a trend from concrete and quasi-concrete systems to more formal, freestanding structures. There is no contradiction in the idea of a system of quasi-concrete objects' exemplifying a freestanding *ante rem* structure.[104]

What are we to make of all this?

It strikes me that Shapiro's final claim here, that there is no *contradiction* in his use of quasi-concrete objects and freestanding systems might very well be true. But so what? The truth is that he has *not* given us freestanding structures, precisely because his definition of 'structure,' as well as the language used to construct the theory of structures, both necessarily involve the use of quasi-concrete notions. Thus, the resulting structures are themselves quasi-concrete. If the whole endeavour of his particular form of structuralism was to avoid the pitfalls of an absolute ontology, and, instead, to develop an ontology of structures which contains no non-structural entities of any kind, then these admissions indicate that he has, rather straightforwardly, failed. I know of no other way to read them. If Shapiro was content to sweep the essential ontological role of these quasi-concrete objects under the rug, then the detour into 'pure' structuralism was ill-motivated in the first place: we might just as well have ignored Benacerraf's qualms and stuck with traditional set-theoretic foundationalism from the beginning, ignoring ontological concerns along the way.

---

[103] [Shapiro 1997, 105].
[104] [Shapiro 1997, 105-106].

In the end, then, what Shapiro's admission of the concreteness of his system shows is that the assertoric form of structuralism (exemplified in, at least, the work of Shapiro and Hellman) is, ultimately, another form of reductionism, which proposes to give us a new array of ontologies, but no good reason to choose between them. After all, they are intertranslatable on Shapiro's reading. If our statements *must* be about some particular non-structural ontology, but we can't choose which one, it seems as though we are not capable of picking out what mathematics is *really* about. At best, we have a variety of things that it might be about, and this, surely, is an insufficient response to Benacerraf's challenge. If we take the assertoric approach to structuralism, then, we still have no clear answer to the sort of non-uniqueness problems faced by Frege and the set-theoretic reductionists.

## 2.4 Algebraic structuralism

### 2.4.1 Categories

Before we get to the philosophical content of algebraic structuralism, let's begin by briefly examining some of its mathematical underpinnings in category theory. What exactly is a category? In a moment I will provide a formal definition of the concept. Before this however, it will be useful to explain what categories are meant to do. Or, perhaps better, what sorts of mathematical activity they are intended to facilitate.

In some ways, the notion of a category itself is a secondary matter in category theory. The chief aim of category theory is to provide a means of 'translating' between different mathematical structures. The emphasis is often on the translations themselves, rather than on the categories (which can themselves be understood as object-like or morphism-like as we see fit). The point of category theory, at least in its initial incarnations, was to provide a theory of transformations between different types of structures. These transformations are called 'functors' and category theory might have been better named 'functor theory'. In one of the founding documents of category theory—Eilenberg and Mac Lane's 1945 paper, "General Theory of Natural Equivalences"—the authors have this to say about the relative priority of categories and functors:

> the whole concept of a category is essentially an auxiliary one; our basic concepts are essentially those of a *functor* and of natural transformation […]. The idea of a category is required only by the precept that every function should have a definite class as domain and a definite class as range, for the categories are provided as the domains and ranges of functors. Thus one could drop the category concept altogether and adopt an even more intuitive standpoint, in which a functor such as 'Hom' is not defined over the category of 'all' groups, but for each particular pair of groups which may be given.[105]

And, further, that

> Perhaps the simplest precise device would be to speak not of *the* category of groups, but of *a* category of groups (meaning, any legitimate such category). A functor such as 'Hom' is then a functor which can be defined for any two suitable categories of groups […]. Its

[105] [Eilenberg and Mac Lane 1945, 247].

> values lie in a third category of groups, which will in general include groups in neither [of the first two categories[106]

I doubt that today many category-theorists would like to entirely dispense with the notion of category, but the emphasis of the above remarks is nevertheless illustrative of the general character of category theory today: the nature of the categories and their objects doesn't matter, it's the transformations between them (and the transformations between transformations, and so on) which reveal new and interesting mathematical structure. So, despite the importance of the notion of category, keep in mind in relation to the following definition that categories are best understood in relation to certain types of transformations.

With that stage-setting out of the way, we can now ask again: what exactly is a category?

I will give a definition of the term presented by Steve Awodey. Others could have been used,[107] but Awodey's is concise and manageable in that it does not involve any dependence on other mathematical terminology. He writes:

> A *category* by definition […] consists of *objects* $A$, $B$, $C$, … and *morphisms* $f$, $g$, $h$, … such that: (i) every $f$ has a unique domain $A$ and a unique *codomain* $B$, written $f\colon A \to B$; (ii) given any $g\colon B \to C$ there is a unique *composite* $g \circ f\colon A \to C$, with composition being associative; (iii) each $B$ has an identity $1_B\colon B \to B$ which is a unit for composition, i.e., $1_B \circ f = f$ and $g \circ 1_B$ for any $f$ and $g$ as stated. A category is *anything* satisfying these axioms.[108]

For our purposes here, there are two relevant features of Awodey's characterization. First: a category contains both objects and morphisms (which are also called maps, mappings, or arrows). This already differentiates it immediately from the traditional account of set theory, which contains only objects, i.e., sets (or, possibly urelements as well) as basic elements. The second thing to note is that a category is anything at all that satisfies these axioms. Implicitly, for instance, a category might contain Frege's functions as its objects, much to his horror. Since these axioms give very few constraints on what can count as morphisms and no direct constraints at all on objects, category theory begins with a fairly indeterminate 'ontology' if it begins with one at all. The resulting generality means that category theory can describe an incredible array of structures. A few characteristic examples will help illustrate this point.

1. The category **Set**, whose objects are sets and whose morphisms are functions between sets.
2. The category **Cat**, whose objects are categories and whose morphisms are functors between categories.
3. The category **Top**, whose objects are topological spaces and whose morphisms are continuous maps.
4. The category **K-Vect** whose objects are vector spaces over a fixed field $K$ and whose morphisms are $K$-linear transformations

---

[106] [Eilenberg and Mac Lane 1945, 247].

[107] E.g., [Mac Lane 1998, 7ff.] or [Lawvere and Schanuel 1997, 21].

[108] [Awodey 1996, 212-213].

These are some of the better known examples of categories. We can also think of categories in relation to what we would normally call logical structures. Thus, for instance, any given deductive system can be understood as a category whose objects are the formulas of the system and whose morphisms are derivations or proofs of those formulas.[109] Taken on their own, this definition and these examples don't yet tell us very much about what makes categories interesting. To see why many have thought of category theory as a new way of doing mathematics, we have to see categories in action.

And for this, we need the crucial notion of a functor.

### 2.4.2 Functors

The term 'functor' was taken by Eilenberg and Mac Lane from the linguistic work of Rudolf Carnap.[110] Though Carnap's philosophical views seem (in some respects) in line with the pluralist approach of many category-theoretically minded philosophers of mathematics, his conception of a functor is somewhat different than the category-theoretic notion. In category theory, a functor is a homomorphism, or structure-preserving morphism, between two categories. More formally, a functor $F$: $\mathbf{A} \to \mathbf{B}$ has the following properties:

1. To each object $X$ of the category **A** it associates an object $FX$ of the category **B**
2. To each morphism $f$ of the category **A** it associates a morphism $Ff$ of the category **B**
3. For each object $X$ of category **A**, $F1_X = 1_{FX}$
4. For any maps $f$ and $g$ of the category **A** such that $f$: $X \to Y$ and $g$: $Y \to Z$, we have $F(f \circ g) = Ff \circ Fg$

In other words, the functor is defined for all the objects and morphisms of the categories in question, and it preserves identity, composition, and the domain/codomain structure. Note here that the categories **A** and **B** needn't be distinct. Thus for instance we have 'endofunctors' which simply map a category to itself—a specific example of which are identity functors, which associate every object with itself and every morphism with itself. Identity functors are usually not of much interest, but there are other functors from categories to themselves which are of mathematical interest. For instance, if we take the category **Set**, we can define a 'power set functor' P: **Set**→**Set** which maps each set to its power set and each function to its image. In fact, some of the most interesting functors are morphisms from a category into or onto itself.

So, what is philosophically interesting about functors? Intuitively, functors provide us the means to 'move between' different categories while examining what we might call their 'shared structure'. In a certain sense, they might be said to transform categories back and forth into one another. We might also say that they allow us to examine bits of shared mathematical structure without getting bogged down in conceptual concerns about the 'nature' of a given category –that

---

[109] This basic insight lies at the heart of the field of categorical logic, first developed in [Lawvere 1963] and [Lawvere 1964].

[110] Though similar in many respects, Carnap's functors are defined only for linguistic expressions. See [Carnap 2002, 14].

is, they allow us a kind of indifference to the particular sorts of things which count as objects and morphisms in our categories. Understood in this way, functors can be a useful conceptual tool for a structuralist approach to mathematics.

In addition to their capacity to isolate shared structure, functors can also be employed to reveal or construct new structure. This feature is perhaps best understood via the notion of a functor category, a concept which exhibits quite well the tendency towards algebraic abstraction so characteristic of category theory.

Recall that a category consists of objects, morphisms, a domain/codomain structure, identity, and composition. The collection of all functors between two categories can themselves be thought of as the objects of another category, i.e., a functor category. In order to see how we might define such a category, we first need another crucial category-theoretic concept: that of the natural transformation. Intuitively, a natural transformation is a way of transforming one functor into another such that the compositionality of the morphisms of the underlying categories is respected.[111] Now, with this notion of natural transformation in hand, we can give a clearer idea of what a functor category is.

We start out with two categories, **A** and **B**. We can then define a category of functors from **A** to **B** (often written 'Funct(**A**, **B**)'). The objects of this category are the functors F: **A**→**B**. The morphisms of this category are the natural transformations between these functors. I've mentioned functor categories here not because they are particularly useful (though they certainly *are* useful) but because they represent exactly the sort of algebraic abstraction which is characteristic of the category-theoretic approach. I'll try to explain what I mean.

Let's place ourselves in Frege's shoes for a moment. Perhaps the most essential core of Frege's 'background ontology' is the division between objects and functions. On his view, this is not a relative distinction but an absolute one. Any formal violation of this distinction would have disastrous results for his attempt to found mathematics. Let's see how the (linguistically similar) distinction between objects and morphisms plays out in the category-theoretic world. We might begin with Frege's universe as a category, let's call it **Frege**. The objects of **Frege** are just Frege's objects (i.e., thoughts, dogs, numbers, etc.) while the morphisms of this category are Frege's functions. We might be interested in some of the structure of this category, so we seek to define some functors from **Frege** to **Frege**. Thus, for instance, we might define the functor *F*: **Frege**→**Frege** which maps every object onto itself, and every function onto itself. We might then bundle together all such functors into a functor category Funct(**Frege, Frege**), whose objects are the functors between **Frege** and **Frege** and whose morphisms are the natural transformations between these functors. Then, if we are so inclined, we might define functors between the categories **Frege** and Funct(**Frege**, **Frege**), then another functor category, and so on, ad infinitum. This example isn't meant to be exciting or revelatory of any interesting mathematical content (and surely it isn't).

What it shows is that the kind of absolute distinction that Frege wants to make between objects and functions just doesn't make any sense in the category-theoretic world. Nor does it make

[111] For a more precise technical definition, see [Mac Lane 1998, 16].

much sense to ask which categories are the real ones, the ones out of which all other categories are 'built'. Instead, what the category-theoretic language of functors, functor categories, and natural transformation shows is that the questions of structure have no 'final' or absolute answer: they must be given relative to a particular mathematical situation. On this model, to think of categories in the static terms of set theory might be *possible* but it certainly seems to miss the point. New structures can be built quite easily, and what might be a morphism of one category is an object of another category, or it may yet play the role of a functor between categories in another context. The question of the exact nature of the entities at issue is beside the point: what matters is the structure which is revealed by particular sorts of transformations.

### 2.4.3 The philosophical aspects of algebraic structuralism

With this rough conception of category theory, we will now have an easier time understanding the philosophical commitments of algebraic structuralists. Let's look at these now and see if this form of structuralism can move us in any significant way beyond the positions dictated by the Frege-Hilbert debate. Algebraic structuralism, inspired in large part by the methodology of category theory, holds that mathematics possesses no final or absolute ontology. On some dispensations, if it is meant to provide foundations at all, it is in a rather different sense than the ontological foundations we've examined thus far. Take, for instance, the following comment by F. William Lawvere, one of the key proponents of the foundational importance of category theory:

> In my own education I was fortunate to have two teachers who used the term 'foundations' in a common-sense way (rather than in the speculative way of the Bolzano-Frege-Peano-Russell tradition). […] The orientation of [the works of his two teachers Samuel Eilenberg and Clifford Truesdell] seemed to be 'concentrate the essence of practice and in turn use the result to guide practice'. I propose to apply the tool of categorical logic to further develop this inspiration.[112]

Of course Lawvere's remarks here are personal and idiosyncratic, but the inclination to oppose the traditional approaches to foundations as 'speculative' and replace them with something more amenable to mathematical practice is a common theme amongst algebraic structuralists influenced by category theory and categorical logic.[113] Like other forms of structuralism, algebraic structuralism holds that mathematics is, in the main, concerned with structures. In contrast to assertoric structuralism, however, algebraic structuralism does not seek a single theory of structures to account for the content of mathematics. Indeed, one way to describe algebraic structuralism is by highlighting its opposition to the idea that mathematics has a single determinate content, or that mathematics concerns itself with a global ontological background or framework.

So, if mathematics is not about a single universe, either of sets or even of abstract structures, what *is* it about? There are a variety of different technical proposals in response to this question,

---

[112] [Lawvere 2003, 213].
[113] Cf., e.g., [Awodey 1996,].

though most are framed, in some way, in category-theoretic terms. Thus, for instance, John Bell writes that

> The fundamental idea is to abandon the unique absolute universe of sets central to the orthodox set-theoretic account of the foundations of mathematics, replacing it by a plurality of local mathematical frameworks – *elementary toposes* – defined in category-theoretic terms. Such frameworks will serve as local surrogates for the classical universe of sets.[114]

Here we see the emphasis on local frameworks versus global foundations which is characteristic of the algebraic approach to structuralism. In place of a single theory of structures, the idea is to investigate a variety of different sorts of connections between different structures. There is not meant to be a final theory of structures, nor of the connections between them, but, instead, several distinct (perhaps incompatible, perhaps complementary) frameworks, each of which will highlight different sorts of connections, transformations, objects, etc.

On Bell's particular view, the most natural way of organizing these frameworks is in terms of category-theoretically defined 'elementary toposes'.[115] My interest here is not in the analysis of the precise details of topos and category theory which make them amenable to the structuralist programme,[116] but instead in the general philosophical attitude towards foundations which is exhibited by those who suggest category theory can act as the basis for algebraic structuralism. For Bell, these elementary toposes are not meant to function as a kind of universal background ontology, in the manner of Hellman's modal theory of possible systems or Shapiro's structure-theory. No, for Bell, these toposes themselves are 'ambient' and variable, in the sense that neither their mutual interrelations nor their internal compositions (*including* their governing logic) are fixed outside of a given mathematical situation.

An even more radical statement of the algebraic approach to structuralism can be found in a 2004 paper by Steve Awodey. There Awodey writes (in response to some of Hellman's criticisms of the category-theoretic approach to structuralism) that

> As opposed to this one-universe, 'global foundational' view, the 'categorical-structural' one we advocate is based instead on the idea of specifying, for a given theorem or theory only the required or relevant degree 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. The laws, rules, and axioms involved in a particular piece of reasoning, or a field of mathematics, may vary from one to the next, or even from one mathematician or epoch to another. The statement of the inferential machinery involved thus becomes a (tacit) part of the mathematics; functional analysis makes heavy use of abstract functions and the axiom of choice, some theorems in algebra rely on the continuum hypothesis; many arguments in homology theory are purely

[114] [Bell 1986, 409].
[115] Bell gives a formal definition of ' topos' at [Bell 1986, 412-413]. For a semi-technical elaboration on some of the interesting relations between toposes (or topoi), category theory, and set theory, see [Bell 1982].
[116] This is already accomplished in [Awodey 1996] and [McLarty 2004].

> algebraic, once given the non-algebraic objects that they deal with; theorems in constructive analysis avoid impredicative constructions; nineteenth-century analysis employed other methods than modern-day analysis, and so on. The methods of reasoning involved in different parts of mathematics are not 'global' and uniform across fields or even between different theorems, but are themselves 'local' or relative. Thus according to our view, there is neither a once-and-for-all universe of all mathematical objects, nor a once-and-for-all system of all mathematical inferences. Are there, then, various and changing universes and systems? How are they determined, and how are they related? Here I would rather say that there are *no* such universes or systems; or rather, that the question itself is still based on a 'foundationalist' preconception about the nature of mathematical statements.[117]

Let's try to unpack this a bit. For Awodey, the algebraic (or categorical) approach to mathematics involves accepting (embracing, really) certain limitations on one's mathematical activity. Instead of presuming to grasp the nature of the set-theoretic universe when proving a theorem about Abelian groups, for instance, his view suggests that in any given mathematical situation, we ought only to consider those 'objects' and bits of 'inferential machinery' which have actually been employed. On this view, we might say, a particular 'structure' is not defined against anything other than the present mathematical situation: we form the structure out of the objects, properties, relations, and logic which we actually employ in the situation at hand and that is the end of the matter as far as ontology and reference-fixing goes. We state as much (and only as much) as we need to about the structures we are presently studying. From this point of view, the problem with the foundationalist approach to mathematics (within which both Awodey and myself would include Shapiro's and Hellman's assertoric forms of structuralism) is that its metaphysical principles require there to be some unifying element which collects together all forms of mathematical reasoning in a single universe (with its own particular collection of objects, relations, properties, and its own governing logic). The assertoric structuralist's complaint is that "these constructions have to take place somewhere! They require some collection principles"[118] and some logic governing these principles. Ultimately, for Awodey, the assertoric structuralist's idea

> that one is 'going up' in a hierarchy, and that this requires stronger and stronger collection principles and existence assumptions rests on the 'foundationalist' conception that the 'objects' involved are fixed and determinate. From a categorical perspective, one is rather 'going down' by specifying more of the ambient structure to be taken into account […].[119]

This contrast presented here gets at the heart of the algebraic structuralist view. Central to this view is the contrast between determinate collections and ambient categories or structures. What is the difference, exactly?

<u>2.4.4 Bottom-up vs. top-down</u>

---

[117] [Awodey 2004, 56].
[118] [Awodey 2004, 62].
[119] [Awodey 2004, 62].

The bottom-up approach is common in the set-theoretic construction of various number domains or algebraic structures. You begin with a very limited ontology—say the empty set or a finite collection of *ur-elemente*. From here, by applications of various 'collection principles' (e.g., the power set axiom, the axiom schema of separation, etc.) you gain access to larger and larger collections, within which you can define more and more structures. This intuitive process is the reason why the set-theoretic universe is often called the iterative or cumulative hierarchy. Importantly, on this view, one's access to a particular structure is dependent upon there being some correlate to that structure in this cumulative hierarchy. So, if we are interested in investigating groups, we must first 'build' them from the bottom up using the elements of the set-theoretic hierarchy in some way. Everything thus remains within the absolute ontology of set theory itself. Nothing in this is necessarily peculiar to *set* theory: we can approach other aspects of mathematics using this foundational attitude as well. Indeed, as we saw, Shapiro's approach to structure theory is bottom-up in exactly this way: he believes that, for his theory to get off the ground, he needs the basic building blocks which will guarantee that the more complex structures exist. If we wish to speak of categories, the category of sets, as understood by a traditional set-theoretic foundationalist, must be fixed, i.e., made non-algebraic,[120] remaining the same for everyone in every situation. In this sense it is anything but ambient. From this determinate category, all the rest are built up. This is the 'bottom-up' approach.

Awodey and Bell, by contrast, wish to take a 'top down' approach which does not even begin with a basic, universal ontology. And, in fact, the more usual approach to category theory is non-foundational in precisely the required sense. Technical definitions of 'category' (of which there are many) are not terribly important in the present context, what is important is the attitude taken towards such definitions. Thus, after giving his version of the category axioms, Awodey writes

> A category is *anything* satisfying these axioms. The objects need not have 'elements', nor need the morphisms be 'functions' although this is the case in some motivating examples [i.e., in the category **Set**]. But also, for example, associated with any formal system of logic is a category, the objects of which are formulas and the morphisms of which are deductions from premises. We do not really care what non-categorical properties the objects and morphisms of a given category may have; that is to say, we view it 'abstractly' by restricting to the language of objects and morphisms, domains and codomains, composition, and identity morphisms.[121]

Now, fairly obviously, this sort of language of abstraction closely mirrors Shapiro's (initial) approximation of the structuralist position. The difference between the two—and it is a crucial difference—is that, for the category-theoretic structuralist, such a language does not and should not be embedded within a single universe of discourse, a single logic, or a single ontology. Or, to

---

[120] This, in broad outline, is precisely the sort of claim that Hellman makes about category-theoretic approaches to structuralism: "As its proponents have maintained, category theory does offer an interesting structuralist perspective on most mathematics as we know it. But it needs to be supplemented and set within a yet wider framework. […] [C]ategory theory—with its non-assertory, algebraico-structural axioms—depends on prior notions of *structure* (collection with relations) and *satisfaction by structures* to make sense of the very notions of 'category' and 'topos'" [Hellman 2003, 154-155]. The objection clearly mirrors the problem of quasi-concreteness in relation to Shapiro's views. Similar claims are made in an earlier paper by Solomon Feferman (i.e., [Feferman 1977]).
[121] [Awodey 1996, 213].

put it in terms familiar from the preceding section, the algebraic structuralist is not at all convinced of the necessity of a background ontology.

The above examples of a category of sets and functions vs. a category codifying a system of formal logic are themselves illustrative of a key difference between the algebraic approach and the assertoric approach. On Shapiro's assertoric approach, he became embroiled in the problem of the quasi-concrete because his (non-structuralist) understanding of the logic animating his project seemed to commit him to sets. For the category theorist, this troubling situation might be easily resolved. On the one hand we have a category (let's just assume, for the sake of our example) that it is the category **Set**. The objects of this category are sets, and the morphisms are functions on those sets. The problem Shapiro has is that his logic seems to be understandable only in terms of this category of sets. But, from the category-theoretic perspective, here we have an interesting example of morphisms between categories. In addition to set we can define another category, let's say **ShLog** (for Shapiro-Logic), whose objects are the formulas of Shapiro's structure theory and whose morphisms are the allowable deductions among them. On this view, we needn't conflate the elements of **Set** and **ShLog**, but can instead describe the morphisms between them (perhaps there is a relation of isomorphism, as Shapiro seems to worry) as the elements of yet another category with a distinct structure. But we needn't take this to be revelatory of some secret or hidden quasi-concrete ontology in our original approach. Instead, it reveals new and interesting *structure* which emerges in this particular situation, or within this particular local framework.

Perhaps the best way of characterizing the difference between this bottom-up approach and the top-down approach recommended by Awodey is to look at the components of each view which are allowed to vary. On the bottom-up approach, there are at least some parts which do not (and cannot) vary. These are the background ontology (usually some universe of sets) and the logic which governs the theory (whatever this might be). No matter what context you work in, if you build your structures out of a foundationally understood theory, then its ontology and logic remain static in every context. The key, on this approach, is to craft a subtle enough theory so that it covers all the desired instances without committing you to anything 'extra'. On the algebraic perspective, by contrast, both of these universally invariant things are allowed to vary; indeed, in Bell's proposal, their variability is an essential component of the notion of an elementary topos.

## 2.4.5 No background ontology? No problem.

As we saw in our examination of Shapiro's work, assertoric structuralism requires some background ontology in order to get off the ground. This background ontology was required in order to ensure that there were structures big enough to satisfy the usual mathematical demands. We also saw that commitment to such a background ontology involved Shapiro (and others) in an ultimately non-structural account of mathematics—exactly the thing he sought to avoid. The question we have to ask now is whether or not algebraic structuralism can really do without a commitment to such a background ontology.

Here is Awodey's statement on this question:

> the structural perspective [of topos theory] is at odds with the idea that all mathematical objects exist in a single, comprehensive universe of sets. The methods of topoi make it more natural to study the 'local' logic of various kinds of structure than to force all kinds of structure into a 'global' logic, say of the category of sets. Moreover, the very idea of 'foundations of mathematics' becomes less significant from a categorical perspective than, say, organizing and unifying the language and methods of mathematics. Such unification often points up related structural phenomena in disparate fields, suggesting fruitful directions for further research.[122]

So the clear answer to the question 'does algebraic structuralism require a single background ontology' is, for Awodey, a straightforward and emphatic 'no'. The point of the approach is to examine structures *in situ*, not to displace them and characterize them in the (usually) foreign terms of object-oriented set theory. Moreover, Awodey's remarks here suggest that, on topos-theoretic approach to 'foundations,' the troublesome connection to a non-structural account of logic (which Shapiro himself struggled with in relation to his 'quasi-concrete' second order language) simply disappears. Logic itself is understood as a potentially variable component of a given topos. It is not pan-theoretically applicable, but a matter of specification in each case.

The position as a whole is summarized by Awodey as follows:

> The structural perspective on mathematics codified by categorical methods might be summarized in the slogan: The subject matter of pure mathematics is invariant form, not a universe of mathematical objects consisting of logical atoms. This trivialization points to what may ultimately be an insight into the nature of mathematics.[123]

The important point here, the point which Shapiro himself tried to get at in *his* form of structuralism, is that, for the algebraic structuralist, there is not and need not be any final accounting of 'invariant form'. The property of invariance is itself defined relative to a particular structure (i.e., a particular local framework or topos). There may be interesting relationships between different frameworks, which lead to the discussion of new and interesting sorts of invariant form (as happened, for instance, in the birth of category theory).[124] Neither do these emergent forms of invariance require any final theory or background ontology. For the topos-theoretic structuralist, ontology really is relative to a given framework. If we wish to remain 'foundationalists' we can seek to axiomatize a category of all categories,[125] one which will encompass every other category. Category theory certainly allows for this possibility, but it is, in a fairly obvious sense, counter to the spirit of the algebraic structuralists endeavour. Even if we do construct such a category of all categories, there is no reason we can't treat this category (and

---

[122] [Awodey 1996, 235].
[123] [Awodey 1996, 235].
[124] Category theory, in its earliest incarnations, arose from the recognition that there were shared bits of invariant form between algebra and topology. See, e.g., [Eilenberg and Mac Lane, 1942] and [Eilenberg and Mac Lane 1945].
[125] Cf., e.g., [Lawvere 1966].

a concomitant logical structure) as one more local framework among others. The aim of category theory is to allow us to move between frameworks, while making it as easy as possible to highlight intrastructural morphisms (functors) which allow us to discern new forms of invariance and to construct new and interesting local frameworks all over again.

2.4.6 Algebraic reference and truth

Of course this shift from an absolute perspective to a radically relativized one does not come without a cost. As we've seen in our examination of Frege, one of the key reasons why he sought fixed references for mathematical terms in the first place was to avoid the obvious relativization of truth which seemed to lie at the basis of Hilbert's view. Frege was deeply committed to the view that mathematics was a science, and that sciences had something to say about the world—i.e., the single world shared by logic, mathematics, physics, chemistry, sociology, and all the rest. For him, the notion that there could be relative truths was simply a non-starter, and resulted in a meaningless, subjective morass from which no scientific findings could emerge. This is, at least partially, an explanation for his somewhat surprising (and perhaps misguided) rejection of the possibility that both Euclidean and non-Euclidean geometry could be 'true'.[126] We saw, too, that Hilbert, in his later proof-theoretic years, was heavily invested in the production of *direct* consistency proofs, proofs which would stake direct claims about the world, and whose truth would be absolute.

So it should be no surprise then that the only way to maintain the fully algebraic version of structuralism is to *relativize* not only truth but meaning and reference as well. These implications were already present in Frege's debate with Hilbert and Korselt, but are spelled out more thoroughly by Bell in a 1986 paper entitled "From Absolute to Local Mathematics". There he writes

> With the relinquishment of the absolute universe of sets, mathematical concepts will in general no longer possess absolute meaning, nor mathematical assertions absolute truth values, but will instead possess such meanings or truth values only *locally*, i.e., *relative* to local frameworks.[127]

And that

> Category theory […] suggests that mathematical concepts and assertions should be regarded as possessing meaning only *in relation to* a variety of more or less *local* frameworks.[128]

These, framed in *positive* terms, are exactly the effects that Frege foresaw in Hilbert's turn of the century views on axiomatics. We are left, then, with the question of whether any significant advance has been made by algebraic structuralism toward answering our question 'what is mathematics about?'. Certainly great technical strides have been made since the Frege-Hilbert

---

[126] Cf. [Frege c. 1900, 169].
[127] [Bell 1986, 409].
[128] [Bell 1986, 410].

correspondence, such that the variability of both 'structure' and 'logic' in a topos is possible in ways which were not even imagined at the close of the nineteenth century.

## 2.5 Conclusion

So, in the end, where do these two forms of structuralism leave us? Are we better off in our general understanding of the content of mathematics now than we were after the Frege-Hilbert debate?

The goal of both the assertoric and algebraic forms of structuralism, as I have presented them here, is to present a clear and persuasive account of mathematical structures. Given such an account, we would be able to generate a satisfactory answer to our guiding question, 'what is mathematics about?'. The answer would be that mathematics is about *structures* and these are to be understood in the way dictated by this particular account.

I think it is clear that there is no victor in the competition between the assertoric and algebraic forms of structuralism in terms of their respective accounts of structure. If we accept the assertoric account, we are pushed into a Frege-style dilemma which forces us to pick out our structures from a single background ontology. While the problems of assertoric structuralism are not identical to Frege's Caesar problem, and certainly need not lead to Russell-style contradictions, from a general philosophical point of view, they are unsatisfying in precisely the same way. Assertoric structuralism, like Frege, claims that mathematics has a specific content (in this case, some theory of structure) but it cannot give us good reasons for picking out this specific content from a variety of equally plausible options. Nor, contra Shapiro, can it rest satisfied with the possibility of a 'shared structure' between these accounts because account of structure given is itself bound to a specific ontology. None of this ought to suggest that the assertoric form of structuralism is a fruitless endeavour[129]—only that it cannot provide an ultimately satisfactory answer to *our* particular question.

Does algebraic structuralism fare any better? Like the assertoric structuralist, the algebraic structuralist also holds that mathematics is chiefly concerned with structure. Unlike the assertoric structuralist, however, the algebraic structuralist does not attempt to provide a single or final theory which pins down the nature of structures. Instead, following the lead of category theory, such structures are only ever determined locally. This has the benefit that the non-uniqueness problems which plagued Frege, set-theoretic foundationalism, and even assertoric structuralism are rather understood to be 'features' of the view. But it also has the disquieting effect of introducing (or, perhaps, accepting an already-present) radical relativity into mathematics.

The idea of a single ontology for mathematics is of course replaced with multiple distinct ontologies, defined relative to a given local framework or topos. This sort of view is already implicit in Benacerraf's work, and is accepted as a starting point for the algebraic approach. Alongside the relativization of ontology, however, we also have the more troubling relativization

---

[129] A feature of assertoric structuralism which has not at all been stressed in this chapter is its capacity to address the deep epistemological question in the philosophy of mathematics. Resnik, in particular, has made significant advances in these areas (cf., e.g., [Resnik 1999]).

of meaning, reference, and truth. Claims made within a local framework have no direct relationship with 'the world' understood as a unified, final, or absolute ontological horizon. Frege, already in his correspondence with Hilbert, worried that the disconnection of axiomatics from any ontological background would have the effect of relativizing truth. The algebraic structuralist simply accepts this fact as an essential feature of mathematical discourse. Similarly, the idea that mathematical statements can be made meaningful only by their reference to a universal background ontology is rejected in favour of the view that both reference and meaning are relative to a given local framework. This fact is both supported by and helps to explain things like the existence of models of set theory within which the axiom of choice holds alongside models within which it does not hold. The implications of this kind of view are wide-reaching and essentially undermine any attempt to understand mathematics as a unified investigation of a single universe.

Does the position of the algebraic structuralist offer us a significant advance beyond the impasse of the Frege-Hilbert debate? It seems that here, too, we have a recapitulation of an already-familiar position. In the assertoric structuralist's case, the position was similar in the relevant respects to Frege's fixed-ontology view. In the algebraic case, the position closely resembles that of Hilbert (at least the Hilbert around the time of the publication of his *Festschrift*). For, recall Hilbert's position in his correspondence with Frege.

His account of mathematical existence seems to be not much different than the ontological relativity presented by the algebraic structuralist. He writes that "if the *arbitrarily* given axioms do not contradict one another with all their consequences, then they are true and the things defined by the axioms exist. This is for me the criterion of truth and existence."[130] Of course Hilbert would later back away from this kind of relativity in his demand for *absolute* consistency proofs, but in this period he was for the most part happy to understand existence in terms of the consistency of a particular theory, and consistency in (admittedly) vague but not clearly absolutist terms. Certainly the category-theoretic/algebraic structuralist approach no longer relies uncritically on a theory-independent notion of consistency for its ontological relativity. But it seems reasonable to consider the newer position a technical refinement and expansion of Hilbert's attitude.

Here, too, we see that, already at the close of the nineteenth century, Hilbert has accepted the fact that his purely relational (we might now say algebraic) approach to axiomatics commits him to the relativity of truth to a given formal theory. Without the definition or context provided by a given theory, there simply is no criterion for truth. The provision of the axiomatic theory (or the description of a structure) gives you the relative ontology, and it is only on the basis of this relative ontology that one can make (relatively) meaningful and truthful claims. Without such a context, there is no obvious way for Hilbert or the algebraic structuralist to speak of meaning, reference, or truth.

In sum, then, it seems that the contemporary algebraic structuralist has been engaged in a project of refinement and increasing precision: the notions of category theory and categorical logic make more perspicuous the relativity of truth and reference, and they allow us to treat relations

[130] [Hilbert to Frege, 29 December 1899; excerpted by Frege in Frege 1980, 39-40].

between various frameworks of truth and reference in a variety of systematic ways. These procedures were in their infancy in Hilbert's day, and it was, somewhat ironically, only through the development of formal logic (with the resources of model- and set-theory) which allowed them their fully algebraic form in contemporary approaches to structuralism.

In any case, it seems that we are again left at a kind of impasse. Or, rather, we are left with two possible choices, both of which have significant implications for our understanding of the nature of mathematics. On the one hand, we can demand with Frege and the assertoric structuralists that mathematics possesses a unified content, and that we give a unique account of this content. This view faces the seemingly intractable non-uniqueness problem, but it has the benefit of maintaining non-relative conceptions of ontology, reference, and truth. On the other hand, we can insist that there is no final accounting of the content of mathematics, that there are different forms of mathematics with interesting relations with one another, but there is no final box within which they all live and breathe. This has the benefit of avoiding the non-uniqueness problem, but has the profound and far-reaching implication that mathematics does not speak directly about a single world, and that the truths of mathematics (and perhaps the rest of science as well) are merely relative to a given situation.

How are we to choose between the two options? As I hope to have shown, neither view straightforwardly wins the day, as each position requires the acceptance of something which the other side imagines to be an obvious compromise or even a failing (e.g., ontological fixity, the relativity of truth). There seems to be no ultimately convincing argument to sway us either way. Either we believe in the non-relativity of truth, and reference, or we do not. The arguments that each side gives have most often been met with a shrug of the shoulders: either you demand a foundation for mathematics or you do not. A good example of this new impasse between forms of structuralism occurs in a debate between Hellman and Awodey regarding the possible usefulness of category theory for structuralism. Hellman suggests, following his assertoric scruples, that category theory requires "a background logic of relations and substantive assumptions addressing mathematical existence of categories themselves"[131] That is, whatever approach is taken, there must be a final reckoning regarding the single logic (and, relatedly, the background ontology) which is taken up by the category-theoretic structuralist. Awodey, the arch-algebraist, responds to this "problem of the home address"[132] by suggesting that it

> asks for something that only seems reasonable from a foundational perspective. This requirement exemplifies a general pattern in the discussion: Hellman sympathizes with the structural viewpoint […] but he thinks that there is still something missing in the overall category-theoretical position. The something is to be provided by his modal-structuralism. But I contend that what is missing is only a correct understanding of the categorical approach.[133]

[131] [Hellman 2003, 157].
[132] [Awodey 2004, 55].
[133] [Awodey 2004, 55].

In effect, both Hellman and Awodey demand something that their opponent cannot give: Hellman cannot picture mathematics without foundations, and Awodey cannot understand why mathematics needs foundations at all. The two are at cross-purposes, and there does not seem to be a satisfying reason to choose one or the other. I do not see that this is a significant advance over the impasse of the Frege-Hilbert debate, though perhaps the battle lines have begun to be drawn more clearly when phrased in the terms of contemporary structuralism.

# Bibliography

Akiba, Ken. (2000). "Indefiniteness of Mathematical Objects," *Philosophia Mathematica* **3**, volume 3: 26-46.

Aquinas, Thomas. (1957). *Summa Theologiae*, 5 volumes (Ottawa: Commissio Piana).

Aristotle. (1971). *On the Heavens*, translated by W. K. C. Guthrie (Cambridge: Harvard University Press).

Aristotle. (1994). *Posterior Analytics*, translated with a commentary by Jonathan Barnes (Oxford: Clarendon Press).

Awodey, Steve. (1996). "Structure in Mathematics and Logic: A Categorical Perspective," *Philosophia Mathematica* (III) **4**: 209-237.

Awodey, Steve and Reck, Erich H. (2002). "Completeness and Categoricity. Part I: Nineteenth-century Axiomatics to Twentieth-century Metalogic," *History and Philosophy of Logic* **23**: 1-39.

Awodey, Steve. (2004). "An Answer to Hellman's Question: 'Does Category Theory Provide a Framework for Mathematical Structuralism," *Philosophia Mathematica* (III) **12**: 54-64.

Balaguer, Mark. (1998). "Non-uniqueness as a non-problem," *Philosophia Mathematica* **3**, volume 6: 63-84.

Beall, J. C. and Restall, Greg. (2000). "Logical Pluralism," *Australasian Journal of Philosophy* **78**: 475-493.

Beall, J. C. and Restall, Greg. (2006). *Logical Pluralism* (Oxford: Oxford University Press).

Beaney, Michael. (1997). *The Frege Reader* (Oxford: Blackwell).

Bell, John. (1982). "Categories, Toposes and Sets," *Synthese* **51**, no. 3: 293-337.

Bell, John. (1986). "From Absolute to Local Mathematics," *Synthese* **69**: 409-426.

Bell, John. (2001). "Observations on Category Theory," *Axiomathes* **12**, issue 1-2: 151-155.

Beltrami, Eugenio. (1868a). "Saggio di interpretazione della geometria non-euclidea," *Giornale di Mathematiche* **VI**: 285–315.

Beltrami, Eugenio. (1868b). "Teoria fondamentale degli spazii di curvatura costante," *Annali. Di Mat.* (*series II*), **2**: 232–255.

Benacerraf, Paul. (1965). "What Numbers Could Not Be," *The Philosophical Review* **74**, No. 1: 47-73.

Benacerraf, Paul (1973). "Mathematical Truth," *Journal of Philosophy* **70**: 661-680.

Benacerraf, Paul. (1996a). "Recantation, or, Any old $\omega$-sequence would do after all," *Philosophia Mathematica* **3**, vol. 4: 184-189.

Benacerraf, Paul. (1996b). "What mathematical truth could not be – I," in A. Morton and S. P. Stich (editors), *Benacerraf and His Critics* (Oxford: Blackwell): 9-59.

Bergmann, Gustav. (1958). "Frege's Hidden Nominalism," *The Philosophical Review* **67**, no. 4: 437-459.

Berkeley, George (1734). *The Analyst* (London).

Bernays, Paul. (1930-1931). "Die Philosophie der Mathematik und die Hilbertsche Beweistheorie," *Blätter für deutsche Philosophie* **4**: 326-67; English translation available in [Mancosu 1998]: 234-265. References are to the translation.

Biermann, K.-R. (1973). *Die Mathematik und ihre Dozenten an der Berliner Universität, 1810-1920* (Berlin: Akademie Verlag).

Blanchette, Patricia. (1996). "Frege and Hilbert on Consistency," *The Journal of Philosophy*, volume XCIII, no. **7**: 317-336.

Blancette, Patricia. (2012). *Frege's Conception of Logic* (Oxford: Oxford University Press).

Bolzano, Bernard. (1817). *Rein analytischer Beweiss des Lehrsatzes, dass zwischen je zwey Werthen, die ein engegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege, (Prague).* translated into English by Steve Russ as "Purely analytic proof of the theorem that between any two values, which give results of opposite sign, there lies at least one real Root of the Equation" by Steve Russ in Bernard Bolzano, *The Mathematical Works of Bernard Bolzano* (Oxford: Oxford University Press, 2004): 251-277.

Bolzano, Bernard. (1837). *Wissenschaftslehre* (Sulzbach).

Bolzano, Bernard. (1930). *Functionenlehre*, edited by K. Rychlik (Prague: Royal Bohemian Academy of Sciences); translated into English by Steve Russ as "Theory of Functions" in Bernard Bolzano, *The Mathematical Works of Bernard Bolzano* (Oxford: Oxford University Press, 2004): 427-572.

Bolzano, Bernard. (2004a). *The Mathematical Works of Bernard Bolzano*, translated by Steve Russ (Oxford: Oxford University Press).

Bolzano, Bernard. (2004b). "On the Mathematical Method," in Bernard Bolzano, *On the Mathematical Method and Correspondence with Exner*, translated by Rolf George and Paul Rusnock (Amsterdam: Rodopi).

Bolzano, Bernard. (1975). *Einleitung zur Grössenlehre. Erste Begriffe der allgemeinen Grössenlehre*, edited by Jan Berg (Stuttgart-Bad Cannstatt: Frommann-Holzboog), volume IIA of the Bernard-Bolzano-Gesamtausgabe; translated into English by Paul Rusnock and Rolf George as "On the Mathematical Method" in *On the Mathematical Method and Correspondence with Exner* (Amsterdam: Rodopi, 2004): 39-82.

Bonola, Roberto. (1955). *Non-Euclidean Geometry: A critical and historical study of its developments*, translated by H. S. Carslaw with an introduction by Federigo Enriques (New York: Dover).

Boole, George. (1854*). An Investigation of the Laws of Thought, on which are Founded the Mathematical Theories of Logic and Probabilities* (Cambridge: MacMillan).

Boolos, George. (1987a). "Saving Frege from Contradiction," *Proceedings of the Aristotelian Society*, New Series, **87**: 137-151.

Boolos, George. (1987b). "The Consistency of Frege's *Foundations of Arithmetic*," in J. J. Thomson (editor), *On Being and Saying: Essays for Richard Cartwright* (Cambridge: MIT Press): 3–20.

Boolos, George. (1993). "Basic Law (V)," *Proceedings of the Aristotelian Society, Supplementary Volumes* **67**: 213-233.

Boolos, George and Heck, Richard. (1998). "*Die Grundlagen der Arithmetik* §§82-83," in Matthias Schirn (editor) *Philosophy of Mathematics Today* (Oxford: Oxford University Press): 315-338.

Bourbaki, Nicholas. (1939). *Éléments de Mathématiques, Volume 1: Théorie des Ensembles, Fascicule de Résultats* (Paris: Hermann).

Burgess, John P. (1997). "Frege and Arbitrary Functions," in William Demopoulos (editor), *Frege's Philosophy of Mathematics* (Cambridge: Harvard University Press): 89-107.

Carnap, Rudolf (1950). "Empricism, Semantics, and Ontology," in *Revue Internationale de Philosophie* **4**: 20-40.

Carnap, Rudolf. (2002). *The Logical Syntax of Language*, translated by Amethe Smeaton, Countess von Zeppelin (London: Open Court).

Cauchy, Augustin-Louis. (1823). "Septième Leçon: Valeurs de quelques expressions qui se présentent sous les formes indéterminées $\infty/\infty$, $\infty^0$, ... . Relation qui existe entre le rapport aux différences finies et la fonction derivée," lecture in *Résumé des leçons données à l'École royale* *polytechnique sur le calcul infinitesimal*; reprinted in Augustin-Louis Cauchy (1882-1974), *Oeuvres Complètes,* série 2, tome 4 (Paris: Gauthier-Villars et fils): 42-46.

Cayley, Arthur. (1859). "A Sixth Memoir upon Quantics," *Philosophical Transactions of the Royal Society of London* **149**: 61-90.

Coffa, J. Alberto. (1986). "From Geometry to Tolerance: Sources of Conventionalism in Nineteenth-Century Geometry," in Robert G. Colodny (editor), *From Quarks to Quasars: Philosophical Problems in Modern Physics* (Pittsburgh: Pittsburg University Press): 3-71.

Coffa, J. Alberto. (1991). *The Semantic Tradition from Kant to Carnap: To the Vienna Station*, edited by Linda Wessels (Cambridge: Cambridge University Press).

Cohen, Paul. (1963) "The independence of the continuum hypothesis," *Proceedings of the National Academy of Sciences of the United States of America* **50**, no. 6: 1143-1148.

Cohen, Paul. (1964). "The independence of the continuum hypothesis, II," *Proceedings of the National Academy of Sciences of the United States of America* **51**, no. 1: 105-110.

Corry, Leo. (2004a). *David Hilbert and the Axiomatization of Physics (1898-1918): From* Grundlagen der Geometrie *to* Grundlagen der Physik (Dordrecht: Kluwer).

Corry, Leo. (2004b). *Modern Algebra and the Rise of Mathematical Structures* (Basel: Birkhäuser).

Corry, Leo. (2006). "Axiomatics, Empiricism, and Anschauung in Hilbert's Conception of Geometry: Between Arithmetic and General Relativity," in *The Architecture of Modern Mathematics: Essays in History and Philosophy*, edited by J. Ferreirós and J. J. Gray (Oxford: Oxford University Press): 133-156.

d'Alembert, Jean le Rond. (1759-1767). "Essai sur les Élémens de Philosophie; ou sur les principes des connoissances humaines; avec les éclaircissemens," reprinted in volume 2 of *Oeuvres Philosophiques, Historiques, et Littéraires de d'Alembert*, 18 volumes, edited by J. F. Bastien (Paris: 1805).

Dehn. Max. (1900). "Die Legendreschen Sätze über die Winkelsumme in Dreieck"*Mathematische Annalen* **53**: 404-439.

De Risi, Vincenzo. (2007). *Geometry and Monadology: Leibniz's* Analysis Situs *and Philosophy of Space* (New York: Springer).

Deskins, W. E. (1995). *Abstract Algebra* (Mineola, New York: Dover).

Diamond, Cora. (1991). *The Realistic Spirit: Wittgenstein, Philosophy, and the Mind* (Cambridge: MIT Press).

Dieudonné, Jean. (1972). "The Historical Development of Algebraic Geometry," *The American Mathematical Monthly* **79**, no. 8: 827-866.

Dummett, Michael. (1973). *Frege: Philosophy of Language* (Harper & Row: New York).

Dummett, Michael. (1982). "Objectivity and reality in Lotze and Frege," *Inquiry: An Interdisciplinary Journal of Philosophy* **25**, no. 1: 95-114.

Dummett, Michael. (1991). *Frege: Philosophy of Mathematics* (Cambridge: Harvard University Press).

Dunnington, Waldo G. (1955). *Carl Friedrich Gauss: Titan of Science, A Study of his Life and Work* (Exposition Press: New York).

Duns Scotus, John. (1950). *Opera Omnia*, volume 1, edited by C. Balić, M. Bodewig, S. Bušelic, P. Čapkun-Delić, I. Jurić, I. Montalverne, S. Nanni, B. Pergamo, F. Prezioso, I. Reinhold, and O. Schäfer (Vatican City: Typis Polyglottis Vaticanis).

Eilbenberg, Samuel and Mac Lane, Saunders. (1942). "Group Extensions and Homology," *Annals of Mathematics* **43**: 757-831.

Eilenberg, Samuel and Mac Lane, Saunders. (1945). "General Theory of Natural Equivalences," *Transactions of the American Mathematical Society* **58**: 231-294.

Einstein, Albert. (1916). "The Generalised Principle of Relativity," in Einstein, Albert and Minkowski, Hermann, *The Principle of Relativity: Original Papers*, translated by M. N. Saha and S. N. Bose (Calcutta: University of Calcutta Press, 1920): 89-163.

Einstein, Albert. (1921). "Geometry and Experience," in Michael Janssen, Robert Schulmann, József Illy, Chistopher Lerner, and Diana Kormos Buchwald (editors), *The Collected Papers of Albert Einstein, Volume 7: The Berlin Years: Wrigings 1918-1921* (Princeton: Princeton University Press, 2002): 382-405.

Ewald, William, editor. (1999). *From Kant to Hilbert: A Source Book in the Foundations of Mathematics* (2 vols.), (Oxford: Oxford University Press).

Feferman, Solomon. (1977). "Categorical Foundations and Foundations of Category Theory," in R. E. Butts and J. Hintikka (editors), *Logic, Foundations of Mathematics, and Computability Theory* (Dordrecht: Reidel): 149-169.

Feferman, Solomon. (1992). "Why a little bit goes a long way: Logical foundations of scientifically applicable mathematics," *PSA 1992*, vol. 2: 442-445.

Feferman, Solomon. (2008). "Lieber Herr Bernays!, Lieber Herr Gödel! Gödel on finitism, constructivity, and Hilbert's program," *Dialectica* **62**: 179-203.

Ferreirós, José. (1999). *Labyrinth of Thought. A history of set theory and its role in modern mathematics* (Basel: Birkhäuser).

Ferreirós, José. (2008). "The Crisis in the Foundations of Mathematics," in T. Gowers (editor), *Princeton Companion to Mathematics* (Princeton: Princeton University Press).

Field, Hartry. (1980). *Science Without Numbers: A Defence of Nominalism* (Princeton: Princeton University Press).

Frege. Gottlob. (1873). *Über eine geometriscche Darstellung der imaginären Gebilde in der Ebene* (Jena: A. Neuenhann); translated into English as *On a Geometrical Representation of the Imaginary Forms in the Plane* in Brian McGuinness, editor, *Collected Papers on Mathematics, Logic, and Philosophy* (Oxford: Basil Blackwell, 1984).

Frege, Gottlob. (1877). "Rezension von A. v. Gall and E. Winter, *Die analytische Geometrie des Punktes und der Geraden und ihre Anwendung auf Aufgaben,*" *Jenaer Literaturzeitung* **4**, number 9: 133–134; translated into English by H. Kaal as "Review of A. v. Gall and E. Winter, *Analytic Geometry of the Point and the Line and its Application to Problems*" in Brian McGuinness, editor, *Collected Papers on Mathematics, Logic, and Philosophy* (Oxford: Basil Blackwell, 1984): 96-97.

Frege, Gottlob. (1878). "Über eine Weise, die Gestalt eines Dreiecks als komplexe Grösse aufzufassen", *Jenaische Zeitschrift für Naturwissenschft* **12** (Supplement): xviii; translated into English by H. Kaal as "On a Way of Conceiving the Shape of a Triangle as a Complex Quantity," in Brian McGuinness, editor, *Collected Papers on Mathematics, Logic, and Philosophy* (Oxford: Basil Blackwell, 1984): 99-100.

Frege, Gottlob. (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens (Halle: Louis Nebert); translated by Terrell Ward Bynum into English as "Conceptual Notation," in Conceptual Notation and Related Articles (Oxford: Oxford University Press, 1972): 101-203.*

Frege, Gottlob. (?1879-1891). "Logic," in *Posthumous Writings*, edited by Hans Hermes, Friedrich Kambartel, and Friedrich Kaulbach; translated by Peter Long and Roger White with the assistance of Raymond Hargreaves (Oxford: Basil Blackwell, 1979): 1-8.

Frege, Gottlob. (1880). "Rezension von: Hoppe, *Lehrbuch der analytischen Geometrie I*," *Deutsche Literaturzeitung* **1** (1880): 210–211; translated by H. Kaal into English as "Review of Hoppe, *Textbook of Analytic Geometry I*," in Brian McGuinness, editor, *Collected Papers on Mathematics, Logic, and Philosophy* (Oxford: Basil Blackwell, 1984):101-102.

Frege. Gottlob. (1882). "Über die wissenschaftliche Berechtigung einer Begriffsschrift," *Zeitschrift für Philosophie und philosophische Kritik* **81**: 48-56; translated into English by Terrell Ward Bynum as "On the Scientific Justification of a Conceptual Notation," in *Conceptual Notation and Related Articles* (Oxford: Oxford University Press, 1972): 83-89.

Frege, Gottlob. (1884a). “Geometrie der Punktpaare in der Ebene,” *Jenaische Zeitschrift für Naturwissenschaft*, **17** (Supplement); translated into English by H. Kaal as “Geometry of Pairs of Points in the Plane” in Brian McGuinness, editor, *Collected Papers on Mathematics, Logic, and Philosophy* (Oxford: Basil Blackwell, 1984): 103–107.

Frege, Gottlob. (1884b). *Die Grundlagen der Arithmetik*: *eine logisch-mathematische Untersuchung über den Begriff der Zahl* (Breslau: W. Koehner); translated into English by Dale Jacquette as *The Foundations of Arithmetic: A Logical-Mathematical Investigation into the Concept of Number* (New York: Pearson Education, Inc., 2007).

Frege, Gottlob. (1891). *Funktion und Begriff* (Jena: Hermann Pohle), translated into English by Peter Geach as “Function and Concept,” in Michael Beaney, editor, *The Frege Reader* (Oxford: Blackwell, 1997): 130-148.

Frege. Gottlob. (1892a). “Ueber Begriff und Gegenstand,” *Vierteljahrsschrift für wissenschaftliche Philosophie*, **16**: 192-205, translated into English by Peter Geach as “On Concept and Object,” in Michael Beaney, editor, *The Frege Reader* (Oxford: Blackwell, 1997): 181-193.

Frege, Gottlob. (1892b) “Ueber Sinn und Bedeutung,” *Zeitschrift für Philosophie und philosophische Kritik*, **103**: 25-50; translated into English by Max Black as “On *Sinn* and *Bedeutung*,” in Michael Beaney, editor, *The Frege Reader* (Oxford: Blackwell, 1997): 151-171.

Frege, Gottlob. (1893). *Grundgesetze der Arithmetik*, volume I (Jena: Verlag Hermann Pohle); partially translated into English by Montgomery Furth as *The Basic Laws of Arithmetic: Exposition of the System* (Berkeley: University of California Press, 1964).

Frege, Gottlob. (1894). “Rezension von: E. Husserl, *Philosophie der Arithmetik I*,” in *Zeitschrift für Philosophie und philosophische Kritik*, **103:** 313–332.

Frege, Gottlob. (1897). “Logic,” in *Posthumous Writings*, edited by Hans Hermes, Friedrich Kambartel, and Friedrich Kaulbach; translated by Peter Long and Roger White with the assistance of Raymond Hargreaves (Oxford: Basil Blackwell, 1979): 126-151.

Frege, Gottlob (c.1900). “On Euclidean Geometry,” in *Posthumous Writings*, edited by Hans Hermes, Friedrich Kambartel, and Friedrich Kaulbach; translated by Peter Long and Roger White with the assistance of Raymond Hargreaves (Oxford: Basil Blackwell, 1979): 167-169.

Frege, Gottlob. (1903). “Über die Grundlagen der Geometrie,” (First series), *Jahresbbericht der Deutschen Mathematiker-Vereinigung* **12**: 319-324; 368-375.

Frege, Gottlob. (1906a). “On Schoenflies: *Die logischen Paradoxien der Mengenlehre,*” in Gottlob Frege (1979), *Posthumous Writings*, edited by Hans Hermes, Friedrich Kaulbach, and Friedrich Kambartel; translated by Peter Long and Roger White with the assistance of Raymond Hargreaves (Oxford: Basil Blackwell): 177-183.

Frege. Gottlob. (c. 1906b). “17 Key Sentences on Logic,” in Gottlob Frege (1979), *Posthumous Writings*, edited by Hans Hermes, Friedrich Kaulbach, and Friedrich Kambartel; translated by Peter Long and Roger White with the assistance of Raymond Hargreaves (Oxford: Basil Blackwell): 174-175.

Frege, Gottlob. (1914). “Logic in Mathematics,” in Gottlob Frege (1979), *Posthumous Writings*, edited by Hans Hermes, Friedrich Kaulbach, and Friedrich Kambartel; translated by Peter Long and Roger White with the assistance of Raymond Hargreaves (Oxford: Basil Blackwell): 203-250.

Frege, Gottlob. (1918). “Der Gedanke. Eine Logische Untersuchung,” in *Beiträge zur Philosophie und deutschen Idealismus* **I**: 143-157; translated into English by P. T. Geach and R. H. Stoothoff as “Thoughts” in *Logical Investigations* (Oxford: Basil Blackwell, 1977): 1-30.

Frege, Gottlob. (1970). *Translations from the Philosophical Writings of Gottlob Frege*, edited by Peter Geach and Max Black (Oxford: Basil Blackwell).

Frege, Gottlob. (1971). *On the Foundations of Geometry and Formal Theories of Arithmetic*, translated with an introduction by Eike-Henner W. Kluge (New Haven: Yale University Press).

Frege, Gottlob. (1980). *Philosophical and Mathematical Correspondence*, edited by Gottfried Gabriel, Hans Hermes, Friedrich Kambartel, Christian Thiel, and Albert Veraart; abridged from the German edition by Brian McGuinness; translated by Hans Kaal (Chicago: University of Chicago Press).

*Freudenthal, H. (1962). “The main trends in the foundations of geometry in the 19th century,” in E. Nagel, P. Suppes, and A. Tarski (editors) Logic, Methodology, and Philosophy of Science Proceedings of 1960 International Congress (Stanford: Stanford University Press): 618-620.*

Gauss, Carl Friedrich. (1863-1929). *Werke,* 12 volumes (Leipzig: Königliche Gessellscgaft der Wissenschaften).

Gauss, Carl Friedrich. (1831). “Anzeige der Theoreia residuorum biquadraticorum, Commentario secunda,” *Göttingsche gelehrte Anzeigen*, April 23, 1831; reprinted in [Gauss 1863-1929, II, 168-178; English translation in [Ewald 1999, 306-313].

Gentzen, Gerhard. (1936). “Die Widerspruchsfreiheit der reinen Zahlentheorie,” *Mathematische Annalen* **112**: 493-565.

Gödel, Kurt. (1931). “Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I,” *Monatshefte für Mathematik und Physik* **38**: 173-198; English translation in [Gödel 1986, pp. 144-195].

Gödel, Kurt. (1986). *Collected Works*, *Volume 1* (Oxford: Oxford University Press).

Goldfarb, Warren. (1997). "Metaphysics and Nonsense: On Cora Diamond's *The Realistic Spirit*," *Journal of Philosophical Research* **XXII**: 57-73.

Grattan-Guiness, Ivor. (2000). "The emergence of mathematical analysis and its foundational progress, 1780-1880," in Ivor Grattan-Guinness (editor), *From the Calculus to Set Theory 1630-1910: An Introductory History* (Princeton: Princeton University Press): 94-148.

Gray, Jeremy. (1989). *Ideas of Space: Euclidean, Non-Euclidean, and Relativistic* (Oxford: Clarendon Press).

Gray, Jeremy. (2007). *Worlds out of nothing: a course in the history of geometry in the 19th century* (London: Springer).

Guicciardini, Niccoló. (2009). *Isaac Newton on Mathematical Certainty and Method*. (Cambridge: MIT Press).

Hale, Bob and Wright, Crispin. (2001). *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics* (Oxford, Oxford University Press).

Halsted, George Bruce. (1894). "Non-Euclidean Geometry: Historical and Expository," *The American Mathematical Monthly* **1**, no. 5: 149-152.

Heath, Thomas L. (1921 [1981]). *A History of Greek Mathematics, Volume I: From Thales to Euclid,* (New York: Dover).

Heath, Thomas L. (1956). *The Thirteen Books of Euclid's Elements, 3 vols.* (2nd Edition), (New York: Dover).

Heck, Richard. (1995). "Frege's Principle," in Jaakko Hintikka (editor), *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics* (Dordrecht: Kluwer): 119-142.

Heck, Richard. (1996). "The Consistency of Predicative Fragments of Frege's *Grundgesetze der Arithmetik*," *History and Philosophy of Logic* **17**: 209–220.

Heck, Richard. (1997). "The Julius Caesar Objection," in Richard Heck (editor) *Language, Thought, and Logic: Essays in Honour of Michael Dummett* (Oxford: Oxford University Press): 273-308.

Heck, Richard. (1999). "*Grundgesetze der Arithmetik* I §10," *Philosophia Mathematica* **3**, vol. 7: 258-292.

Heck, Richard. (2005). "Julius Caesar and Basic Law V," *dialectica* **59**, no. 2: 161-178.

Heck, Richard G. and May, Robert. (2013). “The Function is Unsaturated,” in Michael Beaney, editor, *The Oxford Handbook of the History of Analytical Philosophy* (Oxford: Oxford University Press): 825-850.

Heidegger, Martin. (1977). “The Question Concerning Technology,” 3-35 in *The question concerning technology, and other essays*, translated by William Lovitt (New York: Garland Publishing).

Heine, E. H. (1870). “Über trigonometrische Reihen,” *Journal für die Reine und Angewandte Mathematik* **71:** 353–365.

Hellman, Geoffrey. (1989). *Mathematics Without Numbers: Towards a Modal-Structural Interpretation* (Oxford: Oxford University Press).

Hellman, Geoffrey. (1996). “Structuralism without Structures,” *Philosophia Mathematica* series 3, volume **4**: 100-123.

Hellman, Geoffrey. (2001). “Three Varieties of Mathematical Structuralism,” *Philosophia Mathematica* (III), volume **9**: 184-211.

Hellman, Geoffrey. (2003). “Does Category Theory Provide a Framework for Mathematical Structuralism?,” *Principia Mathematica* (III), volume **11**: 129-157.

von Helmholtz, Hermann. (1977). “On the Origin and Significance of the Axioms of Geometry,” in Hermann von Helmholtz, *Epistemological Writings*, edited with an introduction by R. S. Cohen and Y. Elkana, translated by Malcolm F. Lowe (Dordrecht: D. Reidel): 1-26.

Herbart, Johann Friedrich. (1824-1825). *Psychologie als Wissenschaft*, reprinted in Johann Friedrich Herbart, *Sämtliche Werke in chronologischer Reihenfolge herausgegeben von Karl Kehrbach und Otto Flügel*, 19 volumes (Aalen: Scientia Verlag, 1964): volumes 5 and 6, 177-402 and 1-339.

Hilbert, David. (1900a). “Mathematische Probleme. Vortrag, gehalten auf dem internationalen Mathematiker-Kongress zu Paris,” *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen*: 253-297; excerpted in English translation in [Ewald 1999]: 1096-1105.

Hilbert, David. (1900b). “Über den Zahlbegriff,” *Jahresbericht der Deutschen Mathematiker-Vereinigung* **8**: 180-194; English translation in [Ewald 1999]: 1089-1095.

Hilbert, David. (1903). *Grundlagen der Geometrie*, 2nd edition (B. G. Teubner: Leipzig).

Hilbert, David. (1919-20 [1992]). *Natur und Mathematisches Erkennen: Vorlesungen, gehalten 1919-1920 in Göttingen. Nach der Ausarbeitung von Paul Bernays*, edited with an introduction by David E. Rowe (Basel: Birkhäuser).

Hilbert, David. (1922). “Neubegrundund der Mathematik: Erste Mitteilung,” *Abhandlungen aus dem Seminar der Hamburgischen Universitat*, **1**: 157-177; English translation in William Ewald, (editor), *From Kant to Hilbert: A Source Book in the Foundations of Mathematics* (2 vols.), (Oxford: Oxford University Press, 1999): 1115-1134.

Hilbert, David. (1950). *The Foundations of Geometry*, translated by E. J. Townshend (La Salle, Illinois: Open Court).

Hilbert, David. (1971). *Foundations of Geometry*, translated by Leo Unger from the tenth German edition, revised and enlarged by Paul Bernays (La Salle, Illinois: Open Court).

Hilbert, David and Cohn-Vossen, S. (1999). *Geometry and the Imagination*, translated by P. Nemenyi (Providence: AMS Chelsea Publishing).

Horstein, Leon. (2014). “Philosophy of Mathematics,” in *The Stanford Encyclopedia of Philosophy* (URL: http://plato.stanford.edu/archives/fall2014/entires/philosophy-mathematics/). Accessed August 2014.

Horty, John. (2007). *Frege on Definitions: A Case Study of Semantic Definitions* (Oxford: Oxford University Press).

Hume, David. (1896). *A Treatise of Human Nature*. (Oxford: Clarendon Press).

Huntington, E. V. (1902). “Simplified Definition of a Group,” *Bulletin of the American Mathematical Society* **8**: 296-300.

Huntington, E. V. (1904). “Sets of Independent Postulates for the Algebra of Logic,” *Transactions of the American Mathematical Society* **5**, no. 3: 288-309.

Husserl, Edmund. (1891). “Philosophie der Arithmetik. Logische und Psychologische Untersuchungen,” in E. Husserl, *Philosophie der Arithmetik, edited* by L. Eley (The Hague: Nijhoff, 1970): 5–283.

Husserl, Edmund. (2001). *Logical Investigations: Volume one*, translated by J. N. Findlay with a preface by Michael Dummett (New York: Routledge).

Kant, Immanuel. (1747). “Thoughts on the true estimation of active forces,” excerpt translated into English by William Ewald, in William Ewald (editor), *From Kant to Hilbert: A Source Book in the foundations of Mathematics* (Oxford: Oxford University Press, 1996); 133-134.

Kant, Immanuel. (1787 [1996]). *Critique of Pure Reason*, translated by Werner S. Pluhar (Indianapolis: Hackett). ‘A’ and ‘B’ refer to the pagination of the 1781 and 1787 editions of the Critique, respectively.

Kant, Immanuel. (1950). *Prolegomena to Any Future Metaphysics*, translated by Lewis White Beck (New York: Bobbs-Merrill).

Kant, Immanuel. (1992). "The Dohna-Wundlacken Logic," in *Lectures on Logic* (Cambridge: Cambridge University Press), edited by Michael Young.

Kant, Immanuel. (1999). *Correspondence*, translated and edited by Arnulf Zweig (Cambridge: Cambridge University Press).

Ketland, Jeffrey. (2006). "Structuralism and the identity of indiscernibles," *Analysis* **66**: 303-315.

Khayyam, Omar. (1077). *Risala fi sharh ma ashkal min musadarat kitab Uqlidis* [Commentary on the Difficulties of Certain Postulates of Euclid's Work]; Arabic text with accompanying translation into French available in Rashed, R. And Vahabzadeh B. (1999), *Al-Khayyam Mathématicien* (Paris: Librarie Scientifique et Technique, Albert Blanchard); English translation available in Rashed, R. And Vahabzadeh B. (2000), *Omar Khayyam the Mathematician* (New York: Bibliotheca Persica Press).

Kitcher, Philip. (1976). "Hilbert's Epistemology," *Philosophy of Science* **43,** No. 1: 99-115.

Klein, Felix. (1871). "Ueber die sogenannte Nicht-Euklidische Geometrie," *Mathematische Annalen* **4**: 573-625.

Klein, Felix (1939 [2004]). *Elementary Mathematics from an Advanced Standpoint – Geometry* (London: Dover).

Kline, Morris. (1972). *Mathematical Thought from Ancient to Modern Times* (New York: Oxford University Press).

Kluge, Eike-Henner W. (1971). "Introduction," in Gottlob Frege, *On the Foundations of Geometry and Formal Theories of Arithmetic*, translated with an introduction by Eike-Henner W. Kluge (New Haven: Yale University Press): xi-xlii.

Künne, Wolfgang. (1997). "Propositions in Bolzano and Frege," in *Bolzano and Analytic Philosophy*, edited by Wolfgang Künne, Mark Siebel, and Mark Textor (Dordrecht: Rodopi): 203-240.

Künne, Wolfgang. (2008). "Frege on Truths, Truth and the True," *Studia Philosophica Estonica* **1.1**: 5-42.

Kvasz, Ladislav. (1998). "History of Geomery and the Development of its Language," *Synthese* **116**, no.2: 141-186.

Lagrange, Jean-Louis. (1867). *Oeuvres de Lagrange*, 14 volumes, edited by J. A. Serret (Paris: Gauthier-Villars).

Landry, Elaine. (1999). "Category Theory: The Language of Mathematics," *Philosophy of Science (Proceedings)* **66**: S14-S27.

Lambert, Johann Heinrich. (1786). "Theorie der Parallellinien," *Magazin für die reine und angewandte Mathematik* (Leipzig).

Lawvere, F. William. (1963). "Functorial Semantics of Algebraic Theories," *Proceedings of the National Academy of Science* **50**, no. 5: 869-872.

Lawvere, F. William. (1964). "An Elementary Theory of the Category of Sets," *Proceedings of the National Academy of Sciences of the United States of America* **52**, no. 6: 1506-1511.

Lawvere, F. William. (1966). "The Category of Categories as a Foundation for Mathematics," *Proceedings of the Conference on Categorical Algebra* (New York: Springer): 1-21.

Lawvere, F. William and Schanuel, Stephen H. (1997). *Conceptual Mathematics: A first introduction to categories* (Cambridge: Cambridge University Press)

Lawvere. F. William. (2000). "Comments on the development of topos theory," in Jean-Paul Pier (editor), *Development of Mathematics 1950-2000* (Basel: Birkhäuser): 715-734.

Lawvere, F. William. (2003) "Foundations and Applications: Axiomatization and Education," *The Bulletin of Symbolic Logic* **9**, no. 2: 213-224.

Leitgeb, Hannes and Ladyman, James. (2008). "Criteria of Identity and Structuralist Ontology," *Philosophia Mathematica* (III) **16**, no. 3: 388-396.

Levine, James. (1996). "Logic and Truth in Frege," in *Proceedings of the Aristotelian Society Supplement* 70: 141-175; revised and reprinted in Michael Beaney and Erich H. Reck (editors), *Gottlob Frege: Critical Assessments of Leading Philosophers*, Volume I (New York: Routledge, 2005): 248-269.

Locke, John. (1690). *An Essay Concerning Humane Understanding, In Four Books* (London).

MacBride, Fraser. (2006). "What constitutes the numerical diversity of mathematical objects?," *Analysis* **66**: 63-69.

Mac Lane, Saunders. (1986). *Mathematics: Form and Function* (New York: Springer).

Mac Lane, Saunders and Alexanderson, G. L. (1989). "A Conversation with Saunders Mac Lane," *The College Mathematics Journal* **20**, no. 1: 3-25.

Mac Lane, Saunders. (1992). "The Protean Character of Mathematics," in J. Echeverra, A. Ibarra, and J. Mormann (editors), *The Space of Mathematics* (New York: De Gruyter): 3-12.

Mac Lane, Saunders. (1996). "Structure in Mathematics," *Philosophia Mathematica* **3**, volume 4: 174-183.

Mac Lane, Saunders. (1998). *Categories for the Working Mathematician, Second Edition* (New York: Springer-Verlag).

Mancosu, Paolo. (1998). *From Brouwer to Hilbert: The debate on the foundations of mathematics in the 1920s* (Oxford: Oxford University Press).

McLarty, Colin. (1993). "Numbers Can Be Just What They Have To," *Nôus* **27**, no. 4: 487-498.

McLarty, Colin. (2004). "Exploring Categorical Structuralism," *Philosophia Mathematica* (III) **12**: 37-53.

Meinong, Alexius. (1904). "On the Theory of Objects," translated in R. M. Chisholm, editor, (1960) *Realism and the Background of Phenomenology* (New York: Free Press).

Mendelsohn, Richard L. (2005). *The Philosophy of Gottlob Frege* (Cambridge: Cambridge University Press).

Moore, Adrian and Rein, Andrew. (1986). "*Grundgesetze* Section 10," in Leila Haaparanta and Jaakko Hintikka (editors), *Frege Synthesized: Essays on the Philosophical and Foundational Work of Gottlob Frege*. (Dordrecht: D. Reidel): 375-384.

Moore, E. H. (1902). "Projective Axioms of Geometry," *Transactions of the American Mathematical Society* **3**: 142-158.

Moore, E. H. (1902b). "A Definition of Abstract Groups," *Transactions of the American Mathematical Society* **3**, no. 4: 485-392.

Moore, G. E. (1953). *Some Main Problems of Philosophy* (London: Unwin).

Mueller, Ian. (1969). "Euclid's *Elements* and the Axiomatic Method," *British Journal for the Philosophy of Science* **20**, no. 4: 289-309.

Newton, Isaac. (1845). *Newton's Principia. The Mathematical Principles of Natural Philosophy*, translated by Andrew Motte (New York: Daniel Adee).

Osius, Gerhard. (1974). "Categorical Set Theory: A Characterization of the Category of Sets," *Journal of Pure and Applied Algebra* **4**: 79-119.

Parsons, Terence. (1987). "On the Consistency of the First-Order Portion of Frege's Logical System," *Notre Dame Journal of Formal Logic* **28**: 161–168.

Parsons, Charles. (1990). "The Structuralist View of Mathematical Objects," *Synthese* **84**, no. 3: 303-346.

Parsons, Charles. (1995). "Frege's Theory of Number," in William Demopoulos (editor), *Frege's Philosophy of Mathematics* (Cambridge: Harvard): 182-210.

Pasch, Moritz. (1882). *Vorlesungen über neuere Geometrie* (Leipzig: Teubner).

Pasch, Moritz. (1887). "Über die projective Geometrie und die analytische Darstellung der geometrischen Gebilde," in *Mathematische Annalen* **30**: 127-131.

Plotinus. (2007). *The Six Enneads of Plotinus,* translated by Stephen MacKenna (London: Forgotten Books).

Poncelet, Jean-Victor. (1822). *Traité des proprieties projectives des figures,* 2 volumes (Paris).

Proclus. (1970). *A Commentary on the First Book of Euclid's Elements*, translated with introduction and notes by Glenn R. Morrow (Princeton: Princeton University Press). [Page references correspond to the pagination of the standard edition of the Greek text, edited by Friedlin].

Putnam, Hilary. (1967). "Mathematics without foundations," *The Journal of Philosophy* **64**, no. 1: 5-22.

Quine, W. V. O. (1970). *Philosophy of Logic* (Cambridge: Harvard University Press).

Resnik, Michael. (1965). "Frege's Theory of Incomplete Entities," *Philosophy of Science* **32**, no.3/4: 329-341.

Resnik, Michael. (1980). *Frege and the Philosophy of Mathematics* (Ithaca: Cornell University Press).

Resnik, Michael. (1981). "Mathematics as a Science of Patterns: Ontology and Reference," *Noûs* **15**, no. 4: 529-550.

Resnik, Michael (1982). "Mathematics as a Science of Patterns: Epistemology," *Noûs* **16**, no.1: 95-105.

Resnik, Michael (1997). *Mathematics as a Science of Patterns* (Oxford: Oxford University Press).

Resnik, Michael. (1999). *Mathematics as a Science of Patterns* (Oxford: Clarendon Press).

Ricketts, Thomas. (1996). "Logic and Truth in Frege," in *Proceedings of the Aristotelian Society Supplement* 70: 121-140; reprinted in Michael Beaney and Erich H. Reck (editors), *Gottlob Frege: Critical Assessments of Leading Philosophers,* Volume I (New York: Routledge, 2005): 231-247.

Ricketts, Thomas. (1997). "Frege's 1906 foray into metalogic," *Philosophical Topics* **25**: 169-188.

Ricketts, Thomas. (2010). "Concepts, objects and the Context Principle," in Michael Potter and Thomas Ricketts (editors), *The Cambridge Companion To Frege* (Cambridge: Cambridge University Press): 149-219.

Riemann, Bernhard. (1868). "Über due Hypothesen, welche der Geometrie zu Grunde liegen," *Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen* **13**: 133-152; translated into English by William Kingdon Clifford as "On the Hypotheses which lie at the Foundation of Geometry," reprinted in William Ewald (editor), *From Kant to Hilbert: A Sourcebook in the Foundations of Mathematics*, volume 2 (Oxford: Oxford University Press, 1999): 652-661.

Rosenfeld, Boris A. and Youschkevitch, Adolf P. (1996). "Geometry" in *Encyclopedia of the history of Arabic Science*, volume 2, edited by Rushdī Rāshid (New York: Routledge); 115-159.

Rowe, David. (1989). Klein, Hilbert, and the Göttingen Mathematical Tradition," *Osiris*, 2nd series, volume **5**: 186-213.

Rowe, David. (2000). "The Calm Before the Storm: Hilbert's Early Views on Foundations," in *Proof Theory: history and philosophical significance*, edited by Vincent F. Hendricks, Stig Andur Pedersen, and Klaus Frovin Jørgensen (Dordrecht: Kluwer): 55-94.

Ruffino, Marco. (2003). "Frege's Views on Vagueness," in *Logic, Truth, and Arithmetic: Essays on Gottlob Frege,* edited by Marco Ruffino, special issue of *Manuscrito* **26**, no. 2: 253-277.

Rusnock, Paul. (1993). "Remarks on the Frege-Hilbert Dispute," in Ingolf Max and Werner Stelzner (editors), *Logik und Mathematik: Frege Kolloquium 1993* (Berlin: de Gruyter): 150-161.

Rusnock, Paul. (2000). *Bolzano's Philosophy and the Emergence of Modern Mathematics* (Amsterdam: Rodopi).

Rusnock, Paul. (2013). "Kant and Bolzano on Analyticity," *Archiv für Geschichte der Philosophie* **95**, issue 3: 298-335.

Rusnock, Paul and Kerr-Lawson, Angus. (2004). "Bolzano and uniform continuity," in *Historia Mathematica*, **32**, number 3: 303-311.

Russell, Bertrand. (1903). *The Principles of Mathematics* (London: Norton).

Russell, Bertrand and Whitehead, Alfred North. (1963). *Principia Mathematica*, volume I, second edition (Cambridge: Cambridge University Press).

Saccheri, Girolamo. (1733). *Euclidus ab omni naevo vindicates: sive conatus geometricus quo stabiliuntur prima ipsa universae geometriae principia* (Milan). Translated into English by George Bruce Halsted (1986) as *Euclid Freed of All Blemish or A Geometric Endeavor in which* *are Established the Foundation Principles of Universal Geometry* (Providence: AMS Bookstore).

Saccheri, Girolamo. (1986). *Euclides Vindicatus*, edited and translated by George Halsted, with additional notes by Paul Stäckel and Friedrich Engel (New York: Chelsea Publishing).

Scholz, Erhard. (1982). "Herbart's Influence on Riemann," *Historia Mathematica* **9**: 413-440.

Scott, W. R. (1987). *Group Theory* (Mineola, New York: Dover).

Shapiro, Stewart. (1989). "Logic, Ontology, Mathematical Practice," *Synthese* **79**: 13-50.

Shapiro, Stewart. (1997). *Philosophy of Mathematics: Structure and Ontology* (Oxford: Oxford University Press).

Shapiro, Stewart. (2005). "Categories, Structures, and the Frege-Hilbert Controversy: The Status of Meta-mathematics," *Philosophia Mathematica* (III) **13**, no. 1: 61-77.

Shapiro, Stewart. (2008). "Identity, Indiscernibility, and *ante rem* Structuralism: The Tale of *I* and *–i*," *Philosophia Mathematica* (III) **16**: 1-26.

Simpson, Stephen. (1988). "Partial realizations of Hilbert's program," *Journal of Symbolic Logic* **53 (2)**: 349-363.

Skolem, Thoralf. (1922). "Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre," *Matematikerkongressen i Helsingfors den 4-7 Juli 1922, Den femte skandinaviska matematikerkongressen, Redogörelse* (Helsinki: Akademiska Bokhandeln): 217-232; English translation in [Van Heijenoort 1967, 290-301].

Sluga, Hans. (2002). "Frege on the Indefinability of Truth," in Erich H. Reck (editor), *From Frege to Wittgenstein: Perspectives on Early Analytic Philosophy* (Oxford: Oxford University Press): 75-95.

Tait, William. (2006). "Gödel's correspondence on proof theory and constructive mathematics," *Philosophia Mathematica* **14**, no. 2: 76-111.

Tappenden, Jamie. (2006). "The Riemannian Background to Frege's Philosophy," in J. Ferreirós and J. J. Gray (editors), *The Architecture of Modern Mathematics: Essays in History and Philosophy* (Oxford: Oxford University Press): 97-132.

Tarski, Alfred. (1986). "What are logical notions?," edited by John Corcoran, *History and Philosophy of Logic* **7**: 143-154.

Toepell, Michael. (1986). "On the origins of David Hilbert's 'Grundlagen der Geometrie,'" *Archive for History of the Exact Sciences* **35**, no. 4.

Torretti, Roberto. (1984). *Philosophy of Geometry from Riemann to Poincaré* (Dordrecht: D. Reidel).

Van Dalen, Dirk and Ebbinghaus, Heinz-Dieter. (2000). “Zermelo and the Skolem Paradox,” *The Bulletin of Symbolic Logic* **6**, no. 2: 145-161.

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

Von Fritz, K. (1955). “Die APXAI in der griechischen Mathematik” *Archiv für Begriffsgeschichte* **I**: 13-103.

Von Neumann, John. (1925). “Eine Axiomatisieurng der Mengenlehre,” *Journal für die reine und angewandte Mathematik* **154**: 219-240; English translation in [Van Heijenoort 1967, 393-413].

Von Neumann, John. (2005). *John Von Neumann: Selected Letters* (Providence: American Mathematical Society).

Wallis, John. (1693). *“De Postulato Quinto,”* Opera Mathematica, Vol. II (Oxford): 669-678.

Weierstrass, Karl. (1880). “Zur Functionenlehre,” *Monatsberichte der K. Akad der Wiss.*: 201-230; also in Karl Weierstrass (1894-1927), *Mathematische Werke*, *volume II*: 201-230.

Wehmeier, Kai. (1999). “Consistent Fragments of *Grundgesetze* and the Existence of Non-Logical Objects,” *Synthese* **121**: 309-328.

Wells, Rulon S. (1951). “Frege’s Ontology,” *The Review of Metaphysics*, **4**, no. 4: 537-573.

Weyl, Hermann. (1921). “Über die neue Grundlagenkrise der Mathematik,” *Mathematische Zeitschrift* **10**: 39-79.

Weyl, Hermann. (1970). “David Hilbert and His Mathematical Work,” abridged in Constance Reid, *Hilbert* (Berlin: Springer-Verlag): 245-283.

Wilson, Mark. (1993). “There’s a hole and a bucket, Dear Leibniz,” *Midwest Studies in Philosophy* **18**: 202-241.

Wussing, Hans. (1984). *The Genesis of the Abstract Group Concept: A Contribution to the History of the Origin of Group Theory*, translated by Abe Shenitzer with the editorial assistance of Hardy Grant (Cambridge: MIT Press).

Ziegler, Renatus. (1985). *Die Geschichte der geometrischen Mechanik im 19. Jahrhundert: eine historisch-systematische Untersuchung von Möbius und Plücker bis zu Klein und Lindemann* (Steiner Verlag: Stuttgart).

Zermelo, Ernst. (1908). "Untersuchungen über die Grundlagen der Mengenlehre I," *Mathematische Annalen* **65**: 261-281; English translation in [Van Heijenoort 1967, 199-215].