Università
Ca'Foscari
Venezia

Master's Degree

in Philosophical
Sciences

Final Thesis

# Numbers, Objects and Abstraction

Notes for a philosophical interpretation of mathematics

**Supervisor**
Ch. Prof. Pagani Paolo

**Assistant supervisor**
Ch. Prof. Jabara Enrico

**Graduand**
Fabio De Martin Polo
Matricolation number
988657

**Academic Year**
2018 / 2019

*To my beloved family.*
*To Alice – my second family.*

# Contents


**Introduction** **6**

**I Hunting abstract objects** **11**

**1 From Gödel to Benacerraf** **12**
1.1 Gödel's Extreme Platonism . . . 12
1.1.1 Metaphysical Foundation . . . 13
1.1.2 Epistemology and Mathematical Intuition . . . 14
1.1.2.1 Incompleteness and Philosophy of Mathematics . . . 16
1.2 Against Platonism I: Metaphysics . . . 23
1.2.1 Numbers are not sets! Benacerraf's Argument . . . 23
1.2.2 Metaphysical Misunderstandings . . . 24
1.3 Against Platonism II: Epistemology . . . 26
1.3.1 Benacerraf-Field's Dilemma . . . 26
1.3.2 Another Challenge: What is Mathematical Doxology? . . . 31

**2 Numbers, sets and objects I. Naïve Considerations** **33**
2.1 Short Introduction to Frege . . . 33
2.1.1 Frege's Logical Objects and Extensions . . . 33
2.1.1.1 Frege's Project and its Failure . . . 34
2.1.1.2 Logicism, Extensions and Natural Numbers . . . 40
2.2 Short Introduction to Cantor . . . 43
2.2.1 Naïve Set Theory: Foundations . . . 43
2.2.2 On Infinite and Finite Sets . . . 49
2.2.3 Reconsidering Benacerraf's Thesis I . . . 53
2.2.4 Ontological Remarks I . . . 56
2.3 Little interlude. Back to Philosophy . . . 57
2.3.1 Frege's Notion of Abstraction . . . 58
2.4 Short introduction to Dedekind . . . 59
2.4.1 *Systems* and Logical Abstraction . . . 59
2.4.1.1 Dedekind's Contributions to the Early Development of Set Theory . . . 60

2.4.1.2 Philosophical Revaluation of Dedekind's Notion of Logical Abstraction . . . . . . . . . . . . . . . . . . . . 65
2.4.1.3 Logicist Attitudes: Frege *vs* Dedekind . . . . . . . . 68
2.5 Lightened or Heavy Platonism? . . . . . . . . . . . . . . . . . . . . . . 71

**3 Numbers, sets and objects II. Axiomatic Considerations 76**
3.1 One Number, Different Sets! . . . . . . . . . . . . . . . . . . . . . . . . 76
3.2 Axioms for Set Theory . . . . . . . . . . . . . . . . . . . . . . . . . . . 80
3.2.1 The Meaning of the Axiomatic Method . . . . . . . . . . . . . 80
3.2.2 Towards an Axiomatization. Open Problems: 1880-1930 . . . 82
3.2.3 Zermelo's Axioms . . . . . . . . . . . . . . . . . . . . . . . . . . 85
3.2.3.1 Becoming Precise! (Mathematization) . . . . . . . . 85
3.2.3.2 Ordered Pairs: from Frege to Kuratowski . . . . . . 90
3.2.4 Zermelian Considerations on the Axioms . . . . . . . . . . . . 97
3.2.4.1 Well-Orderings, Choices and Axiomatic Method . . . 97
3.2.4.2 Infinite Sets, Natural Numbers and Axioms . . . . . 103
3.2.5 von Neumann's contribution . . . . . . . . . . . . . . . . . . . 109
3.2.5.1 Becoming Precise! (Consolidation) . . . . . . . . . . 109
3.3 Zermelo's *vs* von Neumann's Sets . . . . . . . . . . . . . . . . . . . . 119
3.3.1 Reconsidering Benacerraf's thesis II . . . . . . . . . . . . . . . 121
3.3.1.1 Logical-mathematical Development of Benacerraf's Paper . . . . . . . . . . . . . . . . . . . . . . . . . . 126
3.3.1.2 Towards a Philosophy of Sets . . . . . . . . . . . . . 128
3.3.1.3 Dedekindian Reflections on "Mathematical Representations" . . . . . . . . . . . . . . . . . . . . . . . . 130

**II Individuating abstract objects 136**

**4 Mathematics by Abstraction 137**
4.1 Three Problems for Abstraction Principles . . . . . . . . . . . . . . . 138
4.1.1 Symmetry/Asymmetry . . . . . . . . . . . . . . . . . . . . . . . 138
4.1.2 "Bad Company" or "Embarrassment of riches" problem . . . . 141
4.1.3 The "Julius Caesar" Problem . . . . . . . . . . . . . . . . . . . 142
4.2 An Interesting Response: Groundedness and Abstraction . . . . . . . 144
4.3 Dependence and Supervenience . . . . . . . . . . . . . . . . . . . . . . 148
4.3.1 Introductory Remarks . . . . . . . . . . . . . . . . . . . . . . . 148
4.3.2 Formal Framework . . . . . . . . . . . . . . . . . . . . . . . . . 149
4.3.2.1 Logical Preliminaries . . . . . . . . . . . . . . . . . . 149
4.3.2.2 Dependence, Identity and Difference . . . . . . . . . 155
4.3.2.3 The Set of Grounded Identity and Difference Facts . 159
4.3.3 Hume's Principle, Basic Law V and Grounded Concepts . . . 164

**5 Philosophy of Abstraction and Platonism 170**
5.1 Towards a Metaphysical Characterization of Abstract Objects . . . . 171
5.1.1 The "Parallelism" Thesis . . . . . . . . . . . . . . . . . . . . . . . 171
5.1.2 The "Indefinite Extensibility" Thesis . . . . . . . . . . . . . . . 174
5.1.3 Our Metaphysical Picture . . . . . . . . . . . . . . . . . . . . . . 176
5.2 Platonism & Ontological Remarks II . . . . . . . . . . . . . . . . . . . 178
5.3 Groundedness, Impredicativity and "Bad Companions" . . . . . . . . 180
5.4 Epistemological Suggestions . . . . . . . . . . . . . . . . . . . . . . . . 182
5.4.1 Frege, Logicism and Epistemology . . . . . . . . . . . . . . . . . 182
5.4.2 Abstract Objects: Benacerraf, Doxology and Epistemology . . 185
5.4.2.1 Suggestion 1 . . . . . . . . . . . . . . . . . . . . . . . . 186
5.4.2.2 Suggestion 2 . . . . . . . . . . . . . . . . . . . . . . . . 188

**Bibliography 193**

# Introduction

*Every good mathematician is at least*
*half a philosopher, and every good philosopher*
*is at least half a mathematician.*

Gottlob Frege

*But every error is due to extraneous*
*factors (such as emotion or education)*
*reason itself does not err.*

Kurt Gödel

⋆ Why mathematics? As a student of philosophy, my main interests lie in the *why* of things. Humans, since the origins, have tried to understand the main characters and the main features of what we encounter in our everyday life, not limiting themselves just to empirical observations but, and especially, by trying to understand things thanks to *reasoning* and to *intellectual* abilities. Not by chance, indeed, it's since the antiquity that philosophers had learnt that, what is visible, is not the totality of what there is, and that a proper philosophical investigation has to take into account this fact. This present work, indeed, is in continuity with my idea of trying to answer *why* questions, but – for space and interests reasons – we have decided to restrict our considerations to one of the oldest disciplines all-over: **mathematics**. As clear, mathematics occupies a very particular position among human knowledge and sciences, for its truths are clearly not of the kind of those of physics or chemistry, but – in this respect – it gives the basis to develop all the empirical sciences, such as physics and chemistry themselves. Obviously, thanks to this particular character, philosophers – since the birth of mathematics itself – have began asking what kind of disciplines we have in front, which kinds of objects it treats, which kinds of relations exist between mathematical entities and human knowledge. This kind of questions are usually covered by a discipline called **philosophy of mathematics** and – as many scholars sustain – its great "birth" has happened exactly with Plato's school. Plato, since the fifth century a.D., elaborated a metaphysical[1] system, which

[1]Actually the term "metaphysics" appeared just after Plato and Aristotle, and its usage has to be attributed to Andronicus of Rhodes, whose main intent was that of giving a complete systematizations of Aristotle's writings. In particular, Andronicus proposed to systematize the books concerning "prime philosophy", after those regarding physics and the natural sciences. This way of systematizing has its reason in Aristotle's methodology, for which – in knowing and

attributed a very special place to mathematics and its objects. Usually, philosophers speak of Plato's theory of forms, since – for the ancient philosopher – our reality is just a "black mirror" of another and absolutely determinate reality, which he called "Hyperuranion". Within this latter, Plato inserted mathematical concepts and geometrical figures, by considering them as **objects** in the proper sense, i.e., as something that is mind-independent, objective and needs to be discovered after a special inquiry. Differently, from the common and, somehow, limited conventional use of the term "object", we point out that in philosophy of mathematics, its usage is very broad and, roughly, it is supposed to refer to whatsoever mathematical entity – functions, variables, numbers, . . . . But, as in the case of natural sciences, determining the character of mathematical objects is not a simple work to accomplish, since mathematical data are generally not conceived as sensorial data and, hence, the inquiry to which they are subject is of a particular kind, i.e., mainly theoretical and not experimental.

⋆ Logic, Old & New Friends. The current research – apart being in continuity with my idea of doing philosophy by considering *why* questions – is also in continuity with the argument to which I devoted my Bachelor's thesis. The work I've done regarded Kurt Gödel's ontological argument and the philosophical assumptions that inspired him in providing such a proof. Confronting with Gödel's deep formal and philosophical work, I was driven in expanding my researches and in considering Gödel's position and his work as good starting point. Furthermore, what Gödel profoundly (and, of course, indirectly) taught me, is that philosophy and mathematics have more in common than what is mainly thought and that their collaboration is fruitful for both. So, importantly for this feature, he suggested us that, if we want to understand philosophically mathematics, a theoretical – and mainly not experimental – methodology has to be adopted. In the context of this present work, we adopt, indeed, several formal methods, coming from **logic**, in order to prove things and to see, in which sense, in most cases, mathematical theorems and propositions are supposed to "enclose" philosophical claims.

Having spent much time on Gödel's philosophy – after obtaining my Bachelor's Degree – I've began my Master's studies by focusing my attention mainly on logic. This "second" part of my studies has brought me in searching a new adventure where to have the opportunity of deepening and, once for all, learning logic and its fundamental features. My choice fell on the University of Munich (Germany), where I took some classes at the Munich Center for Mathematical Philosophy, for one semester. There, I've deepened my logical knowledge, and, in particular, my understanding of Gödel's celebrated Incompleteness Theorems – gaining consequently an always

apprehending – we go from the given and most evident things (physics) up to the most hidden and mysterious features of reality (prime philosophy). In this sense, the books concerning prime philosophy should – in accordance to Aristotle's methodology – put after those of physics. The term devoted to represent the study of prime philosophy became, indeed, from Andronicus on out, "metaphysics", which in ancient Greek literally signifies "beyond physics".

clearer insight for what concerns his philosophy of mathematics. Additionally, always in the context of my studies abroad, another perspective, in particular, captured my attention – Gottlob Frege's works. His pioneer work, within logic and the philosophy of mathematics, has to be considered not only for his historical importance, but, also and especially, for his innovative way of conceiving the work of philosophers. His leading idea of "purifying" philosophical and mathematical languages from the errors coming from the ambiguities of natural languages has brought him in formulating one of the first logical **formal** vocabularies ever. We will still have time to return on Frege's and Gödel's philosophical ideas. For the moment, it is enough to focus the attention upon what I think that Frege and Gödel brilliantly suggested us from a methodological point of view:

- GÖDEL-STYLE. Mathematical statements are philosophically meaningful and an appropriate research within the philosophy of mathematics has to take into account – as fundamental data – mathematical results.
- FREGE-STYLE. Natural languages are full of ambiguities and their usage is not that healthy for formulating philosophical precise claims or mathematical correct results. Formal languages are, indeed, helpful in seeing things from a clearer and preciser perspective.

In the whole of the present work we will try to follow these two suggestions, by – additionally – explaining why we think that Frege's and Gödel's works in the philosophy of mathematics are of profound inspiration for our analysis.

⋆ HUNTING & INDIVIDUATING. We've said that a formal methodology will be adopted along the entire thesis and, hence, logical languages and systems will be deepened and discussed. Now, from a strictly philosophical point of view, we have – for clarity – distinguished two main areas of the philosophy of mathematics: from one side we will clarify metaphysical and ontological features of mathematics; from the other side, instead, we will focus our attention on the problems concerning the knowledge and apprehension of mathematics and of what it involves. For precision let's say that we believe that the two disciplines – the **metaphysics** and the **epistemology** of mathematics – are very close and strictly related and that we distinguish them just for clarity and to avoid misunderstandings. In this sense, we may say that the first ones try to argue in favour or against an existential characterization of mathematics and its objects, while the latter – usually in accordance with the first one – justifies their knowledge and explains how humans are supposed to get in touch with mathematical entities. Metaphysics/ontology and Epistemology, so, finally, furnish the contexts of our reflections with respect to our subject of investigation, namely mathematics and its objects.

With respect to the main theme of our inquiry, it could be said that the following thesis is about one of the most discussed arguments, not only within the philosophy of mathematics, but also in all other philosophical areas. Roughly, if we pay attention to all of ours, for example, numerical statements – such as "There are two pens", and so on – we may easily see that we are treating the number-word "two" as referring to something *existent*. We've said that it is possible to track back the birth of philosophy of mathematics to Plato's school and that, importantly, his leading ideas have in, some forms, survived and influenced generations of philosophers after him[2]. Not by chance, indeed, two of the most celebrated sustainer of contemporary **Platonistic** philosophies of mathematics are the two main sources of inspiration of the present work – again, Frege and Gödel. In this sense, roughly, our work takes seriously Plato's idea of mathematical entities as existent, self-subsistent and mind-independent, i.e., as entities of a particular kind and very different – in its metaphysical characterization – from concrete or physical bodies. These particular guys – conventionally called **abstract objects** – indeed, will occupy our entire reflections: we will try to consider the possibility of admitting these objects within a philosophy of mathematics and, so, consequently, arguing on how their admission should be conceived of – metaphysically/ontologically and epistemologically. Exactly this objective has motivated us in dividing our work in two main parts. In the first part – titled Hunting Abstract Objects – we will consider some criticism that has been moved to mathematical Platonism, by trying to consider its tenability from both perspectives, philosophical and logical-mathematical. In this sense, our first part is a **hunt**, which – at the end – is supposed deliver a plausible answer for what concerns our main philosophical doubt concerning, at least, the question of the tenability and possibility of having abstract objects within a philosophy of mathematics. Differently, the second part – whose title is Individuating Abstract Objects – takes the conclusions of the first part as starting points and develops a methodology which is meant to help to **individuate** and recognize different and well-defined abstract objects.

As obvious, we do not cover *any* topic and, indeed, our considerations will be restricted to a particular context – furnished in Chapter 1. Starting from this setting, we will – as the title remarks – try to develop a tenable **philosophical interpretation of mathematics**. We precise, since now, that this version of our following study will be subject to further deepening and developments, and – as we hope – we wish to have given, at least, an idea of the direction which we think might be considered and deepened in contemporary philosophy of mathematics.

Anyway, we do not want to drain all fun out of the reading, and so we stop our introduction here.

Venezia
June, 2019

[2]Instructively, the English mathematician, logician and philosopher A. N. Whitehead remarked: «The safest general characterization of the European philosophical tradition is that it consists of a series of footnotes to Plato».

**Acknowledgements.** I would like to offer my special thanks to professor Jabara Enrico, who assisted, helped and encouraged me during the entire writing of my thesis. As a student of philosophy, I am especially thankful to have had prof. Jabara's assistance – for his suggestions and discourses "fomented" my passion and interest towards the wonderful world of mathematics. Additionally, I would like to express my deep gratitude to professor Pagani Paolo, whose support and indications have been fundamental, not only for the development of my thesis, but also and especially for my formation as individual and, maybe, as future philosopher. Importantly, I wish to thank various friends and colleagues for their sincere interest to this work, especially, dott. Cecconi Alessandro, dott. Levorato Riccardo, dott. Rossi Niccolò and dott. Fanti Rovetta Francesco.

As obvious, I wish to thank my parents – Sandro and Manuela – for having given me the opportunity to live the life I am living and for loving me unconditionally everyday of my life. My special thanks are extended to my brilliant brother Simone, for his tolerating and supporting me in every moment of my studies.

Finally, if "life is nothing without friends", I wish to thank the person to which this work is dedicated, Alice – whose friendship is the diamond that let my days constantly shine.

# Part I
# Hunting abstract objects

# Chapter 1

# From Gödel to Benacerraf

**Overview.** In this first chapter our aim is twofold. Firstly, we wish to draw a path that does follow a theoretical development of the problems to be considered. In this sense, – as the title suggests – we will focus exactly on the philosophical considerations, proposed by one of the major exponents of the so-called "Platonistic" school in the philosophy of mathematics, – Kurt Gödel –, by, linking them to the challenges that the French philosopher Paul Benacerraf moved to a supposed sustainer of the Platonistic claims. Secondly, since our purposes are not merely exegetical, we will especially pose the attention on the foundations and on the tenability of the claims we will analyse and submit to judgement. Hence, in order to be clear and "clean", as much as possible, starting from the very beginning, formal tools will be adopted and any question will be "divided" into metaphysical and epistemological concerns.

## 1.1 Gödel's Extreme Platonism

Kurt Gödel is usually acknowledged through the major mathematicians and logicians ever and this is an undoubtedly and widely accepted fact. What, maybe, is less known is that Gödel's interests were broader then they concerning the range of pure mathematics and, indeed, he spent his late career, especially, in doing philosophy and practising physics. His philosophical production, unfortunately, is not uniform and the writings that compose it have, in most cases, the character of personal notes. Anyway, from the writings that we possess, it is possible to reconstruct the fundamental claims Gödel, explicitly or implicitly, endorsed and to establish effectively something like a "Gödelian" philosophy of mathematics. In order to be precise, consider that Gödel's philosophical concerns were very inclusive – philosophy of mind and metaphysics were, for example, two of his main interest areas –, even if we will restrict here our considerations just to his metaphysics and epistemology of mathematics.

### 1.1.1 Metaphysical Foundation

With respect to the philosophy of mathematics Gödel held a position that we will call "Extreme Platonism". The two fundamental aspects of his philosophical position are *Platonism*, by one side, and its *extreme* version, by the other.
First of all, let's define generally what a platonist would endorse about mathematics. The fundamental claims are the following two:

(P1) Mathematical objects exist[1].

(P2) Mathematical objects are abstract objects.

So, for instance, if we say, with respect to the set of natural numbers, that "there are infinite prime numbers", and we endorse a platonistic view of mathematical objects, we would intend that there are infinite abstract entities that are characterized as "prime numbers". In addition, it is possible to define abstract objects by saying that, unlike the concrete ones, they do not have a collocation in space and time. So, for example, the table in front of me is not abstract in virtue of its being in a determined space (world, nation, city, library) and its being inserted into a specific time (year, day, etc.). Differently, the natural number 2 is not spatio-temporal located and, therefore, we can think of it as a sort of platonic universal, which we have to "catch" in some way.
Gödel agreed with the general assumptions of mathematical Platonism and tried to develop a tenable account. Indeed, «his platonistic view was more sophisticated than that of the mathematician in the street»[2]. Gödel thought that, as physical theories should explain physical objects and their connections, so mathematics has to explain its particular objects and their fundamental relations. Additionally, Gödel extended the parallelism between mathematical and physical universe, by considering that, if concrete things are not *free* creations of human mind, and if, the thesis that mathematics describes (abstract) objects is true, mathematical entities shouldn't be considered either "ideas" or "mental constructions". Indeed,

> Classes and concepts may, however, also be conceived as real objects, namely classes as "pluralities of things" or as structures consisting of a plurality of things and concepts as the properties and relations of things existing independently of our definitions and constructions. It seems to me that the assumption of such objects is quite as legitimate as the assumption of physical bodies and there is quite as much reason to believe in their existence. They are in the same sense necessary to obtain a satisfactory system of mathematics as physical bodies are necessary for a satisfactory theory of our sense perceptions [...][3].

[1] Gödel enumerate 14 propositions in a list titled "My philosophical viewpoint", that, posthumous, has been transcribed by Cheryl Dawson and published by Hao Wang. In this list, at the second position, instructively, Gödel wrote: «Concepts have an objective existence» (Wang 1996, p. 316). For remarks consider, among others, Kennedy 2007, §3, Horsten 2007, §3.1 and Linnebo 2009c.

[2] Horsten 2007, p. 16.

[3] Gödel 1944, pp. 456–457.

Therefore, Gödel's first step of philosophical foundation of mathematical Platonism consists in accepting the thesis for which mathematics (as sciences) is concerned with the description of a realm composed by entities which are not spatio-temporally located (different from physical objects). In his own words:

> [With Platonism] I mean the view that mathematics describes a non-sensual reality, which exists independently both of the acts and of the dispositions of human mind and is only perceived, and probably perceived very incompletely, by the human mind.[4]

The second step into Gödel's foundation of mathematical Platonism is concerned with the epistemological issues arisen from the metaphysical status he attributed to mathematical objects. Indeed, we may ask to a platonistic philosopher: "Since we are believing that numbers, sets, functions, etc. are abstract objects, how can we apprehend them?" "How are we humans supposed to acquire knowledge of 'things' which aren't immediately given by our sensorial experience?" Gödel himself fell in troubles while trying to answer these questions and the explanation he gave is the main reason why we are calling his position "extreme".

### 1.1.2 Epistemology and Mathematical Intuition

On our purpose, the main source for Gödel's epistemological account of Platonism is his article devoted to Cantor's set theory[5]. First of all, when Gödel tried to explain how human minds can achieve mathematical knowledge introduced the notion of *mathematical intuition*[6]. Following Gödel, in order to explain how this particular human faculty works, we have to consider the parallelism between sciences and mathematics again. We may think that humans achieve knowledge of the physical world by perception, that is with innate sensorial faculties. Roughly, I know that the object in front of me is a concrete entity, because my visual perception and any scientific inquiry would confirm me that what I am actually seeing is, for example, a table that has a specific location in space and time. For Gödel, as for many Platonists, the way we apprehend mathematical objects is parallel to the way we know physical bodies. Actually, this is not Gödel's answer, indeed, he thought that this parallelism could just provide a useful way to answer the question. Moreover, he thought that, even if we are allowed to speak of mathematical *perception*, we must not forget the very nature of mathematical entities, namely their being abstract. Therefore, the link between physics and mathematics becomes just heuristic: if we want to understand how mathematicians arrive to their results, we have to consider how scientists arrive to their scientific discoveries. Once then we've

[4] Gödel 1951, p. 323.

[5] Gödel 1947 and Gödel 1995.

[6] For an interesting discussion on the connections between Platonism and mathematical intuition in Gödel's philosophy of mathematics, see Parsons 1995. In addition, it is to notice that Gödel's epistemology of mathematics was strictly influenced by Gottfried Leibniz and Edmund Husserl's works. The austrian logician thought that some metaphysical intuitions of Leibniz and Husserl's phenomenology could provide good basis to clarify some aspects of the platonistic position (Berto 2008, especially part II).

understood that mathematics studies abstract objects that we perceive, we can forget about the parallelism with natural sciences, and focus our attention on the notion of *mathematical perception*. Additionally, the parallelism is very useful to draw attention on the notion of "certainty" that a Platonist would endorse. Our sensorial experience is, at different degrees, fallible and might be corrected by successive perceptions. The same can happen in mathematics. For instance, even if Frege's intuitions on the "third realm" were very sophisticated, further researches, such as Russell's, have shown that one of the most basic intuitions of his system entailed a contradiction. For Gödel, this situation reflects what happens in natural sciences.
Abstract entities and mathematical intuition are the two principal elements of Gödel's Platonistic conception. Much of his late philosophical work has been devoted to the understanding of how it is possible for humans to access the Platonic heaven of mathematics. Since mathematical intuition provides us with a sort of certainty with respect to mathematical theories, Gödel asks himself which kind of "data" and "evidence" are necessary to believe in theories? For instance, what kind of *evidence* do I need, to affirm that the "Zermelo-Fraenkel set theory with the Axiom of Choice" (ZFC), really describes the mathematical universe composed by abstract mathematical objects as sets, classes, functions and so on? Let's focus on the notion of *evidence*. We begin with a distinction that Gödel himself proposed[7]:

1. "Intrinsic evidence"
2. "Extrinsic evidence"

The first kind of evidence is the certainty we get by considering the axioms of a theory. Roughly, axioms codify the basic truths of the particular branch of mathematics we are studying, so that further and more complicated truths can be obtained. In other words, if we look at the axioms of ZFC, for instance, we are looking at the theory's internal and fundamental articulation. Hence,

> [...] despite their remoteness from sense experience, we do have something like a perception of the objects of set theory, as is seen from the fact that the axioms force themselves on us as being true. I don't see any reason why we should have less confidence in this kind of perception, i.e. in mathematical intuition, than in sense perception [...][8].

The second kind of evidence is different from the first one and do not rely on the internal articulation of a mathematical theory[9]. Indeed, it is difficult to understand how an intuition can be responsible of our way of determining the truth or the falsehood of specific mathematical axioms. Gödel exactly noticed that «mathematical intuition might not be strong enough to provide compelling evidence for axioms»[10]

[7]See also Linnebo 2011, pp. 172–175.
[8]Gödel 1947, pp. 483–484.
[9]This second kind of evidence is strictly linked to the philosophical meaning Russell was ascribing to axioms. See chapter 3, section "The meaning of the axiomatic method".
[10]Horsten 2007, p. 16.

and, therefore, he tried to develop an account of *evidence* based upon the notion of "true consequence". Let's focus on Gödel's argumentation:

> [...] even disregarding the intrinsic necessity of some new axiom, and even it has no intrinsic necessity at all, a probable decision about its truth is also possible in another way, namely, inductively by studying its "success". Success here means fruitfulness in consequences, in particular in "verifiable" consequences [...][11].

So, the ability to discover always new truths from a basic set of axioms, via proofs, is exactly what Gödel had in mind. If we want to declare that our ZFC axiomatization effectively describes the set theoretic universe, then we should consider how many new truths of that universe are provable from the ZFC axioms. Therefore, what this second kind of evidence requires is not an intuition of the internal necessity of whatever mathematical axiom, but that the logical consequence, that brings us from the axioms to new theorems, is well-established:

> There might exist axioms so abundant in their verifiable consequences, shedding so much light upon a whole field, and yielding such powerful methods for solving problems [...] that, no matter whether or not they are intrinsically necessary, they would have to be accepted at least in the same sense as any well-established physical theory.[12]

This inspired Gödel in searching new axioms to be added to ZFC, which could have decided highly independent questions, such as the *continuum hypothesis*. Generally speaking, this means that, in principle, according to Gödel's Platonism, every mathematical problem has a solution and, so, for example, Cantor's continuum conjecture must be either true or false. But, actually, further set-theoretical researches have shown that the *continuum hypothesis* is an *independent* question, that is its truth (or falsehood) does not depend upon a particular axiomatization of set theory. So, if we are Gödel-like Platonists, we are convinced that every mathematical problem has a definite solution and, hence, that the mathematical universe is perfectly complete and determined. What is fallible are our intuitions and evidences for new axioms: the *continuum hypothesis* «undecidability from the axioms assumed today can only mean that these axioms do not contain a complete description of that reality»[13].

#### 1.1.2.1 Incompleteness and Philosophy of Mathematics

First of all, we've seen that Gödel's Platonism held that there are mathematical abstract entities that have to be considered as the mathematician's data. Additionally, we've seen that mathematical axioms stand to mathematical data (abstract objects) as physical laws stand to sensorial bodies (concrete objects). The parallelism between natural sciences and mathematics has shown that as scientists acquire their data by (direct) perception, so mathematician enter the "third realm" by (indirect) intuition. This indirect mathematical intuition is based upon of two basic kinds of evidence:

[11] Gödel 1947, p. 477.
[12] Gödel 1947, p. 477.
[13] Gödel 1947, p. 476.

"intrinsic" and "extrinsic". An intrinsic evidence is acquired by considering just the axioms of the theory we're studying and nothing else. Differently, extrinsic evidence is achieved by considering the "consequences" that the axioms of a determined theory have, that is the more new theorems are provable from the axioms of the theory we are studying, the more the axioms themselves are to be considered valid. Additionally, we're told that, it «should be noted that mathematical intuition need not be conceived of as a faculty giving an *immediate* knowledge of the objects concerned»[14]. Indeed, since our way of "extrapolating" new informations from the mathematical reality, is by intuition of the axioms' validity and their consequent application in proofs, the way we perceive that reality should be considered *incomplete*[15]. Our mathematical concepts and objects are not sufficiently defined and fine-grained, and this implies that some questions, for instace, some set-theoretical issues, cannot be "immediately" decided. In other terms, the fact that mathematical perceptions are not immediate and, at different degrees, fallible, implies – according to Gödel – the "incompleteness claim", i.e., that position for which, in conclusion, our access to the mathematical universe is not (and cannot be) definitely complete. But, as Gödel pointed out:

> This seems to to be an indication that one should take a more conservative course, such as would consist in trying to make the meanings of the terms "class" and "concept" clearer, and to set up a consistent theory of classes and concepts as objectively existing entities.[16] [...] the certainity of mathematics is to be secured not by proving certain properties by a projection onto material systems – namely, the manipulation of physical symbols – but rather by cultivating (deepening) knowledge of the abstract concepts themselves which lead to the setting up of these mechanical systems, and further by seeking, according to the same procedures, to gain insights into the resolvability, and the actual methods for the solution, of all meaningful mathematical problems.[17]

Gödel thought indeed that one of his main results "purpotes" his incompleteness claim, that is, in other terms, the conviction that our ordinary methodologies are not subtle and fine-grained enough to describe the entire "mathematical realm". The logician, indeed, in 1931, proved his famous Incompleteness Theorems for Arithmetic: let's focus for a moment upon the first one[18]. At the beginning of the century many mathematicians and philosophers were convinced that a suitable logical theory could provide the base to derive the basic truths of arithmetic, that is of Number Theory. In order to see how it is possible to define number theory with axioms consider the following definitions:

---

[14]Gödel 1947, p. 484.

[15]Always in the list of his main philosophical propositions, Gödel affirmed that «[t]here is incomparably more knowable *a priori* than is currently known», Wang 1996, p. 316.

[16]Gödel 1944, p. 468.

[17]Gödel 1961, p. 383.

[18]See Gödel 1931, Button and Walsh 2011, Berto 2008, and, especially, Smith 2013.

**Definition 1.** The Robinson[19] Arithmetic $\mathsf{Q}$ is given by the universal closures of the following eight axioms:

Q1. $S(x) \neq 0$

Q2. $S(x) = S(y) \rightarrow x = y$

Q3. $x \neq 0 \rightarrow \exists y[x = S(y)]$

Q4. $x + 0 = x$

Q5. $x + S(y) = S(x + y)$

Q6. $x \times 0 = 0$

Q7. $x \times S(y) = (x \times y) + x$

Q8. $x \leq y \longleftrightarrow \exists v[x + v = y]$

If we add to $\mathsf{Q}$'s axioms Q1-Q8 the following Induction Schema we get the theory of Peano[20] arithmetic, $\mathsf{PA}$:

$$[\varphi(0) \wedge \forall y\, (\varphi(y) \rightarrow S(\varphi(y))) \rightarrow \forall y\, \varphi(y)]$$

In order to give an intuitive idea of Gödel's theorems it's enough to consider just $\mathsf{Q}$'axioms and some other notions. First of all, consider that our theory is composed by symbols: we're presented with a logical vocabulary (variables, connectives, quantifiers, ...) and with a non logical vocabulary, in this case, the language composed by the constant 0 and the "Successor" unary function, $S(x)$. The $n^{th}$ application of the successor function to zero, $S...S(0)$, intuitively, gives us whatever natural number, greater than 0, we desire. Gödel's idea, in order to show the limits of formal theories of arithmetic, established a particular coding strategy which transforms each symbol of our (logical and non logical) language into a natural number. The first task is to enumerate the symbols of our language, for instance:

[19]Raphael Mitchel Robinson (1911 – 1995) has been an expert in mathematical logic, set theory, geometry, number theory, and combinatorics. In 1937 he has set out a simplified version of the John von Neumann (1923) axiomatic set theory and he has co-worked with Alfred Tarski at Berkeley's mathematics department. In 1950 Robinson proved that a formalized theory need not to have infinitely many axioms. "Robinson arithmetic", $\mathsf{Q}$, indeed, is finitely axiomatizable because it lacks the so-called Schema of Induction. In any case, $\mathsf{Q}$, like Peano arithmetic $\mathsf{PA}$, is incomplete and undecidable in the sense of Gödel' Incompleteness theorems (see the next sections and Gödel 1931).

[20]Giuseppe Peano (1858–1932), has been a mathematician and logician, he served as professor of mathematics at the University of Turin from 1890 to 1932. In 1891 he founded the *Rivista di matematica* ("Review of mathematics"). Peano is especially known for having developed, around 1903, an international auxiliar language, called *Latino sine flexione*, based on Latin. Peano's contributions to mathematics include the simplification of logico-mathematical symbolism, the first statement of vector calculus (*Elementi di calcolo geometrico*, published in Turin in 1891) and some important results concerning the ordinary differential equations. He also obtained crucial and subtle results within the debate around the foundations of arithmetic and of sets. Indeed, Peano's postulates (1899) gave a set of five axioms for the theory of natural numbers that allowed arithmetic to be constructed as a formal-deductive system.

$$s_1, s_2, s_3, s_4, \ldots$$

Suppose to have an arbitrary finite string of the symbols:

$$\sigma = s_{k_1}, s_{k_2}, s_{k_3}, \ldots, s_{k_n}$$

According to Gödel's encoding strategy, we have to encode $\sigma$ with a number indicated as follows:

$$\ulcorner \sigma \urcorner = \pi_1^{k_1} \times \pi_2^{k_2} \times \pi_3^{k_3} \times \cdots \times \pi_n^{k_n}$$

In order to see more closely what Gödel had in mind, consider the following arbitrary coding definition:

| $\neg$ | $\wedge$ | $\vee$ | $\rightarrow$ | $\leftrightarrow$ | $\forall$ | $\exists$ | $=$ | ( | ) | 0 | $S$ | $+$ | $\times$ | $x$ | $y$ | $z\ldots$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 3 | 5 | 7 | 9 | 11 | 13 | 15 | 17 | 19 | 21 | 23 | 25 | 27 | 2 | 4 | 6... |

As it is clear we've encoded each symbol with an even number, except variables, that have been assigned with an odd one. For the sake of the argument, take the following formula, $\exists y S(x) = y$ and consider it our $\sigma$ to be coded. According to the previous mentioned strategy we have to:

1. write the position of the symbol we want to code using a prime number, indicated as $\pi_k$;
2. write the symbol-code as an exponent of the prime number: $\pi_k^{k_n}$;
3. multiply the number sequence $\pi_1^{k_1} \times \pi_2^{k_2} \times \pi_3^{k_3} \times \cdots \times \pi_n^{k_n}$ in order to get the "Gödel number" of the string of symbols.

So, for example, $\exists y S(x) = y$ is coded in the following manner:

| $\exists$ | $y$ | $S$ | ( | $x$ | ) | $=$ | $y$ |
|---|---|---|---|---|---|---|---|
| $2^{13}\times$ | $3^4\times$ | $5^{23}\times$ | $7^{17}\times$ | $11^2\times$ | $13^{19}\times$ | $17^{15}\times$ | $19^4$ |

In order to avoid too difficult technical tools, we say that Q is strong enough to capture any primitive recursive function (p.r.). First of all, p.r. functions are a subset of the recursive ones and we define them as follows:

1. The initial functions:

$$S(x) \text{ (Successor)}$$
$$Z(x) = 0 \text{ (Zero)}$$
$$I_i^k(x_1, \ldots, x_k) = x_i \text{ for each } i, \text{ and for each } k,\ 1 \leq i \leq k$$
$$\text{(Projection or Identity function)}$$

2. If $f$ can be defined from the p.r. functions $g$ and $h$ by composition (namely, by substituting $g$ into $h$), then $f$ is p.r.

3. If $f$ can be defined from the p.r. functions $g$ and $h$ by primitive recursion, then $f$ is p.r. This means that, if[21]:

$$f(\overline{x}, 0) = g(\overline{x})$$
$$f(\overline{x}, S(y)) = h(\overline{x}, y, f(\overline{x}, y)$$

then $f$ is defined from $g$ and $h$ by primitive recursion.

4. Nothing else is a p.r. function.

This means that $\mathsf{Q}$, given a p.r. function $f : \omega \to \omega$, there is an open wff $\varphi(x, y)$ such that:

(i) $f(n) = m$ iff $\mathsf{Q} \vdash \varphi(\mathsf{n}, \mathsf{m})$

(ii) $\mathsf{Q} \vdash \exists! y\ \varphi(\mathsf{m}, \mathsf{y})$

Notice that, since our coding function is recursive too, $\mathsf{Q}$ captures it with some formula $\varphi(x, y)$, however, «$\mathsf{Q}$ is so weak that it can prove almost nothing *about* this formula»[22].

The same applies to deductions, which have to be considered as sequences of sentences $\sigma_1, \sigma_2, \ldots, \sigma_n, \sigma_{n+1}$, where $\sigma_{n+1}$ is the conclusion of the derivation. Now, the Super Gödel Number of a proof or deduction is coded as follows:

$$\pi_1^{\ulcorner \sigma_1 \urcorner} \times \pi_2^{\ulcorner \sigma_2 \urcorner} \times \cdots \times \pi_n^{\ulcorner \sigma_n \urcorner} \times \pi_{n+1}^{\ulcorner \sigma_{n+1} \urcorner}$$

With this subtle methodology, Gödel constructed his famous undecidable sentence, that is a sentence that has no proof nor refutation in $\mathsf{Q}$ and which shows us that $\mathsf{Q}$ is incomplete. Consider the property of "being a proof of $x$". To be more precise, the property which tells us if there is a super g.n. that numbers, according to our coding schema, a proof in $\mathsf{Q}$. Call this the "Provability predicate": $\mathsf{Prov(x)} =_{\text{def}} \exists \mathsf{vPrf(v, x)}$. Now we define the following wff[23]:

$$\mathsf{U(y)} =_{\text{def}} \neg \exists \mathsf{xPrf(x, \ulcorner y \urcorner)}$$

Consider the open wff $\mathsf{U(y)}$ and construct its "diagonalization". In order to do this job we have to find the g.n. of $\mathsf{U}$, namely $\ulcorner \mathsf{U} \urcorner$, and substitute it ($\mathsf{U}$'s g.n) for the free variable $\mathsf{y}$ in $\mathsf{U}$:

$$\mathsf{G} \equiv \exists \mathsf{y}(\mathsf{y} = \ulcorner \mathsf{U} \urcorner \wedge \mathsf{U(y)})^{24}$$

Let's call the "diagonal" wff $\mathsf{G}$. By some further specification we can see that the foregoing sentence is equivalent to the following:

$$\mathsf{G} \equiv \neg \exists \mathsf{xPrf(x, \ulcorner U \urcorner)}$$

That is, "$\mathsf{G}$ is true if and only if it is unprovable".

[21] We allow abbreviations of the following form: $x_1, \ldots, x_n =_{\text{def}} \overline{x}$

[22] Button and Walsh 2011, p. 131.

[23] We allow the abbreviation of "well-formed formula(s)" by writing "wff(s)".

[24] This wff is equivalent to the construction Gödel himself used in 1931: $\mathsf{G} \equiv \mathsf{U}(\ulcorner \mathsf{U} \urcorner)$. The capital letter $\mathsf{G}$ indicates that we're presented with an "undecidable Gödel sentence".

By inspection, hence, if Q proves G, than it would prove something false, which is not the case. Recall, indeed, that "G is true iff it is not provable", or, in other terms, that "G is true just in case no Super Gödel number numbers its proof". Since Q proves just correct theorems (Soundness Theorem), and since G is true, then G itself has not a proof in Q. But, since G is true, ¬G must be false and therefore must not have a proof in Q too. This situation leads us immediately to Gödel's First Incompleteness Theorem, namely:

**Theorem 1.1.1** (First Incompleteness Theorem)**.** No consistent, computably enumerable theory which interprets Q is arithmetically complete. ■

An arithmetical theory[25], which is consistent (no contradiction ca be derived in it) and computably enumerable (there is an algorithm which enumerates the theory's sentences), is *incomplete* iff $T \nvdash \varphi$ and $T \nvdash \neg\varphi$. This is Gödel's most famous result and, indeed, much of his philosophy of mathematics has been inspired by this important achievement.

Recall, now, that humans acknowledge portions of the mathematical universe by indirect perception, that is by what Gödel called *mathematical perception*. This latter is based upon two distinct kinds of evidence: the first is called *intrinsic* evidence and has no meaningful help in this situation. The second kind, instead, has the name of *extrinsic* evidence, since it's main concerns are the "consequences" that theories have and could be helpful in explaining Gödel's view. Now, from a philosophical perspective, we may ask ourselves: what is the suggestion that the *incompleteness* result for formal arithmetical theories is delivering? Gödel's own answer was given during his *Gibbs Lecture*, where he tried to sum up the major philosophical connections between the Incompleteness theorems and the foundations of mathematics. Indeed:

> [...]the following disjunctive conclusion is inevitable: Either mathematics is incompletable in this sense, that its evident axioms can never be comprised in a finite rule, that is to say, the human mind (even within the realm of pure mathematics) infinitely surpasses the powers of any finite machine, or else there exist absolutely unsolvable diophantine problems[26].

According to Gödel's conception, clearly, the second option was not tenable: there are not *absolutely unsolvable* mathematical problems, since human minds have an always more specific and detailed access to the realm of mathematics. Human minds always surpass the finite ability of a computer[27] and, in this sense, mathematics is *incomplete*

---

[25]We've established Gödel's First Incompleteness Theorem by adopting Q. This is not actually necessary, since the incompleteness result applies to *any* consistent and computably enumerable formal theory.

[26]Gödel 1951, p. 310.

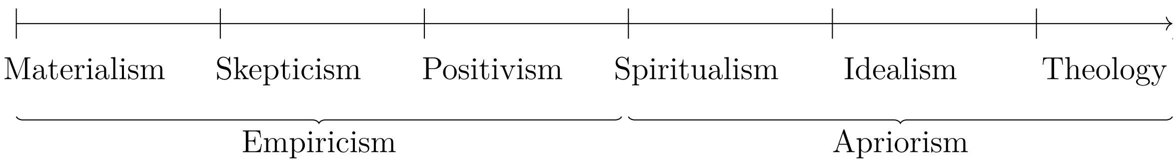
[27]Consider that Gödel was writing his essay during the 50s, therefore his argument was directed against the predecessor of our computer, namely the Turing machine.

and *incompletable*. There cannot be a finite rule which captures all mathematical truths, since human minds are infinite, that is, humans have capabilities that can bring them to discover always new mathematical truths. In this sense, indeed, our mathematical perception is *incomplete*:

1. We perceive just parts of the realm of abstract mathematical objects;
2. Human capabilities are infinite[28];
3. New mathematical perceptions can always be obtained.

This situation, for Gödel, showed two fundamental aspects of his position: (i) everything in mathematics, in principle, has or will have a solution and (ii) our knowledge of the mathematical universe can always be deepened. Indeed, in the case of the Incompleteness Theorems, which are also called *limitative theorems of arithmetic*, we're presented exactly with this: suitable logical systems and reasoning are not enough, in order to get a "complete" system of all arithmetical truths, and, therefore, arithmetic should be considered *inexhaustible* and not as enclosed in an artificial system[29].

In this spirit, Gödel provided also his own interpretation of the historical development of mathematics and mathematical practice, trying to show how an *idealistic* conception, as opposed to a *materialistic* one, of mathematical entities could be more useful to a mathematicians work. To see how did Gödel understand the development of mathematical reasoning during the centuries, it is important to consider different philosophical approaches to mathematics and systematize them according to their different degrees of affinity with metaphysics. We have to imagine a right-oriented line which divides the philosophical convictions about mathematics in the following manner:




According to Gödel, so, the philosophy of mathematics has had different interpretations, in particular each of the above mentioned philosophical conceptions tried to develop a tenable account of mathematics. Despite Gödel's platonistic spirit, mathematics, from the Renaissance, has tended to the left direction of the arrow and the major consequence has been that many philosophers had tried to develop empiric philosophical accounts of mathematics itself. Gödel, instead, insisted that «mathematics by its nature as an a priori science, always has, in and of itself, an inclination toward the right [...]»[30].

[28] «Human reason can, in principle, be developed more highly (through certain techniques).» And, «[t]here are systematic methods for the solution of all problems (also art, etc.)», Wang 1996, p. 316.

[29] At this point of the discussion, the links between "realism" and "rationalism" – fundamental for Gödel's perspectives – are explicit.

[30] Gödel 1961, p. 377.

As we will discover soon, exactly the metaphysical foundation of Platonism and the appeal to mathematical intuition lead Gödel's view into troubles. In particular, it is interesting to notice that the major problems connected to this version of Platonism are not only epistemological, but also ontological. Let's think for a moment to the following doubt: assumed that it is almost *impossible* to justify our knowledge of abstract mathematical entities, how can we, then, be sure and affirm that they *exist*? In more logical terms: how are we supposed to claim that the existential quantifier ($\exists$) of our theories "really" ranges over a domain of abstract objects, if we cannot have any clear perception of them?[31]

## 1.2 Against Platonism I: Metaphysics

### 1.2.1 Numbers are not sets! Benacerraf's Argument

Does really a mathematician work with abstract objects? Is it possible to be a Platonist while constructing new mathematical theories? In some articles[32], Paul Benacerraf proposed a challenge to platonistic philosophers and argued that mathematics is not to be considered as an incomplete description of the realm composed by particular abstract entities[33].

For the sake of the argument, suppose that two young adolescents studied the basic notions of arithmetic in a particular way: they were told that each natural number greater than 1 could be identified with a *set*[34]. Therefore, each one of the young logicians, for example, considered the number "2" as the set composed by two "things", the number "27" as the set containing 27 "elements", and so on. In other words, each adolescent considered arithmetic as a branch of set theory (the mathematical study of sets) and, therefore, numbers were just the elements of the $\omega-$set[35]. So, in the case of number theory, they earned «the numbers merely involved, learning new names for familiar sets. Old (set-theoretic) truths took on new (number-theoretic) clothing»[36]. They will say that the statement "there are $n$ $F-$things" means that "there is a relation which associates one-to-one the $F-$things

[31]We've called Gödel's position "Extreme Platonism" for two main reasons: (i) the parallelism between sciences and mathematics and (ii) the introduction of the notion of mathematical intuition. In any case, there are also "Moderate" versions of Platonism, which we'll encounter and describe in the next paragraphs.

[32]Benacerraf and Putnam 1983, Benacerraf 1965 and Benacerraf 1973.

[33]In this section I'll explain the philosophical assumptions of Benacerraf's criticism (See also Plebani 2011) In the next paragraphs we'll observe the same results in a precise formal way.

[34]Since the aim of this chapter is to give just an informal description of Benacerraf's argument, we will consider sets as simple "collections of things".

[35]We call $\omega$ the first transfinite ordinal, which has to be distinguished from $\aleph$, which is the first transfinite cardinal. An ordinal number indicates simply the well-ordering of a set, that is, given the first element of the sequence, it is always possible to reach the element immediately succeeding it and going up to $\omega$. A cardinal number, instead, simply states the number of elements (the cardinality or magnitude) of a set: $\aleph_0$ is to be considered the number stating the cardinality of the set of the natural numbers, $\mathbb{N}$.

[36]Benacerraf 1965, p. 273.

to the numbers starting from 1". So, they will be able to define and apprehend "addition", "multiplication", "exponentiation", and derive the basic arithmetical rules, namely arithmetic's fundamental laws (the so-called Dedekind-Peano Axioms for Number theory). Now, as the development of mathematics has stated, there is not one correct set theory, but there are different ways to *describe* the set-theoretical universe. Indeed, consider now that the two adolescents have taken two different set-theory classes: one has been told about Zermelo's theory and the other one has studied von Neumann's proposal. Consider, in addition, that both set theories satisfy the prerequisites for a reduction of arithmetic to logic, but they differ in the *way* this reduction is done. The boy considers von Neumann's construction and for him, for instance, the number "two" will graphically be as follows: $2 = \{\varnothing, \{\varnothing\}\}$. The girl, instead, apprehended Zermelo's reduction and proposed the following use of sets: $2 = \{\{\varnothing\}\}$. The contrast raises when we consider the "membership relation, $\in$". For the boy, clearly, $0 \in 2$ since $2 = \{\varnothing, \{\varnothing\}\}$. Differently, according to the girl's use of set theory $2 = \{\{\varnothing\}\}$, and hence $0 \notin 2$.
How it is possible that, even if both set theories – Zermelo's and von Neumman's – allow the reduction of numbers to sets, they do not agree upon a question concerning the primitive relation of their languages, namely $\in$? Maybe, one could still claim that both reductions are correct, that something as *the* number theoretical or *the* set theoretical universe do not exist, and assert that, no matter how big, the mathematical realm is enormous. But actually this poses a big philosophical problem: if we agree that the mathematical universe is very vast and we consider that the number 2 is $2 = \{\{\varnothing\}\}$ and $2 = \{\varnothing, \{\varnothing\}\}$, aren't we saying that one single "element" (i.e., 2) is identical to two different sets? One number can, of course, not be both sets and, more generally, it is not possible for one single element being identified with two other and different objects. Hence, both reductions must fail and the two young logicians must – according to Benacerraf – be convinced that numbers are not sets.

### 1.2.2 Metaphysical Misunderstandings

According to Benacerraf, we should extend «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 other (not already known to be a number)»[37]. The whole situation has arisen from the mistaken assumption that ascribed numbers other properties, different from the arithmetical ones they necessarily have. These latter properties are the fundamental and unique properties we have to consider while speaking of the *nature* of numbers and could be defined as the properties they have just in virtue of their being arranged in the $\mathbb{N}-$progression:

> To *be* the number 3 is no more and no less than to be preceded by 2, 1 and possibly 0, and to be followed by 4, 5, and so forth. And to *be* the number 4 is no more no less than to be preceded by 3, 2, 1, and possibly 0, and to be followed by… *Any* object

[37]Benacerraf 1965, p. 290.

> can *play the role of* 3; that is any object can be the third element in some progression. What is peculiar to 3 is that it defines the 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.[38]

The two young logicians example and this metaphysical claim, are useful to explain what Benacerraf's point is. First of all, it is possible to reconstruct the argument's form into two ways, *restricted* or *general*. Let's consider the first way:

(P1) Numbers have just arithmetical properties.

(P2) No set has just arithmetical properties.

(C) Numbers are not sets.

So, as we have seen numbers are defined uniquely by the properties their bear standing into the number-series and, therefore sets, that have non-arithmetical properties, cannot be identified with numbers. For the sake of the argument, consider the following questions: are there types of objects which have only arithmetical properties, except the numbers themselves? According to Benacerraf, the correct answer is "no":

> [...] numbers are not objects at all, because in giving the properties (that is, necessary and sufficient) of numbers you merely characterize an *abstract structure* – and the distinction in the fact that "elements" of the structure have no properties other than those relating them to other "elements" of the same structure.[39]

Becarraf's slogan is, therefore, numbers have only *structural properties*, namely those properties they have thanks to their standing into a specific arithmetical *structure*. Additionally, but without paying too much attention on it, Benacerraf defined the notion of *abstract structure* as a "system of relations" and, indeed, only «when we are considering a particular sequence as being, not the numbers, but *of the structure of the numbers* does the question of which element is, or rather *corresponds* to, 3 begin to make any sense»[40]. In this spirit, we can rewrite the previous argumentation and obtain its more general version :

(P1) Numbers have just structural properties.

(P2) No object has just structural properties.

(C) Numbers are not objects.

From an ontological point of view, therefore, Benacerraf is trying to operate a sort of "ontological reduction": mathematical entities are not objects independent from a mathematician's work but, instead, they are dependent from the structure in which they are inserted. There is no special abstract object which is a particular natural

[38] Benacerraf 1965, p. 291.
[39] Benacerraf 1965, p. 291.
[40] Benacerraf 1965, p. 292.

number, such as 3, but, for instance, there is a "system of relations" we obtain by elaborating the "less-than" relation, in which 3 represents the relations (preceding 2,1 and 0 and succeeding 4, 5, ...) that any third member of a progression has in virtue of its being part of the progression itself.
From an epistemological point of view, with this characterization, it is not necessary to set up an account of mathematical intuition or perception which should help us in trying to understand how we can apprehend numbers. According to Benacerraf, indeed, mathematics is the study of *abstract structures* (i.e., "systems of relations") and so his epistemological account we'll be different from the gödelian one. In this context, we have just to focus on the structure of the mathematical theory we want to study and consider just the appropriate relations: for instance, «[a]rithemtic is the science that elaborates the abstract structure that all progressions have in common merely in virtue of being progressions. It is not a science concerned with particular objects [...] there is no unique set of objects that are the numbers. Number theory is the elaboration of *all* structures of the order type of the numbers. The number words do not have single referents»[41].
This kind of objection is the one we've called metaphysical since it tries to reduce the *nature* of the mathematical objects to their "being part" of a *structure*. The consequence is strictly ontological: the structure has *ontological priority* than the objects and the relations it contains, that are, therefore, *ontologically dependent* from the first one. There is no platonic heaven to describe, instead there are many systems of relations that realize *abstract structures* and which make sense of our ordinary assertions. If we speak of *the* number three we are not picking any particular abstract entity out of the domain of natural numbers, instead we are simply considering *any* object which could occupy the third position in a structure, that is the position immediately preceding 4 and immediately succeeding 2 in a well-ordered progression. Therefore, any $n$ belonging to the $\mathbb{N}-$sequence is ontologically dependent by the progression itself, while the $\mathbb{N}-$sequence itself is ontologically independent from the elements it contains. Hence, any collection which is infinite, contains specific relations, has a smallest element, etc., is an *abstract structure* for the number-theoretical progression, no matter whether it contains the signs "1", "2", ....

## 1.3 Against Platonism II: Epistemology

### 1.3.1 Benacerraf-Field's Dilemma

In this section, we'll consider the epistemological issues arisen from platonistic philosophies, such as Gödel's[42]. First of all, Benacerraf wrote his epistemological considerations in an article which was originally devoted to discuss the notion of "mathematical truth" and is quite difficult to access[43]. In any case, Benacerraf

[41] Benacerraf 1965, p. 292.
[42] See Benacerraf 1973, Plebani 2011, Linnebo 2009c.
[43] For clear reconstructions see Linnebo 2006 and Yap 2009.

divided the discussion into his paper in two separate parts: in the first part we're presented with some semantical considerations regarding the notion of mathematical truth[44], the following parts, instead, are devoted to the philosophy of mathematics, in particular against Platonism again. Benacerraf wrote:

> [My concerns are:] (1) the concern for having a homogeneous semantical theory in which semantics for the propositions of mathematics parallel the semantics for the rest of the language, and (2) the concern that the account of mathematical truth mesh with a reasonable epistemology.[45]

This time, Benacerraf's criticism is, thereby, moved by epistemological issues. In the paragraph titled "Knowledge"[46], Benacerraf explained the epistemological theory he would have assumed in order to develop his argument, i.e., the *causal theory of reference*. Suppose we have an agent $X$ and some statement $p$, we say that, for $X$ to know that $p$ is true requires that there is relation between the referents of $p$ (names, predicates and quantifiers) and agent $X$. On this account, for example, I'll know that the statement "The table in front of you is brown" is true, when the concrete brown table in front of me will have same *causal relation* with me, that is when I'll see it, when I'll touch it, and so on. Roughly, if the objects to which I stand in causal relation confirm me what the statement is saying, making it true, then I'll get an explanation or description of the world: «[t]he proposition $p$ places restrictions on what the world can be like [...] In brief, in conjunction with our knowledge, we use $p$ to determine the range of possible relevant evidences»[47]. By adopting a causalist position in epistemology, Benacerraf formulated his criticism and the argumentation can be given as follows:

(P1) For $X$ to know that $p$, $X$ must stand in a causal relation with the referents of $p$.

(P2) No human subject stands in causal relation to mathematical abstract objects.

(C) No human subject can know a proposition whose referents are mathematical abstract objects.

As before, Benacerraf's position is anti-gödelian and refuses the identification of mathematical entities with abstract objects. Since the previous argumentation was based upon a very delicate epistemological theory, much criticism has been deserved to it and, indeed, in his famous book *Realism, Mathematics and Modality*[48], H. Field presented another very interesting version of the dilemma. Field's challenge is

[44]In Benacerraf's article we're presented with two problems: «the first is that the semantics for mathematical propositions ought to mirror the semantics for the rest of natural language, in the sense that truth conditions for both should be similar; the second is that there must be a reasonable epistemology accompanying the account of truth», Yap 2009, p. 159.

[45]Benacerraf 1973, p. 403.

[46]Benacerraf 1973, pp. 412–414.

[47]Benacerraf 1973, p. 413.

[48]Field 1989, pp. 230–239 and Linnebo 2009c, pp. 17–18.

always directed to undermine Platonism, but not on the basis of "causality", as did Benacerraf, but starting from a new notion, namely from that of "reliability". First of all, let's understand why "causality" is not the best way to formulate the dilemma and why, according to Field, it would be better to embrace the notion of *reliable*:

> Benacerraf formulated the problem in such a way that it depended on a casual theory of knowledge. The present formulation does not depend on *any* theory of knowledge in the sense in which the causal theory is a theory of knowledge: that is, it does not depend on any assumption about necessary and sufficient conditions for knowledge. [...] we do need – and do have, at least in outline – an explanation of between the facts about electrons [numbers] and our "electron"["number"] beliefs (i.e., the beliefs we would express using the word "electron"["number"]).[49]

In other words, Field argues that in order to find a plausible explanation of human knowledge of abstract entities – as postulated by physical and mathematical theories – we have to, firstly, consider the "reality" to which they belong to and then, consequently, noting if their introduction helps us in forming "true" beliefs about that "reality" itself. In this sense, roughly speaking, if we find a meaningful explanation on how (mathematical) abstract objects are helpful to our minds in generating "true" beliefs or ideas of the portion of the (mathematical) universe they are supposed to describe, then Platonism will turn out as epistemologically tenable. So, generally, from the quote above, it can be implied that, according to Field, a fine-grained Platonist should provide an adequate account of reliability[50]. This latter is generally based upon the agreement into the scientific community, that is: if a mathematician accepts something about, for instance, the number series, we trust him, since his mathematical knowledge is for sure very deep and we feel reliable in his faculties. We can state the following claim as follows:

$(Rel_{\mathrm{m}})$: $\forall S$(mathematicians accept $S \rightarrow S$ is true)

Field's example is the following: assume that in some cases, at least in outline, we have an almost satisfying explanation of reliability for our beliefs in some physical entities or phenomena. For instance, we have some beliefs about the existence and nature of "electrons", based on empirical observations, that have some influences and consequences upon our brain. In other words, this means that our brains receive "inputs" that force us in forming "true" beliefs upon the external world. According to Field, a Platonist philosopher, in believing in a sort of "other reality", thinks that reliability in mathematical statements can be explained in the same way. That is, $(Rel_m)$ can be explicated in terms of:

$(Rel_{\mathrm{p}})$: $\forall S$(physicists accept $S \rightarrow S$ is true)

[49] See Field 1989, pp. 232–233. In the foregoing passage, we've inserted the term "number", beside the the word "electron", present in the original text, to clarify how Field's objection applies also to mathematical entities.

[50] See especially Plebani 2011, pp. 126-128 and Linnebo 2006, pp. 548-553.

Here is exactly the problem : if physicists, or scientist in general, are reliable in the sense that their theories give us an intuitive idea of the external world; then, mathematicians, in which sense are reliable, if there is no concrete correspondence between us, mathematical theories, abstract objects and the external world? In Field's view, a Platonist would argue that the parallelism between concrete objects and mathematical entities is correct but just "heuristic", and that the way in which we form true mathematical beliefs is different, even if parallel, to that belonging to sciences. Let's explain this important point: if scientific theories and their correspondence with the concrete reality are the starting point to justify why we feel reliability in scientists, then, in mathematics, the consistency of a theory is the first step in justifying reliability in mathematicians' work. Recall that for Gödel-like Platonists a consistent mathematical theory, that is a theory in which no contradiction can be obtained, corresponds to a correct, although incomplete, description of the universe of abstract and existing mathematical entities. But is it really sufficient to affirm that our reliability depends on the "personal" logical faculties of a mathematician? Moreover, the consistent theories that mathematicians produce are the *real* descriptions of the mathematical universe? More generally: are Platonist's explanations and definitions of reliability – as based on the mathematicians' abilities to choose the correct and consistent theory – satisfying? Field clearly states that Platonists' explanations of reliability are not enough: there is difference in establishing that reliability, as in the case of scientific theories, is based upon the intuitive ideas that we form in us about the concrete reality, from saying that the reliability is derived from consistency, as for mathematical statements. In the first case, we have an almost "clear" description of the world, while for the second, it is not clear why consistency allows us to produce "true" beliefs concerning the "third realm" – assumed that we are not in direct touch with that realm itself. Likewise,

> [...] it is impossible to give a scientific explanation of mathematical reliability claim. Since mathematical objects do not participate in the causal order, [($Rel_{\mathrm{m}}$)] clearly cannot be explained in the same way as [($Rel_{\mathrm{p}}$)]. [...] According to Field, this radical separation of platonic entities from our physical universe makes it impossible to give any kind of explanation of [($Rel_{\mathrm{m}}$)][51].

In any case, even if Field's argument is also directed against the parallelism between sciences and mathematics, he assumes the parallelism in order to show that mathematical reliability cannot be successfully be explained in terms of physical reliability. Therefore, concludes Field, there cannot be a justification *at all* of mathematical reliability in terms of scientific or empirical reliability.

It is clear that one could criticize Field's attempt to search scientific explanation of mathematical reliability, by saying that mathematics is very different from sciences and, indeed, it is useful to point out and to consider that, «Field's challenge fails as an *objection* to mathematical platonism. But this failure does not undermine its

[51]Linnebo 2006, pp. 552–553.

force as a challenge»[52]. In this spirit, according to Field's *challenge*, the situation a Platonist faces – slightly modifying Benacerraf's dilemma – seems to be the following:

(P1) Mathematicians are reliable, in the sense that for almost every mathematical sentence $p$, if mathematicians accept $p$, then $p$ is true.

(P2) For belief in mathematics to be justified, it must at least in principle be possible to explain the reliability described in (P1).

(P3) Platonism cannot explain reliability.

(C) Platonism is not tenable.[53]

Field's challenge against Platonism is so complete. We've seen that the parallelism between physical bodies/knowledge - mathematical entities/knowledge is wrong if based upon the parallelism between the notion of reliability for sciences and for mathematics. According to Platonists, a consistent theory, considered the description of a portion of the mathematical universe, could produce true beliefs of the abstract objects it describes. Therefore consistency explains reliability in mathematical theories and works. For Field, instead, consistency cannot be the unique criterion to explain reliability, since no clear explanation can be given of the link between mathematical consistent theories, the universe they are supposed to describe and our way to generate true mathematical beliefs with respect to that universe. Moreover, argues Field, physical theories, unlike mathematical ones, have a more clearer and preciser notion of what counts as evidence and, hence, – at different degrees of certainty – physics, for example, truly describes portions of the natural world. Differently, the requirement of a tenable explanation of mathematical reliability has not be accomplished yet and Platonism rests without justification. In conclusion, it is important to notice that, even if, Field's starting point has been Benacerraf's dilemma, it was directed not only to the difficulty arisen from the abstract nature of mathematical objects, but also to the fundamental assumption that mathematics describes a non-physical "reality":

> [...] whatever mathematical objects are, truths about them must be knowable. Any account of their nature which fails to explain how that could be, is an adequate one. And according to both Field and Benacerraf, mathematical Platonism is inadequate for that very reason. Platonism, argues Benacerraf, is a view on which truth conditions on the statements of mathematics are given in terms of objects whose nature places them beyond our cognitive reach. Regardless of how well those truth conditions harmonize with a semantics for the rest of the language, they are useless if we could never know that they have been met.[54]

[52] Linnebo 2006, p. 553.
[53] «If these three premises are correct, it will follow that mathematical platonism undercuts our justification for believing in mathematics», Linnebo 2009c, p. 17.
[54] Yap 2009, p. 162.

### 1.3.2 Another Challenge: What is Mathematical Doxology?

There is another challenge that has been formulated by Vann McGee[55] and that is useful to consider:

> The problem is sometimes posed as a problem in mathematical epistemology: How can we know anything about mathematical objects, since we don't have any casual contact with them? But to put it as a problem in epistemology is misleading. The problem is really a puzzle in mathematical doxology: Never mind knowledge, how can we even have mathematical beliefs? Mathematical beliefs are beliefs about mathematical objects. To have beliefs about mathematical objects, we have to refer to them, we have to pick them out; and there doesn't appear to be anything we can do to pick out the referents of mathematical terms[56].

So, according to McGee's perspective, there is an even more important reading of Benacerraf's challenge, that is the interpretation concerning the *doxological* questions about mathematics. McGee argues that, if we put Benacerraf's problem in epistemological terms, then we are seeking for an explanation of the way we may achieve mathematical knowledge starting from mathematical objects. But – and here is the crucial point – this is just the second step in our epistemological foundation of mathematics, indeed, before treating the epistemic connections between mathematical abstract objects and us, we must find a faithful explanation of the kind of "reference" invoked by mathematicians. For clarity consider the following example. Let's take the number word "14". In order to establish the most accurate epistemological account, it should be considered if our sentences, such as "14 is a natural number", really refer to an abstract object or not. Indeed, in order to do this job it's worthy to spend some considerations on the (problematic) notion of mathematical reference. So, Benacerraf, according to McGee, has not just formulate the so called "epistemological access problem", but has proposed one of the most puzzling problems in the philosophy of mathematics. It can be posed as follows:

- *Doxology*: How do we pick out (refer to) mathematical objects?
- *Epistemology*: How do we acquire and possess mathematical knowledge?

Hence, for example, from a doxological point of view, a Platonist could be asked:

(D1) How do we refer and pick out mathematical abstract objects?

For McGee, a Platonist will generally be unable to provide tenable answers to the doxological questions and, therefore, he will not able to set up a consistent and "moderate" epistemology of mathematics.

In conclusion it might be said that, for the moment, that, following Button and

[55]McGee 1993 (published in Armour-Garb and Beall 2005, pp. 111–142) and Button and Walsh 2011, pp. 145-146.
[56]McGee 1993, p. 135.

Walsh, we've called Gödel's philosophy of mathematical knowledge "extreme" since it postulates a human faculty that guarantees the access to, and justifies the existence of, the platonic heaven of abstract mathematical entities. The doxology and epistemology we're seeking for are better called "moderate", since our aim is to study mathematical knowledge without involving particular neural processes or fantastic human faculties. Indeed, we think that Putnam's suggestion for which the «appeal to mysterious faculties seems both unhelpful as epistemology and unpersuasive as science»[57] should be taken seriously[58]. So, to avoid this kind of "psychologistic" appeal to an almost inexplicable human faculty, the best way we have to understand mathematical statements (and their philosophical implications) is to set down a "formal theory". *Why*? Logical languages are very helpful to clarify our understanding of mathematical and philosophical reasoning and, therefore, in the following chapter we are going to analyse some notions, such as the one of "object" or "set", formally. In this sense, it could be said that we aren't seeking for a "doxology (and an epistemology) by acquaintance", but for a "doxology (and an epistemology) by description". This means exactly that we aren't trying to describe a neural cognitive process between human minds and abstract entities, instead we believe that formal languages (i) can provide good tools to describe and analyse the logic underlying mathematical statements and (ii) clarify some ontological and epistemological issues concerning mathematics itself.

[57]Putnam 1980, p. 471.

[58]Putnam defines very briefly its idea: «the "moderate" position [...] tries to avoid mysterious "perceptions" of "mathematical objects"»(Putnam 1980, p. 466), indeed, ask yourself «[w]hat neural process, after all, could be described as the perception of a mathematical object? Why of one mathematical object rather than another?»(Putnam 1980, p. 471).

# Chapter 2

# Numbers, sets and objects I. Naïve Considerations

**Overview.** As we have seen in the previous chapter, Benacerraf pointed out that the natural numbers such as "$1, 2, 3, \ldots$" may not ontologically and logically be reduced to sets. Moreover, Benacerraf extended his argumentation by affirming that every ontology of mathematics which includes any sort of abstract objects is incoherent for ontological and epistemological (-doxological) reasons. In this chapter, we'll deepen some aspects of set theory and its philosophy in order to consider Benacerraf's reductive argument from another perspective. At the end of the present chapter, we should be engaged, another time, with the question of whether something like abstract objects "exist" and whether their introduction within an ontology of mathematics may be useful in order to fix our "reference" with respect to mathematical entities.
Since in this chapter our main concern will be with the so called "naïve set theory", we will start with some historical-philosophical remarks that will be useful within the final discussion. First of all, we will consider and explain Frege's logical and philosophical attempt to provide an adequate account of "logical objects" and its fundamental conception of "extensions". Secondly, our analysis will be devoted to the pioneering mathematical and philosophical work of the two brilliant German mathematicians Richard Dedekind (1831-1916) and Georg Cantor (1845-1918).

## 2.1 Short Introduction to Frege

### 2.1.1 Frege's Logical Objects and Extensions

Gottlob Frege (1848-1925) is considered as the "inventor" of our formal logic and of its application within the study of metaphysics, of language, of mathematics and so on. In this section, we shall investigate just a part of the immense fregean work and, indeed, our reconstruction will focus only on Frege's ontological conception of "logical objects".

#### 2.1.1.1 Frege's Project and its Failure

Roughly, Frege rigidly distinguished "concepts" and "objects" and he denied that concepts were "individuals", since individuals were just objects. Additionally, before giving examples, we have to distinguish two types of relations occurring between these two types of entities[1]:

(i) "Falling under": when we say that "$\overline{x}$ falls under $\overline{y}$" we are relating some objects ($\overline{x}$) to some concepts ($\overline{y}$).

(ii) "Being in": when someone affirms that "$\overline{x}$ is in $\overline{y}$" he is relating some concepts ($\overline{x}$) to some objects($\overline{y}$).

Frege[2] thought that there are such logical objects and the list may include: truth-values, courses-of-values, extensions, numbers, directions, shapes and so on. All these abstract objects could be, according to Frege's works, reduced to "courses-of-values". The main principle governing the aforementioned reduction, namely Basic Law V, failed, since the English logician Bertrand Russell discovered that from Frege's system a contradiction could be entailed. The principle that undermined Frege's logic attempted to systematize the notion of "course-of-value of a function" and "extension of a concept". The course-of-values of a function $f$ could be considered as the set of ordered pairs belonging to $f$ itself, indeed, logically, the function outputs a result $y$ for every argument $x$ to which the function is applied: $f(x) = y$. When $f$ represents a concept, Frege called its course-of-values its "extension" and considered it as the set of all objects that fall under the concept that $f$ represents. In this sense, the extension "collects" all the objects that truly fall under the concept or, in more fregean terms, an extension collects the objects that the concept $f$ maps to the logical object representing "the true". For example, the extension of the concept "$x$ is a student in Venice" can be seen as the set or collection consisting of all those individuals which are truly students in Venice. For a more mathematical example consider the concept "$x$ is a positive even integer greater than 0 and less than 10". Hence, its extension is the set of all logical objects, namely the integers, that satisfy the condition that the concept is asserting; in this case, 2, 4, 6 and 8.
For our purpose, let's suppose that we have primitive function terms $f, g, h, \ldots$ in our formal language and that their functional applications such as $f(x), g(y), h(z), \ldots$ are allowed. In *Grundgesetze*, §9[3], Frege introduced his primitive notation for courses-of-values and extension by adopting Greek letters such as $\epsilon$ and $\alpha$. Additionally, he denoted by $\epsilon\prime$ and $\alpha\prime$, the fact that the object variables $\epsilon$ and $\alpha$ are bound in the expressions $f(\epsilon)$ and $g(\alpha)$, respectively, and that these resulting expressions indicate course-of-values:

[1] Recall that, by overlining variables, such as $\overline{x}$ and $\overline{y}$, we abbreviate $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$, respectively.

[2] My reconstruction is principally due to Zalta 2018. For Frege's preliminary works on arithmetic, logic and their philosophy, we refer to the following English translations: Frege 1879, pp. 5–82, and Frege 1884. Anyway, our main focus has been directed on and inspired by Frege 1893/1903.

[3] Frege 1893/1903, pp. 14–16.

(i) $\epsilon\!\prime f(\epsilon)$.

and

(ii) $\alpha\!\prime g(\alpha)$.

indicate, respectively, the course-of-values of the functions $f$ and $g$. Two fregean examples are the following:

(ia) $\epsilon\!\prime(\epsilon^2 - \epsilon)$.

denote the course-of-values of the function represented by the open formula:

(iia) $x^2 - x$.

Likewise, he adopted:

(ib) $\alpha\!\prime(\alpha \times (\alpha - 1))$.

to denote the course-of-values of the function represented by the following open formula:

(iib) $x \times (x - 1)$.

Importantly, then, Frege noticed that, if the objects falling under the functions $x^2 - x$ and $x \times (x - 1)$ are the same, then the extensions of those two functions are the same (and viceversa). So, formally, he noticed that

$$\forall x(x^2 - x = x \times (x - 1))$$

holds if and only if:

$$\epsilon\!\prime(\epsilon^2 - \epsilon) \equiv \alpha\!\prime(\alpha \times (\alpha - 1))^4.$$

Frege generalised the previous equivalence and, in §20[5], he embodied it within his famous and inconsistent principle, namely Basic Law V. Let's represent it formally:

**Principle 1** (Basic Law V for Functions)**.** $\epsilon\!\prime f(\epsilon) = \alpha\!\prime g(\alpha) \longleftrightarrow \forall x(f(x) \equiv g(x))$.

In current terms: the course-of-values of the function $f$ is identical to the course-of-values of the function $g$ if and only if $f$ and $g$ map every object to the same value. Going a step further, recall that the extension of a concept is the set of objects that fall under that concept and, indeed, Frege defined what does it mean for an object to be the member of an extension or set. Although Frege used the $x \cap y$ to designate the "membership relation", we will follow the more usual practice, using $x \in y$. So:

**Definition 2.** $x \in y =_{\text{def}} \exists F(y = \epsilon F \wedge F(x))$

In other words, $x$ in an element of $y$ just in case $x$ falls under a concept $F$ of which

[4]We adopt the symbol $\equiv$ to indicate the "material equivalence" between the two concepts involved.

[5]Frege 1893/1903, pp. 35–36.

$y$ is the extension. For example, it is possible to show that $1 \in \epsilon[x : x + 2^2 = 5]$, starting from the premiss $1 + 2^2 = 5$. The notation $[x : x + 2^2 = 5]1$ indicates that the property of "being an $x$ such that $x$ added to $2^2$ gives as result 5" – our $F$ in the general definition – is witnessed by 1.

**Example.** Let $[1 + 2^2 = 5]$ be $[x : x + 2^2 = 5]1$. If $[x : x + 2^2 = 5]1$ then $1 \in \epsilon[x : x + 2^2 = 5]$.

*Proof.*

(a) $[x : x + 2^2 = 5]1$ — Premise

(b) $\epsilon[x : x + 2^2 = 5] = \epsilon[x : x + 2^2 = 5]$ — Axiom of identity $x = x$

(c) $\epsilon[x : x + 2^2 = 5] = \epsilon[x : x + 2^2 = 5] \wedge [x : x + 2^2 = 5]1$ — From (a)-(b) by $\wedge$−introduction

(d) $\exists F(\epsilon[x : x + 2^2 = 5] = \epsilon F \wedge F(1))$ — From (c) by $\exists$−introduction

(e) $1 \in \epsilon[x : x + 2^2 = 5]$ — From (d), by definition of $\in$

■

Thus, given the premise that "1 falls under the concept $F$", namely $[1 + 2^2 = 5]$, one can prove that 1 is a member of the extension of the concept "being an $x$ that added to $2^2$ gives 5 as result", that is $1 \in \epsilon[x + 2^2 = 5]$. From this example it should already be clear that the number 1 represents the logical object which truly falls under the concept $F$, belonging therefore to its extension, i.e. $\epsilon F$.

Most of all work has been done and the last thing we have to observe, in order to engender the Paradox, is the special version of Basic Law V for concepts[6] and some few other derived laws:

**Axiom** (Basic Law V for Concepts). $\epsilon F = \epsilon G \longleftrightarrow \forall x(F(x) \equiv G(x))$

**Corollary 1** (Existence of extensions). $\forall F \exists x(x = \epsilon F)$

**Corollary 2** (Law of Extensions). $\forall F \forall x(x \in \epsilon F \longleftrightarrow F(x))$

In current terms, the first axiom says that the extension of the concept $F$ is identical to the extension of the concept $G$ if and only if the objects falling under $F$ fall also under $G$. In contemporary se theory, we could have said that the set of the $F$s is identical to the set of the $G$s if and only if $F$ and $G$ are materially equivalent:

[6]Without explaining at length Frege's conception, we will simply assume that, according to him, concepts are special cases of functions.

$$\{x \mid F(x)\} = \{x \mid G(x)\} \longleftrightarrow \forall y(F(y) \equiv G(y)).$$

To see that the first corollary is a consequence of BLV, notice that when we instantiate the variable $F$ to $P$ in BLV, we can establish:

$$\epsilon P = \epsilon P \longleftrightarrow \forall x(P(x) \equiv P(x))$$

Since the right side of this instance of BLV can be derived by logic rules alone, it follows that $\epsilon P = \epsilon P$. But, then, by existential generalization, it follows that:

$$\exists x(x = \epsilon P)$$

Now the first first corollary follows by universal generalization on $P$:

$$\forall F \exists x(x = \epsilon F)$$

The second corollary, the so-called Law of Extensions, asserts that an object is a member of the extension of a concept if and only if it falls under that concepts. In order to see this, consider our previous example involving the derivation of $1 \in \epsilon[x + 2^2 = 5]$ from the premise $[1 + 2^2 = 5]$.
Also with these few informations on Frege's deep work, it is possible to focus the attention on the discover of the existence of the paradoxical Russell Set within Frege's *Grundgesetze*. There are two strategies to derive a contradiction from Frege's system and we will present both of them.

**Theorem 3** (Russell′s Paradox 1)**.** BLV entails a contradiction.

*Proof.* For the moment, consider the the concept *being the extension of a concept which you don't fall under* and call it $Q$; formally:

$$Q = [x : \exists F(x = \epsilon F \wedge \neg F(x))]$$

Further, by corollary 1, we know that the extension of the concept *being the extension of a concept which you don't fall under*, exists; namely:

$$\epsilon Q$$

Now suppose,

$$\underbrace{[x : \exists F(x = \epsilon F \wedge \neg F(x))]}_{Q}(\epsilon Q)$$

That is, $Q(\epsilon Q)$. In current terms: the extension of $Q$ falls under the concept *being the extension of a concept which you don't fall under*. Considering this, instantiate the free $x$ with $\epsilon Q$ in order to get:

$$\exists F(\epsilon Q = \epsilon F \wedge \neg F(\epsilon Q))$$

Letting $P$ be such a concept, we obtain

$$\epsilon Q = \epsilon P \wedge \neg P(\epsilon Q)$$

Now, by applying BLV to the first conjunct it follows $\forall x(Q(x) \equiv P(x))$. But, since $\neg P(\epsilon Q)$, it follows $\neg Q(\epsilon Q)$, contrary to the starting hypothesis.
Reverse the argument and suppose:

$$\underbrace{\neg[x : \exists F(x = \epsilon F \wedge \neg F(x))]}_{\neg Q}(\epsilon Q)$$

That is, $\neg Q(\epsilon Q)$. Now, instantiate again $x$ with $\epsilon Q$ to get:

$$\neg\exists F(\epsilon Q = \epsilon F \wedge \neg F(\epsilon Q))$$

By simple logical transformations, this means that:

$$\forall F(\epsilon Q = \epsilon F \rightarrow F(\epsilon Q))$$

By instantiating the second-order variable $F$ with $Q$ itself, it follows:

$$\epsilon Q = \epsilon Q \rightarrow Q(\epsilon Q)$$

Hence, $Q(\epsilon Q)$. But, another time, this is contrary to the initial hypothesis.
In both case we've derived a contradiction. ■

In this way, without employing foreign strategies to Frege's *Grundgesetze* (but simply a different notation), we've derived the famous contradiction which lead Frege's dream into troubles.
Anyway, since our main concern remains set theory and its philosophy, we will now focus on the method Russell himself adopted in 1902[7]. Roughly, we will initially derive a fundamental principle governing properties (absent, but implicit, in Frege's *Grundgestze*) and which will bring us directly in front of the set-theoretical version of the paradox. Recall that Frege established his system within a second-order logic, quantifying hence not only on individuals but also on properties. In this sense, the domain on which the quantifiers range will be composed by $n-$ary predicate or relation letters $F^n, G^n, \ldots (n \geq 0)$ designating arbitrary properties. Indeed, consider that a second-order logic contains a so-called "comprehension principle for properties" that guarantees the *existence* of an $n$-place relation or property corresponding to any open formula $\varphi$ with $n$ object variables $x_1, \ldots, x_n$. We'll state it, even if Frege did not formulate it explicitly within his entire work:

**Axiom** (Comprehension Principle for Concepts)**.** $\exists F\forall x(F(x) \longleftrightarrow \varphi)$

The foregoing principle governs $1-$place relations, but it is possible to extend it and consider $n-$ary relations. So, for any schematic condition $\varphi$ there is a corresponding property $F$. Anyway, what we actually need in order to engender Russell's Paradox

[7]See Russell 1902, pp. 124–125, and Frege 1902, pp. 126–128 (both texts are collected in van Heijenoort 1967). For philosophical commentary, see, among others, Linnebo 2011, pp. 33–37.

is the version of the Comprehension Principle, not for concepts, but rather for extensions. Let's see how to derive it from the definitions and theorems we've already provided:

*Proof.*

(a) $\forall F \forall x(x \in \epsilon F \longleftrightarrow F(x))$ Given

(b) $\forall x(x \in \epsilon F \longleftrightarrow P(x))$ From (a), instantiation of $\forall F$ with $P$

(c) $\exists y \forall x(x \in y \longleftrightarrow P(x))$ From (b), $\exists-$introduction on $\epsilon P$

(d) $\forall F \exists y \forall x(x \in y \longleftrightarrow F(x))$ From (c), $\forall-$introduction on $P$ ■

This last step is exactly the principle we were seeking for. The theorem in the last step of the previous derivation is Frege's "Naïve Comprehension Principle for Extensions"; it affirms that for any concept $F$, there in an extension $y$ which collects as elements all and only those objects that fall under $F$, namely all the $x_1, \dots, x_n \in \epsilon F$. Now, consider the schematic version of the Naïve Comprehension Principle for Extensions, namely:

$$\exists y \forall x(x \in y \longleftrightarrow \varphi(x))$$

Recall that the Law of Extension, from which we derived the Naïve Comprehension Principle/Schemata for Extensions, has been proved with the help of Basic Law V. Being this latter an axiom of the *Grundgestze* system it should be considered as truth-entailing in all of its applications, that is, nothing inconsistent or incoherent should be derived from it. Let's consider, finally, what Russell's logical inquiry discovered:

**Theorem 2.1.1** (Russell′s Paradox 2)**.** Basic Law V entails a contradiction.

*Proof.* Consider the Naïve Comprehension Principle for Extensions and let $\varphi = y \notin y$, that is the "not-membership" relation. Hence,

$$\exists y \forall x(x \in y \longleftrightarrow y \notin y)$$

Let $a$ be such an $y$ to get:

$$\forall x(x \in a \longleftrightarrow a \notin a)$$

Since $a$ is in the scope of the universal quantifier we obtain:

$$a \in a \longleftrightarrow a \notin a$$

Contradiction. ■

As it is clear, the last sentence that have we derived in the foregoing proof entails a contradiction. If $(a \in a \longleftrightarrow a \notin a) = (a \in a \longleftrightarrow \neg a \in a)$ then, using a simple propositional logic vocabulary, it is easy to see that the logical form of Russell's conclusion is $A \longleftrightarrow \neg A$. Recall that the Naïve Comprehension Principle/Schemata for Extensions has been derived from the Law of Extensions, while the latter was implied by Basic Law V. In this sense, all the theorems which involved the application of Basic Rule V, implicitly used an inconsistent rule, being thus incorrect. All the strategies Frege developed in order to derive the Dedekind-Peano axioms for arithmetic from his logical system failed since the derivation of the main principle governing the existence of numbers as abstract logical objects applied to Basic Law V. This principle is nowadays known as "Principle of Hume" and its introduction would have allowed Frege to show that arithmetic is reducible to logical truths alone.

What have we achieved up to now?

Recall that our main purpose in this chapter is to deepen our set-theoretical knowledge (and its philosophical foundations) in order to discuss Benacerraf's ontological reductive argument. Let's summarize our main results:

1. Frege developed an interesting theory of extensions: pick a property $F$ and analyse its course-of-values, thus determining its extension, that is all the objects that have the considered property.

2. Frege's theory of extensions is considered as a "naïve set theory", which is to contrasted with its axiomatic version. This means that we do not have all precise rules to form and define sets, but every collection of elements is considered as a set itself.

3. Frege's ontology of mathematics is full of abstract objects (such as extensions and numbers), each of them governed by a logical principle (such as Basic Law V and Hume's Principle).

Now, Frege was not interested in the "mathematics" of sets but, rather in its philosophical foundation and reduction to logic. Indeed, some years before Frege's works appeared, two brilliant German mathematician had independently developed some interesting and almost pure mathematical considerations concerning sets, namely R. Dedekind and G. Cantor. Even if their contributions to the early development of naïve set theory and to debate concerning the foundations of mathematics are fundamental, their theories of sets are, in any case, exposed to Russell's Paradox.
In the next section, while discussing some issues concerning sets and their elements, we will present our "naïve" ontological objection to Benacerraf's reductive arguments exposed in "What numbers could not be" (1965).

#### 2.1.1.2 Logicism, Extensions and Natural Numbers

Frege's main purpose was to derive *all* the basic truths for number theory from a small package of logical concepts and axioms. In particular, Frege aimed to give clear proofs – into his *Grundgestze*'s system – of the so called Dedekind-Peano axioms for natural numbers. Let's state them in ordinary and formal language ($xSy$ is the formal version of "$y$ is the successor of $x$" or of "$y$ is preceded by $x$", and $\mathbb{N}(x)$ indicates that $x$ is a natural number)[8]:

(**A1**) 0 is a natural number

$$\mathbb{N}(0)$$

(**A2**) Successors of natural numbers are natural numbers

$$\mathbb{N}(x) \wedge xSy \to \mathbb{N}(y)$$

(**A3**) If a natural number is succeeded by two numbers, then they are the same number

$$\mathbb{N}(x) \wedge xSy \wedge xSz \to y = z$$

(**A4**) If a natural number is preceded by two numbers, then they are the same number

$$\mathbb{N}(y) \wedge xSy \wedge zSy \to x = z$$

(**A5**) Any natural number has a unique successor

$$\mathbb{N}(x) \to \exists! y\ xSy$$

(**A6**) Principle of mathematical induction. Let, $m, n$ be restricted variables ranging over natural numbers and $F$ be an arbitrary predicate letter:

$$\forall F[F(0) \wedge \forall m, n(F(m) \wedge mSn \to F(n))] \to \forall n F(n)$$

It's important to be precise and to say that Frege thought that arithmetic, i.e. the mathematical study of the natural numbers sequence, could be reduced to and, consequently derived from, logic alone. So, generally conceived, Frege's *logicist* project contains both, the following two logical theses:

($L_1$) All arithmetical concepts are definable in terms of, and thanks to, the presence of logical concepts.

($L_2$) All arithmetical truths are derivable from logical truths,

and the following ontological claim:

($L_3$) Logical and arithmetical objects can be thought of as particular abstract objects.

Concerning ($L_3$), Frege thought that "$1, 2, 3, \ldots$" should be considered as indicating the "number" of objects falling under some concept. So, if someone utters "there are 3 books", the number-sign 3 refers to the extension of the concept "being a book" under which fall three elements. Frege thought that, in order to affirm the existence of the "numbers as extensions", it was solely necessary to establish a general rule regarding their "equipotency". For instance, if the objects falling under the concept

[8] See Linnebo 2011, p. 33; Zalta 2018, p. 39.

"being a book" and the objects falling under the concept "being a pen" are in a biunivocal correspondence, then they have the same number of elements. Frege, indeed, thought that the notion of sameness of cardinality, or in fregean terms of "equinumerosity" between sets, is principal for his logicist account of arithmetic and, indeed, it's fundamental to state it formally. Let $F$ and $G$ be two arbitrary predicate letters and let $\sim$ be the 2−relation symbol representing equinumerosity:

**Definition 3** (Equinumerosity)**.** $F \sim G =_{\text{def}} \exists R[\forall x(F(x) \rightarrow \exists! y(G(y) \wedge xRy)) \wedge \forall x(G(x) \rightarrow \exists!(F(y) \wedge yRx))]$

In other words, that $F$ and $G$ are equinumerous means that there is a binary relation $R$ which establishes a one-to-one correspondence (bijection) between the objects that fall under $F$ and those that fall under $G$. Hence, it seems natural to reason in the following way: two concepts $F$ and $G$ will be equinumerous for definition if the "number" of elements falling under $F$ is identical to the "number" of objects falling under $G$. This latter claim will be, indeed, embodied within the so called Hume's Principle. It is possible to represent the situation we've described as follows: Recall

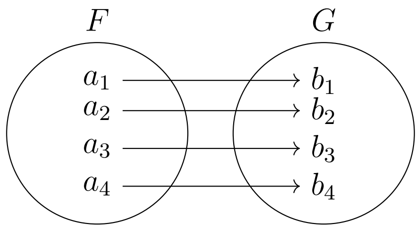



Figure 2.1: Mapping of $R$

that in *Grundgesetze*[9], Hume's Principle has been derived from his Basic Law V, determining thus its inapplicability during any proof and has not been assumed as an axiom. In any case, since it has played a fundamental role in Frege's project, let's state it adopting an additional operator, $\#F$, indicating "the number of the objects falling under $F$" (or simply, "the number of the $F$s"):

**Proposition 4** (Hume′s Principle )**.** $\#F(x) = \#G(x) \longleftrightarrow F \sim G$

In this case, the number of the $F$s is identical to the number of $G$s if and only if $F$ and $G$ are equinumerous. To see this clearly compare what Hume's Principle is asserting with the mapping presented in figure 2.1. With this framework, indeed, Frege defined what a "cardinal number" is:

**Definition 4** (Cardinal Number)**.** $Card(x) =_{\text{def}} \exists F(x = \#F)$

[9]Frege defined firstly Hume's Principle in his *Grundlagen* (see, Frege 1884, pp. 84–85). For commentary in his *Grundgesetze*, see, for instance, Frege 1893/1903, pp. 56–57.

So, a cardinal number is the logical object $x$ which is the number of some concept $F$. This considerations and, in particular, Hume's Principle are «the basic principle[s] upon which Frege forged his development of the theory of natural numbers»[10]. This characterization has been renamed "Frege's definition of cardinal numbers as equivalence classes". For with the previous formal assessment Frege has been able to derive the Dedekind-Peano Axioms for Arithmetic. In particular, in order to show the infinity of numbers Frege made use of a sequence of representative concepts such as "the concept under which fall $n$ objects":

$$\mathbb{N}_0 : [x \mid x \neq x]$$
$$\mathbb{N}_1 : [x \mid x = \#\mathbb{N}_0]$$
$$\mathbb{N}_2 : [x \mid x = \#\mathbb{N}_0 \vee x = \#\mathbb{N}_1]$$
$$\mathbb{N}_3 : [x \mid x = \#\mathbb{N}_0 \vee x = \#\mathbb{N}_1 \vee x = \#\mathbb{N}_2]$$
$$\ldots$$

In other terms, $\mathbb{N}_0$ is the the set of all not self-identical members, i.e. the set which has no elements in it and every $\mathbb{N}_i$ represents the set under which all the preceding members of the collection $\mathbb{N}_{i-1}$, $(i > 0)$ fall. Thus enabled Frege to the following characterization of finite cardinals within the *Grundgesetze*:

- $\#\mathbb{N}_0 = 0$
- $\#\mathbb{N}_1 = 1$
- $\#\mathbb{N}_2 = 2$
- $\#\mathbb{N}_3 = 3$

  $\ldots$

Although, equinumerosity and "naïve" sets will occupy this entire chapter, we will not follow Frege's formalization of extensions and sets, but we will follow the more usual practice and notation. Anyway, while considering from a closer point of view neo-fregean proposals – at least, with respect to abstraction principles – we will see that HP and BLV still play fundamental roles.

## 2.2 Short Introduction to Cantor

### 2.2.1 Naïve Set Theory: Foundations

Recall that set theory is the mathematical study of "sets", that is the study of "collections of things". In order to see, from a different perspective Benacerraf's argument, we still need new technical tools. Ontological and epistemological considerations, such as the existence and knowledge of abstract objects and the possibility to reduce the natural numbers to set-theoretical representations, will be discussed when all the preliminary, formal and historical remarks have been presented. Without

[10]Zalta 2018, p. 34.

assuming Frege's theory of extensions and of logical objects, let's begin with the pure mathematical notion of *set.* In Cantor's own word

> By an "aggregate" (*Menge*) we are to understand any collection into a whole (*Zusammenfassung zu einem Ganzen*) $M$ of definite and separate objects of our intuition or our thought. These objects are called the "elements" of $M$.[11]

The idea that a set represents a collection is perpetuated even in other modern definitions. For instance, Frege, Schoenfield and Thomason, respectively, say that:

> A class, in the sense in which we have so far used the word, consists of objects; it is an aggregate, a collective unity, of them; if so, it must vanish when these objects vanish. If we burn down all the trees of a wood, we thereby burn down the wood.[12]

> A *set* is a collection of objects[13].

> A *set* is many things considered as a unity... there are nine planets, taken severally. But taken together, as one, there is a single object: the set (*collection, multitude, aggregate, class*) of the planets[14].

The description they propose are telling us something about two fundamental properties of sets. Let $A$ be designating any set[15]:

1. **Membership relation**: Every set $A$ has a member or element. In symbols, $x \in A$ iff $x$ belongs to (or is a member of ) $A$.
2. **Extensionality Property**: Every set is determined by its members. If $A$ and $B$ are sets, then they are identical iff they have the same elements. Formally,

$$A = B \iff \forall x(x \in A \longleftrightarrow x \in B)$$

The result we are going to analyse are all based on the Extensionality Property and on the following principle, which we have encountered in different fashion already in Frege's work:

**Principle 2** (General Comprehension Principle)**.** For each n-ary definite condition $P$:

$$\exists A(A = \{\overline{x} \mid P(\overline{x})\})$$

whose members are exactly all the $n$-tuples satisfying $P(\overline{x})$, such that

$$\forall \overline{x}\ (\overline{x} \in A \iff P(\overline{x})).$$

[11] Cantor 1915, p. 85.
[12] Frege 1895, p. 212.
[13] Schoenfield quoted from Oliver and Smiley 2006, p.126
[14] Thomason quoted from Oliver and Smiley 2006, p.126
[15] The formal remarks are due to the brilliant exposition of set theory of Moschovakis 2006. For Italian readers, a very useful introduction to Cantor can be found in Costantini 2016 (Second Part).

By the Extensionality Property it follows that at most one set $A$ satisfies the condition (2.1) and we can call the set $A$ the "extension" of the definite condition $P$. Since we are working with formal methods, we want to avoid "vague" conditions, which have nothing to do with scientific reasoning. Consider, for instance, the following set:

$$A =_{\text{def}} \{x \mid x \text{ is a good citizen}\}$$

The condition of "being a good citizen" could be longly debated and therefore the membership relation is not always clearly determined. In this case, in order to avoid strange and arbitrarily defined sets, we define what is for $P$ to be a "definite condition". We say that

**Definition 5.** An $n$-ary **condition** $P$ is **definite** if for each $n$-tuple $\overline{x}$ of objects, it is determined **unambiguously** whether $P(\overline{x})$ is true or false.

In the same way, we do not want "casual" assignments within our theory, but we are aiming to define unambiguously when an object can be assigned to a determined value. This motivates the next restriction:

**Definition 6.** An $n$-ary **operation** $F$ is **definite** if it assigns to each $n$-tuple $\overline{x}$ of objects a unique, **unambiguously** determined object $w = F(\overline{x})$.

Thus, assuming that the laws of biology will not betray us, the following non-mathematical operation is definite:

$$F(x) =_{\text{def}} \begin{cases} \text{the egg of } x, & \text{if } x \text{ is an oviparous} \\ x, & \text{otherwise} \end{cases}$$

In this manner, we are ensuring that $F(\overline{x})$ will get a determined value for each $\overline{x} = x_1, \ldots, x_n$. In practice, we will not consider the values for $F(\overline{x})$ with not oviparous' $\overline{x}$ , thus we're allowed to define the operation simply as $F(x) =_{\text{def}}$ the egg of $x$.
Additionally, we need to assume the existence of an "empty set[16]" and specify the condition of "being a function" and there are «[...] no problems as mathematicians have always made these assumptions, explicitly or implicitly»[17]:

**Definition 7.** Let $A$ and $B$ be two sets. We define a **function** $f$ from $A$ to $B$ as follows:

[16]«Somewhat peculiar is the empty set ∅ which has no members. The extensionality property implies that *there is only one empty set*» (Moschovakis 2006, p. 2). Additionally, it is useful to precise that «we define the ∅ as the unique set with no members; the empty set. Whilst there are *philosophical* discussions to have about ∅'s existence, there are no *technical* discussions to be had» (Button and Walsh 2011, p. 30). For philosophical criticism see, for instance, Oliver and Smiley 2006.
[17]Moschovakis 2006, p. 21.

$$\text{Function}(f, A, B) \iff f : A \to B$$

In order to be clearer we will employ different notation when different functions are considered:

- Mappings such as $a \mapsto f(a)$ are useful while considering functions without officially naming them. For example, $x \mapsto x + 1$ (for $x \in \mathbb{N}$) is the function which associates each natural number to its immediate successor. If we "name" the function $f$, then it is defined by the formula $f(x) = x + 1$ (for $x \in \mathbb{N}$).
- $A \rightarrowtail B$ is an "injective function" or an "injection" (*one-one*):
$$\iff \forall a, b \in A \; (a = b \implies f(a) = f(b))$$
- $A \twoheadrightarrow B$ is a "surjective function" or a "surjection" (*onto*):
$$\iff \forall b \in B, \; \exists a \in A \; (b = f(a))$$
- $A \to B$ is a "bijective function" or a "bijection" (*one-to-one correspondence*):
$$\iff \forall b \in B, \; \exists! a \in A \; (b = f(a))$$

These preliminary remarks are fundamental to the understanding of Cantor's set theory. In particular, all the results we are going to analyse are centred around the few technical definitions we have briefly introduced. So, if our aim is to discuss some features regarding the "size" of sets and how big they could be, it is important to introduce some of their peculiar characteristics. As in the case of Frege's theory, let's begin by introducing Cantor's equinumerosity notion:

**Definition 8.** Two sets $A$ and $B$ are **equinumerous** or **equipotent** iff there is a one-to-one correspondence between them, written $A \sim B$.

Since we've introduced the notion of *equipotency*, we can prove one of its main particularity, namely that it corresponds to an equivalence relation. First of all recall that from a mathematical point of view a generic relation $R$ is said to be an equivalence relation iff the following three conditions can be verified:

1. For all $x$, $xRx$ (Reflexivity)
2. For all $x$ and $y$, $xRy \implies yRx$ (Symmetry)
3. For all $x, y$ and $z$, $xRy$ and $yRx \implies xRz$ (Transitivity)

Now, we can state the following result concerning the equinumerosity relation between sets:

**Proposition 5.** The euinumerosity relation is an equivalence relation.

*Proof.* Consider that our $R$ is the relation $A \sim B$. We must prove that:

(i) for each set $A$, $A \sim A$.

(ii) for each set $A, B$, $A \sim B \implies B \sim A$.

(iii) for each set $A, B, C$, $A \sim B$ and $B \sim C \implies A \sim C$.

For (i): If $A \sim A$, then there must be a bijective function $f : A \to A$. As clear $f$ is the identity function that for each $a \in A$, $a \mapsto a$: the mapping is one-one (injective) and onto (surjective), hence one-to-one (bijective). Therefore, $\sim$ is a reflexive relation.
For (ii): If $A \sim B$, then it follows that there is a bijection $f : A \to B$. Since $f$ is bijective map, then there exists also the bijection $\overleftarrow{f} : B \to A$. Hence, $B \sim A$.
For (iii): Suppose that $A \sim B$ and that $B \sim C$. By definition of $\sim$, there are two biunivocal correspondences $A \xrightarrow{f} B \xrightarrow{g} C$. Hence, the composition of $f$ and $g$ is a one-to-one function $\underbrace{A \xrightarrow{f} B \xrightarrow{g} C}_{f \circ g}$. Therefore, $A \sim C$. ■

Consider, for the moment, that the

> [...] radical element of Cantor's definition is the proposal to accept the existence of such a correspondence as the characteristic property of equinumerosity for all sets, despite the fact that its application to infinite sets leads to conclusion which had been viewed as counterintuitive.[18]

As we see from the quote above, we've encountered the notion of "infinite" set, and, in order to introduce discussions around set sizes, we need to be able to say the conditions under which a set is "less or equal in size" to another one. The next definition will very be useful:

**Definition 9.** A set $A$ is **less** or **equal in size** to a set $B$ iff it is equinumerous to some subset of $B$. Formally,

$$A \leq B \iff \exists C\big(C \subseteq B \land A \sim C\big)$$

Since in this chapter our aim is to discuss set theory "philosophically", we have to outline its fundamental results. Unlike Frege's theory, all the formalism here introduced should be considered as «a useful device which is compatible with every philosophical approach to the subject»[19]. Recall, indeed, that while discussing Frege's theory of extensions we've introduced his ontological conception of objects and concepts. Here, instead, we won't assume and defend *any* position concerning mathematics, logic and their connection to philosophy, but all the philosophical remarks will be outlined at the end of the present section.
Now, first of all, let's introduce some other useful symbols: let $\Lambda$ be a non-empty set,

[18] Moschovakis 2006, p. 7.
[19] Moschovakis 2006, p. 30.

finite or infinite, and suppose that for each $\lambda \in \Lambda$ we are given a set $A_\lambda$. We say that we have a *family* of sets $\mathcal{F} = \{A_\lambda \mid \lambda \in \Lambda\}$[20].

**Definition 10.** A **partition** of a set $A$ is family $\mathcal{F} = \{A_\lambda \mid \lambda \in \Lambda\}$ of non-empty subsets of $A$, $A_\lambda \subseteq A$, such that:

1. $A_\lambda \cap A_\mu \neq \varnothing \implies A_\lambda = A_\mu$
2. $\bigcup_{\lambda \in \Lambda} A_\lambda = A$

We say that the $A_\lambda$s partition the given set $A$.

Consider now the following non-mathematical example. Let $A = \{a, b, c, d, e, f, g, h\}$ and let $A_1 = \{a, b, c, d\}$, $A_2 = \{a, c, e, f, g, h\}$, $A_3 = \{a, c, e, g\}$, $A_4 = \{b, d\}$ and $A_5 = \{f, h\}$. It follows that:

1. $A_1 = \{a, b, c, d\}$ and $A_2 = \{a, c, e, f, g, h\}$ are not partitions of $A$, since they are not mutually disjointed.
2. $A_1 = \{a, b, c, d\}$ and $A_5 = \{f, h\}$ are not partitions of $A$ since they leave out, for instance, the element $e$.
3. $A_3 = \{a, c, e, g\}$, $A_4 = \{b, d\}$ and $A_5 = \{f, h\}$ correspond to a partition of $A$.

A more mathematical example is the following. Consider the $\mathbb{N}$ set. If we establish that $\mathbb{N}_e = \{x \mid x \div 2 = 0\}$ and that $\mathbb{N}_o = \{x \mid x \div 2 \neq 0\}$, then we've partitioned the set of the natural numbers into the partitions of the even and the odd numbers. Much, but not all, of the formal work is done. Finally, let's introduce other useful notions and achievements:

**Definition 11.** Given a generic equivalence relation $R$ on a set $A$, the **equivalence class** of an element $x \in A$ is defined as follows:

$$[x] = \{y \mid \langle x, y\rangle \in R\}$$

or

$$[x] = \{y \mid x \equiv y\}$$

The notation "$a \equiv b$" has to be read "$a$ is equivalent to $b$". The next theorem, as we will see, immediately follows:

**Theorem 2.2.1.** If $R$ is an equivalence relation on a set $A$, then the equivalence classes generated by $R$ are partitions of $A$.

*Proof.* We have to prove that:

(i) for all $a \in A$, $a$ belongs to some equivalence class generated by $R$.

[20]For technical introductions to set partitions, see, among others, Carlucci Aiello and Pirri 2005, p. XX. (for Italian speaker) and Biggs 2002, pp. 126–141.

(ii) each pair of $A_\lambda$, $A_\mu$ ($\lambda \neq \mu$) is disjoint.

For (i): Suppose $a \in A$ and that $R$ is an equivalence relation. By definition, $[a] = \{y \mid \langle a, y\rangle \in R\}$ and, hence, $a \in [a]$.
For (ii): If $[a] \cap [b] \neq \varnothing$, then there exists an element $c \in [a] \cap [b]$. Since $R$ is an equivalence relation, then it is also transitive, so, if $\langle a, c\rangle \in R$ and $\langle c, b\rangle \in R$, then $\langle a, b\rangle \in R$, namely $b \in [a]$. If we consider an element $x \in [b]$, namely $\langle b, x\rangle \in R$, by transitivity of $R$ we would get $\langle a, x\rangle \in R$, that is $x \in [a]$. Hence, $b \subseteq [a]$.
If $\langle b, c\rangle \in R$ and $\langle c, a\rangle \in R$, then $\langle b, a\rangle \in R$, namely $a \in [b]$. Always considering a general element $x \in [a]$, $\langle a, x\rangle \in R$, we get, by transitivity of $R$, $\langle b, x\rangle \in R$, that is $x \in [b]$. Hence $a \subseteq [b]$.
Finally, $b \subseteq [a]$ and $a \subseteq [b] \implies [a] = [b]$. ■

Now, in order to indicate the "bigness" of a set, let's explain the notion of "cardinality" or "magnitude". With the next two definitions, we've set up much of the formalism we need to conclude our "mathematical" section:

**Definition 12.** Given a set $A$, we call **cardinality** or **magnitude**, the equivalence class determined by $A$ with respect to the equinumerosity relation, $A/\sim$. We write $|A|$ to indicate this class. Let's establish that:

(i) $|A| = |B| \iff A \sim B \iff \exists f : A \to B$, with $f$ bijective.

(ii) $|\varnothing| = 0$.

In other words, the class $[a]$ is to be considered as the set of elements of $A$ that stand in an equivalence relation with $a$. Consider informally for a moment what we've achieved and let $n$ be a natural number. We can notice that all the sets containing exactly $n$ members belong to the same equivalence class. For the sake of the argument, consider the set $A$ containing two books and the set $B$ containing two bottles. Since $A$ and $B$ are equipotent, that is, there is a bunivocal map from $A$ to $B$, the two sets share the same cardinality. In other words, having $A$ and $B$ just two elements, the equivalence class of the magnitude of $A$, $|A|$, is identical to the equivalence class of the magnitude of $B$, $|B|$. In this case, in particular, $A$ and $B$ share the equivalence class determined by the equinumerosity relation and, indeed, $[2] = \{\{A\}, \{B\}, \dots\}$, the equivalence class of 2 is the set that contains all the sets containing just two elements (all the sets with magnitude or cardinality 2).

### 2.2.2 On Infinite and Finite Sets

As briefly sketched in the introduction to this chapter, the early developments of set theory have been characterized by the works of Frege, Cantor and Dedekind. Up to now, the reconstruction we've proposed is strictly related to that of Cantor. Additionally, recall that our final aim is to reconsider how identities between set theoretical representations of natural numbers may be conceived of. For the sake of the argument, indeed, we have to understand how the notions of finiteness and infiniteness have been brought within the discussion concerning sets. In what follows, we will care about the distinction between *finite* and *infinite* sets. In this spirit, we will first consider a Cantor-style approach:

**Definition 13** (Cantor)**.** A set $A$ is said **finite** if there exists a natural number $n \in \mathbb{N}$ such that:

$$A \sim \{z \in \mathbb{N}| \mid z < n\} \text{ and } A = \{0, 1, 2, \dots n-1\},$$

otherwise $A$ is **infinite**. Thus, the empty set is finite $\varnothing = \{z \in \mathbb{N} \mid z < 0\}$.
We say that $A$ is **countable** or **denumerable**, if it is finite or equinumerous with the set of natural numbers $\mathbb{N}$, otherwise it is **uncountable** or **non − denumerable**.

Indeed, we may denote any finite set $A$ as follows:

$$A_n = \{m \in \mathbb{N} \mid m \leq n-1\}.$$

By focusing on the previous definitions it trivially follows that:

**Remark.** If A is finite, then either A=∅ or A has an enumeration, that is a surjective function f: $\mathbb{N} \to A$ such that

$$A = f(\mathbb{N}) = \{f(0), f(1), f(2), \dots\}.$$

■

The foregoing definitions and results allow us to prove the mathematical version of the ancient principle for which "the whole is greater than its parts"[21]. In particular, we are able to show that a subset of a finite set is finite:

**Theorem 2.2.2.** Let $A$ be a finite set. If $B$ is a subset of $A$, then $B$ is also finite.

*Proof.* We have to prove that: If $A$ is finite and $a \in A$, then $A/\{a\}$ is also finite.
If $A$ is the empty set the case is trivial.

We next prove the general case by induction.

**Base case**:
If $n = 1$, then $A/\{a\} = \varnothing$ is finite.
If $n > 1$, the function $f$ is restricted to $\{m \in \mathbb{N} \mid m \leq n-1\}$ and yields bijection into $A/\{a\}$. Hence, $A/\{a\}$ is finite and has $n-1$ elements.

**Inductive step**:

[21] Historically, this principle has been found in-between Euclid's axioms and postulates for geometry.

Now we need to show that, if our theorem is valid for sets with $n$ elements, it is also true for sets with $n+1$ elements.
Let $A$ have $n+1$ elements, and $B \subseteq A$.
We have two case:

(i) If $B = A$, then we're done.

(ii) If $B \subset A$, then $\exists a \in A/B$. This means that $B = A/\{a\}$. Since $A/\{a\}$ has $n$, then it follows that $B$ is finite.

■

From this important theorem, a corollary – regarding a feature of the cardinality of subsets of finite sets – can be stated:

**Corollary 6.** Let $A$ and $B$ finite sets such that $A \subseteq B$. If $|B| = n$ elements, then $|A| \leq n$.

*Proof.* Let $A \subseteq B$. We have two cases:

(i) $A \neq B$.

(ii) $A = B$.

For (i): If $A \neq B$, then $A \subset B$. So, if $A$ is a proper subset of $B$ and $|B| = n$, then $|A| < n$.
For (ii): If $A = B$ and $|B| = n$, then $|A| = n$.
Hence, from (i) and (ii), $|A| \leq n$. ■

Now, if we consider our definition of "finite" set, it is easy to see that we are assuming the existence of the $\mathbb{N}$ set. But, precisely without an axiomatic setup of natural numbers, the set $\mathbb{N}$ cannot be defined axiomatically and, hence, we have to assume its existence. Indeed, the definition we have given is the one Cantor established, but if we want to get rid of the set $\mathbb{N}$, we can use a Dedekind-style approach to infinite and finite sets. Instead of citing directly from Dedekind's 1888[22] essay we will give his proofs in a more conventional fashion. Let $A$ be a set:

**Definition 14** (Dedekind)**.** A set $A$ is **Dedekind – finite** if there is no injection

$$\pi : A \rightarrowtail B, \text{with B} \subsetneq \text{A}$$

from $A$ into a proper subset $B$ of itself. If $A$ is not Dedekind-finite, then it is **Dedekind – infinite**.[23]

[22] We will return on R. Dedekind's 1888 essay in the section titled "Dedekind on *Systeme* and Logical Abstraction", where the German mathematician's contributions to set theory, logic and philosophy of logic are analysed.

[23] «Definition. A system $S$ is said to be *infinite* when it is similar to a proper part of itself; otherwise $S$ is said to be a *finite* system» (Dedekind 1888b, p. 806). For technical commentary see Moschovakis 2006, pp. 48–49.

With this definition, then, it is possible to prove the a statement similar to Cantor's theorem that, if $A$ is a finite set and $B$ is a subset of $A$, then $B$ is also finite, namely:

**Theorem 2.2.3.** If $A$ is Dedekind-finite, then every subset of $A$ is also Dedekind-finite[24].

Now, the theorem follows:

*Proof.* **Proof by contraposition**
*To prove*: If $B \subsetneq A$ is Dedekind-infinite, then $A$ Dedekind-infinite.

- Let $A$ be a set and let $B \subsetneq A$ be Dedekind-infinite. We show that $A$ is infinite. By Dedekind-infinity there is an injective function (one-to-one) $\pi : B \rightarrowtail B'$, where $B' \subsetneq B$. By extending $\pi$ to a function $\psi : A \rightarrowtail A$, put

$$\psi(a) = \begin{cases} \pi(a), & \text{if } a \in B \\ a, & \text{if } a \notin B. \end{cases}$$

  $\psi$ is injective and, hence, we can consider whether an element is in $B$ or don't. Since $B' \subsetneq B$, there is an element $a \in B$ but not in $B'$. So, $a \in B$ determines that $\pi(a) = \psi(a)$. Otherwise, if $a \notin B$ and $\psi(a) = a$, then $\psi(a) \notin B$. So, finally, if $B \subsetneq A$ and the set of images of $\psi$, formally $\Psi(A)$, is not in $B$, this means that $\Psi(A) \subset A$. This allows us to conclude that $\psi$ in an injection from $A$ to a proper subset of itself, namely:

$$\psi :\ A \rightarrowtail \Psi(A)$$

- Now it is possible to state the contrapositive of the result:

  If $A$ is Dedekind-finite, then every $B \subsetneq A$ is Dedekind-finite.

∎

Within the axiomatic framework and, especially thanks to Axiom of Choice[25], Cantor's and Dedekind's notion of "finiteness" will become equivalent. Additionally, another interesting element of Dedekind-finite sets is the following:

(i) Suppose $A$ contains a proper subset $A'$ with a bijection $\pi : A \to A'$. This, in particular, is an injection.

(ii) Suppose $A$ contains a proper subset $A'$ with an injection $\psi : A \rightarrowtail A'$. Then, there is an injection $\phi : A' \rightarrowtail A$. Thus, applying the Cantor-Shroeder-Bernstein

[24]Dedekind 1888b, p. 807, Theorem 68 and Moschovakis 2006, p. 48. My own proof.

[25]I'll discuss the axioms of set theory in the next chapter. See Chapter 3, section "Zermelian considerations on the axiom", first paragraph "Well-orderings, choices and axiomatic method".

Theorem[26], which does not require the Axiom of Choice, there is a bijective function $\pi : A \to A'$.

This means that Dedekind-finite sets can be formulated either involving an injective function or a bijection from a set to a proper subset of itself. Thus, without any axiom and by showing (i) and (ii), the two conditions are equivalent.

**Where are we now?** In this section, mostly devoted to Cantor's sets, we have seen that:

1. Any property may determine a collection;
2. Equinumerosity is an equivalence relation between sets;
3. A set may be partitioned by considering equinumerosity between its elements.

Finally, we have seen that there are two way to deal with finite or infinite collections, one devoted to Cantor, the other proved by Dedekind. While formulating our argument as based upon this few achievements, we will consider finite sets in more Dedekindian sense. Recall that Benacerraf argued that sets cannot be useful in defining natural numbers: different and contrasting set theoretical representations of a single natural number are available and, hence, every reduction, of the second to the first, should be rejected. Since we are aiming to consider whether some sets can be useful in representing natural numbers, their presence should not be assumed from the beginning. In the next section, indeed, we will:

(i) Construct finite sets by using Dedekind's notion of finiteness;

(ii) Establish between them an equivalence relation, i.e. equipotency;

(iii) Partition sets;

(iv) Revaluate the notion of identity, invoked by Benacerraf, and faithful representations.

### 2.2.3 Reconsidering Benacerraf's Thesis I

Bencacerraf's article "What numbers could not be" (1965) showed us that:

(**B**) : Numbers are not sets and, moreover, numbers are not abstract objects

In this chapter, we've introduced a formal framework connected to the so-called "naïve set theory" and some elements of the algebra of sets. These tools allow us to discuss Benacerraf's problem from a different perspective than in the foregoing

[26]The theorem states that $\leq$ is a *partial order* relation. As we will see, this result is a fundamental logical tool since it will allow us to order the cardinals. Generally, it shows that, given two arbitrary injective functions $g : A \to B$ and $f : B \to A$, there exists a bijection $h : A \to B$. The theorem is fundamental from a set-theoretical point of view since if we have that $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$, that is $A$ and $B$ are equipotent. It is interesting to notice that, although Dedekind proved this theorem twice during his life (1887 and 1892) without the axiom of choice, his name is not mentioned and explicitly connected to this result.

chapter. Recall that we considered that «numbers are not objects at all, because in giving the properties (that is, necessary and sufficient) you merely characterize an *abstract structure*»[27], so that to be the number "2" is no more and no less than *any object* preceded by 1, 0, and followed by $3, 4, 5, \ldots$. We called these suggestion as Benacerraf's ontological reduction: numbers are just "places" of *any* progression, standing in well-defined relations. These "ontological" suggestions follow, according to Benacerraf, why many philosophers have identified mistakenly numbers with sets. Benacerraf's thesis is against the identification of number with sets. Since we have two good, but different, set theoretical representations, such as

(i) $\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \ldots$ (Zermelo)

(ii) $\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \ldots$ (von Neumann),

it is not possible to identify any number-sign with them. Our first considerations are settled within naïve se theory, we do not have axioms that allow us to form new sets starting from "empty set", $\varnothing$, and, hence, thanks to the General Comprehension Principle, every property determines a starting point to construct further sets.
We begin our (re-)consideration by giving some definitions:

**Definition 15.** Let $\mathcal{F}$ be the class containing all finite sets and let $\mathcal{F}/\sim$ the quotient set of $\mathcal{F}$, that is the set partitioned by the equivalence relation $\sim$ (equipotency).

Firstly, consider that our $\mathcal{F}$ contains all the sets that cannot have an injection within a proper subset of themselves, i.e. they're finite.
For the sake of the argument, consider the $\mathcal{F}/\sim$, that is the set partitioned by the equivalence relation of "equipotency". We have proved that two sets $A$ and $B$ are equipotent, written $A \sim B$ iff there is a bijection $f : A \to B$. In addition, we've proved that the equipotency relation is an equivalence relation, that is reflexive, $A \sim A$, symmetric, $A \sim B \implies B \sim A$, and transitive, $A \sim B$ and $B \sim C \implies B \sim C$. In order to draw our conclusions we need to define the equivalence classes generated out by the equipotency relation on $\mathcal{F}$:

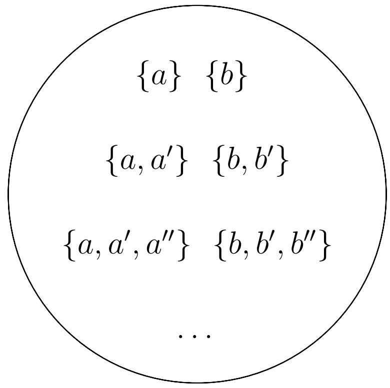
**Definition 16.** Given the equipotency $\sim$ relation on the set $\mathcal{F}$, the equivalence class of an element $x \in \mathcal{F}$ is defined as follows:

$$[x] = \{y \mid y \sim x\}$$

Starting from these achievement we might be apply Theorem 2.1 and say that the equivalence relation, $\sim$, partitions our set $\mathcal{F}$. Suppose to have the following situation:

[27]Benacerraf 1965, p. 291.

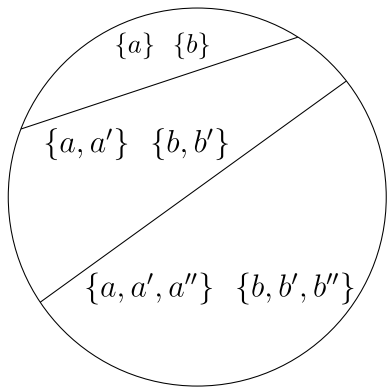
$\mathcal{F}$




As it results from the Venn diagram representing $\mathcal{F}$, we have a set that contains all the finite sets. Recall that, by our definition, a set is finite iff it has no injection into a proper subset of itself. In other terms, for any set $B \subsetneq A$, there is, at least, an element in $A$ to which no element of $B$ is associated.
In Cantor-style, – but, by assuming the existence of $\mathbb{N}$ – a set is finite iff there is a bijective function $f : A \to \mathbb{N}_n$. In this way, the sets contained in $\mathcal{F}$ are all in a biunivocal correspondence with some subset of the natural numbers, that is $\mathbb{N}_n \subseteq \mathbb{N}$. In addition, recall that $\mathbb{N}_n$ is a notation to indicate the "set of numbers in $\mathbb{N}$ less than $n$", that is $\mathbb{N}_n = \{0, 1, 2, 3, \ldots, n-1\}$, for every $n \in \mathbb{N}$.

However, consider now $\mathcal{F}/\!\sim$, that is the set $\mathcal{F}$ partitioned by the equivalence relation of equinumerosity between its elements. In this case, the equivalence class will be determined by the number of elements that sets share, namely all the sets containing $n$ members will belong to the same equivalence class, $[n]$. The previous image can be transformed into the following:

$\mathcal{F}/\sim$




The equivalence relation of equinumerosity between sets partitions our $\mathcal{F}$ into equivalence classes determined by the number of elements that sets share. For example, consider the sets determined by their having just two elements, namely:

$$\{a, a'\}$$

$$\{b, b'\}$$

$$\{c, c'\}$$

$$\{d, d'\}$$

$$\dots$$

Suppose that two of them represent respectively a collection of 2 apples and a set composed by 2 oranges. Since our class $\mathcal{F}$ includes all the finite sets, then the two collections of apples and oranges belong to it. Consider the figure of $\mathcal{F}/\sim$ and the partition $\mathcal{F}_2$ containing $\{a, a'\}$, $\{b, b'\}$, $\{c, c'\}$, $\{d, d'\}$, and so on. This partition is determined by the equivalence class based on the equipotency relation between the two sets involved. Roughly, the set containing 2 apples, $\{a, a'\}$, and the set containing 2 oranges, $\{b, b'\}$ (and all the sets containing just 2 elements), belong to the same partition of $\mathcal{F}$ and this means that:

(i) $|\{a, a'\}| = |\{b, b'\}| = |\{c, c'\}| = \cdots \iff \{a, a'\} \sim \{b, b'\} \sim \{c, c'\} \sim \dots$

Since the sets are equinumerous it is possible to conclude:

(ii) $[\{a, a'\}] = [\{b, b'\}] = [\{c, c'\}] = [\{d, d'\}] = \dots$

In other terms, the equivalence class to which the sets in $\mathcal{F}_2$ belong is the following:

(iii) $[2] = \{\{a, a'\},\ \{b, b'\},\ \{c, c'\}, \{d, d'\}, \dots\}$

Similarly, the equivalence class, to which the sets of any partition $\mathcal{F}_n$ belong, is so determined:

$$[n] = \{\{a, a', \dots, a^n\},\ \{b, b', \dots, b^n\},\ \{c, c', \dots, c^n\}, \{d, d', \dots, d^n\}, \dots\}$$

This means that, in our example, "2 is the class containing all the sets in $\mathcal{F}$ that contain just two elements". More generally, $[n]$ is the class containing all the finite sets in $\mathcal{F}$ that are equinumerous each with respect to the others.

### 2.2.4 Ontological Remarks I

As should already be clear this argumentation does not aim to refute Benacerraf's ontological reductive argument. This, probably, won't even happen while having the axiomatic version of set theory and the other techniques to discuss ordinal and cardinal numbers "as sets". Indeed, what we count to do is just to sketch different perspectives from which Benacerraf's argumentation can be judged and evaluated. What we've achieved in this chapter is, indeed, a mere logical consideration concerning numbers and equinumerosity: since a natural number indicates the cardinality of a determined set, and since a set can be partitioned by the equinumerosity relation, it is possible to affirm that numbers are equivalence classes of equipotent sets. Our argument uses naïve set theory and, indeed, any collection of things (such as apples, oranges, . . . ) is assumed to form a set. A version of our argument can be extended to the axiomatic version of set theory to consider exactly the two set-theoretical

reductions that Benacerraf's article treated. In any case it is important to recognize that doubting of Benacerraf's conclusions, for which sets are not logically reducible to sets, allows us – I'll maintain – to investigate again, more deeply and with a different care, some ontological features of mathematical objects.
Anyway, up to now, we have seen that numbers can be considered as equivalence classes of sets. What does this exactly mean? What are we operating here? We see that, by saying that natural numbers are equivalence classes of sets we are allowing someone to affirm that "two is the class which includes all the sets containing two elements", whatsoever elements they are. In this case, even if the set of two oranges apparently does not share anything with the set of two apples, they actually share the fact of having just two elements. In other words, what the two collections share is the equivalence class, namely the partition of $\mathcal{F}$ obtained by the equipotency relation. In this sense, hence, a number $n$ could be considered as the class consisting exactly in those sets that have $n$ elements as members. Thus, $n$ is not "one" particular set, but corresponds to the partition – belonging to class of all finite sets, $\mathcal{F}$ – composed of equipotent sets[28]. As we have immediately noticed, the argument for which numbers *are* sets, is based – according to Benacerraf – upon the identification of numbers and sets. Are we sure that our set theoretical reductions imply the "total" identification of what we represent, starting from the empty set, and the natural numbers? In this context, for instance, Zermelo's or von Neumann's ordinals should be thought *as the* natural numbers themselves, that is: are there no salient differences between the two mathematical entities considered? While trying to answer this question we will introduce exactly the sets Benacerraf's article treats (Zermelo's and von Neumman's) and we will see that – in order to develop whatsoever philosophical claims starting from mathematics itself – we should carefully distinguish between "identity statements" and "faithful representations".

## 2.3 Little interlude. Back to Philosophy

The previous part of this second chapter has been engaged with the more logical and mathematical arguments of Benacerraf's paper and we have briefly argued how we think our reflections might undercut Benacerraf's starting point. Recall that the French philosopher derived – from his analysis of sets and numbers – his desired philosophical conclusion, i.e., the impossibility for an ontology of mathematics to contain abstract objects. In contrast, here, we begin developing some rough and introductory considerations on the path that, historically, applied to, and brought into, the philosophy of mathematics, the discussion concerning abstract objects[29].

[28]In the succeeding parts of the present work we will reconsider Benacerraf's argument (**B**) also in the axiomatic versions of von Neumann and Zermelo. This latter developments will, in some sense, strengthen these preliminary "ontological remarks".

[29]Since our aims are not strictly exegetical, we will briefly introduce the debate and return to some of its contemporary developments within Chapters 4-5 of the present work.

### 2.3.1 Frege's Notion of Abstraction

Again, in this little interlude, the main source of inspiration has to be traced back to Frege's work. To see, how initially Frege thought that logical objects might be introduced and characterized, let's consider that in contemporary literature laws, such as Hume's Principle, Basic Law V or the General Comprehension Principle for extensions, are called *abstraction principles.* The word "abstraction" is philosophically and logically meaningful, therefore we restrict our usage to the following:

> According to the philosophical tradition, to abstract is to "extract" from a class of things a feature that these things have in common when they are equivalent in some respect. For instance, we abstract the colour *red* from a collection of things that are chromatically equivalent.[30]

Let's take, for instance, Hume's Principle:

$$\#F = \#G \longleftrightarrow F \sim G$$

If we agree that one way in which concepts may be equivalent is expressed by the equinumerosity relation, then, if we abstract on equivalent equinumerous concepts we obtain the identity between their numbers (or cardinalities).
Another fregean examples concerns the "abstraction" of the concept of "direction" from the sole notion of "line":

**Example** (Directions)**.**

$$d(\ell_1) = d(\ell_2) \longleftrightarrow \ell_1 \,||\, \ \ell_2$$

In other words, the directions of two lines ($d(\ell)$) are identical if and only if the two lines involved are parallel ($||$). The example concerning the directions of lines, can be generalized. Let $R$ be a generic equivalence relation between the lines considered and $=$ be the associate identity predicate on the abstracted objects (in our case, the directions), which holds of them, just in case $R$ holds of the lines from which the directions are abstracted. Formally, the previous abstraction principle can be rewritten as follows:

$$\S(\ell_1) \ = \ \S(\ell_2) \longleftrightarrow R\,(\ell_1,\, \ell_2).$$

Historically, hence, in Frege's ontology, where there's place for abstract individuals, logical objects are introduced thanks to rules that

> [...] allow us to talk about directions – just as more familiar objects – as presented in different ways, as identified and distinguished, and as objects of various predications. At the very least, this shows how it can be permissible to talk *as if* there are mathematical objects – in this case, directions.[31]

[30]Linnebo 2011, p. 30. For a clear treatment of the origins and of some contemporary discussions of abstraction principles, see Ebert and Rossberg 2017, pp. 3–33. See also our Chapters 4-5.
[31]Linnebo 2011, p. 127.

Notice that the object represented by $d(\ell)$, the direction of $\ell$, inherits all the properties and relations that characterize the line $\ell$ itself. In this sense, the identity of the two directions is nothing more and above the parallelism of the two lines in terms of which the directions are specified. Problems explicitly arise while considering Basic Law V:

$$\epsilon' f(\epsilon) = \alpha' g(\alpha) \longleftrightarrow \forall x(f(x) = g(x)),$$

or, with respect to concepts,

$$\epsilon F = \epsilon G \longleftrightarrow \forall x(F(x) = G(x)).$$

That is, when two concepts, such as $F$ and $G$, are coextensive, then they have the same extension (i.e., the same objects falling under them). By considering the class whose members are all and only those objects that are not members of themselves we get a contradiction. Indeed, what post-fregean set theorists learned from Frege's failure is, that it should be considered «dangerous to abstract on concepts in a way that yields objects»[32] and developed several methodologies to avoid paradoxes. Anyway, not only philosophers applied to abstraction techniques and, in the same vein, Cantor's implicit use of the General Comprehension Principle for sets bares the same paradoxical conclusions as Frege's Basic Law V:

**Theorem 2.3.1** (Russell′s Paradox 3)**.** General Comprehension Principle entails a contradiction.

*Proof.* If the General Comprehension Principle for sets holds then the "set of all sets" exists:

$$U = \{x \mid x \text{ is a set}\},$$

and consider its peculiar property of "being a member of itself", $U \in U$.
Applying another time the General Comprehension Principle for sets let $R$ be:

$$R = \{x \mid x \text{ is a set} \longleftrightarrow x \notin x\}.$$

It follows, from the definition of $R$,

$$R \in R \longleftrightarrow R \notin R$$

Contradiction. ■

## 2.4 Short introduction to Dedekind

### 2.4.1 *Systems* and Logical Abstraction

Historically and theoretically, both, set theory and abstraction techniques, deserve to be analysed also from another perspective, i.e., R. Dedekind's. In this context, the

[32] Linnebo 2011, p. 129.

following explanations will follow two parallel, even if distinguished, lines. From one side, we will clarify some logical and mathematical aspects concerning Dedekind's profound intuitions regarding "systems" (his term for sets)[33] and how they can be helpful in the discussion of the conception that views "numbers as sets". Secondly, we will consider his *logicist* attitude and his "innovative" notion of *logical abstraction*[34]. Now, it's useful to point out that Dedekind, unlike Frege, has not always been considered as an official "philosopher", but recent works have tried to revaluate his philosophical contributions to the foundations of mathematics[35]. Our investigations on Dedekind "as philosopher" will, indeed, be conducted in comparison with Frege's conceptions[36] and aims to rehabilitate Dedekind's figure into the philosophical debate around mathematics[37].

#### 2.4.1.1 Dedekind's Contributions to the Early Development of Set Theory

Like Cantor, also Dedekind contributed much to the development of naïve set theory. Like Frege, Dedekind thought that the basic truths concerning natural numbers could be derived from irreducible logical notions. Thus, in his famous essay of 1888, the German mathematician, starting from the notions of *System*[38] (set), *Ding* (object) and *Abbildung* (mapping, function), elaborated a theory of "systems" and, consequently, showed how to characterize the set of the natural numbers:

> If we scrutinize closely what is done in counting a set or number of things, we are led to consider the ability of the mind to relate things to things, to let a thing correspond to a thing, or to present a thing by a thing, an ability without which no thinking is possible. Upon this unique and therefore indispensable foundation [...] the whole science of numbers must, in my opinion, be established. [...]
> In accordance with the purpose of this memoir I restrict myself to the consideration of the series of so-called natural numbers[39].

Indeed, Dedekind – in explicit accordance with the authors we've discussed at the beginning of the section devoted to the foundations of naïve set theory – defined a *System*, or set, by saying that:

> It very frequently happen that different things $a, b, c, \dots$ for some reason can be considered from a common point of view, can be associated in the mind, and we say that they form a *System* $S$; we call the things $a, b, c, \dots$ *elements* of the system $S$, they are *contained* in $S$; conversely, $S$ *consists* of these elements. Such a system $S$

[33] Dedekind 1890a, 1888b.

[34] Much of the considerations we will express on this point are due to Tait 1996, Linnebo and Pettigrew 2014 and Reck 2018.

[35] For an introductive text consider, for instance, Potter 2000, pp. 82–104. For good commentary see Reck 2003, Reck 2009, Reck 2016, Reck 2017.

[36] See, in particular, Reck 2013a, Reck 2013b.

[37] In particular, we will follow Dedekind 1888b, Dedekind 1890a.

[38] Dedekind's original essay was written in German.

[39] Dedekind 1888b, pp. 791–792.

> (an aggregate, a manifold, a totality) as an object of our thought is likewise a thing; $S$ is completely determined when, for every thing, it is determined whether it is an element of $S$ or not[40].

So, as before, a system is a collection of elements and the members that it collects are characterized by the membership-relation. Likewise, in Dedekind's treatment of systems we find a statement corresponding to the fundamental Extensionality property of sets:

> The system $S$ is hence the same as the system $T$, in symbols $S = T$, when every element of $S$ is also an element of $T$, and every element of $T$ is also an element of $S$.[41].

The text continues by explaining all the fundamental notions, such as inclusion, intersection, union, etc[42]. In order to discuss his treatment of arithmetic it is useful to focus for a moment upon the concept of *ähnliche Abbildung* (similar map), that is Dedekind's German characterization of an injective function (one-to-one correspondence). In the essay we are first introduced to the general notion of map[43] and, just after a few pages, Dedekind gives us a definition of "similar mapping":

> A mapping $\phi$ of a system $S$ is said to be *similar* [*ähnlich*] or *distinct*, when two different elements $a, b$ of the system $S$ there always correspond different images $a' = \phi(a),\ \ b' = \phi(b)$[44].

This is the dedekindian condition for $\phi$ to be a one-to-one correspondence. He continues explaining that his statement is correspondent to a "negative" claim, namely:

$$a \neq b \rightarrow [a' = \phi(a)] \neq [b' = \phi(b)].$$

In any case, he noticed that the "contrapositive" direction of his definition holds:

$$[a' = \phi(a)] = [b' = \phi(b)] \rightarrow a = b.$$

Now, put the set of images $\phi(S) = S'$. The *inverse* function is defined as the rule that $\overleftarrow{\phi}(a') = a$. In other words, $\overleftarrow{\phi} : S' \rightarrow S$ is the map that applied to $a'$, equivalent to $\phi(a)$, gives as result $a$. Now, $\overleftarrow{\phi}$ is clearly one-to-one or *similar* and $\overleftarrow{\phi}(S') = S$. Finally, $\phi \circ \overleftarrow{\phi}$ is the identity function on $S$.

As we have seen earlier, Dedekind's definition of "finiteness" and "infiniteness" is very useful since allows us to have a "good" notion of finite sets without employing

[40]Dedekind 1888b, p. 797.

[41]Dedekind 1888b, p. 797.

[42]Dedekind employs a different notation for sets and their relations, but we will follow the contemporary usual mathematical practice.

[43]See Dedekind 1888b, p. 799. The essence of his definition is the following: a function or map $\phi$ with *domain* $S$ is a rule that assigns to any element $s \in S$ a value $\phi(s)$, called the *image* of $s$. We are allowed to say that $\phi$ maps $s$ to $\phi(s)$.

[44]Dedekind 1888b, p. 801.

and assuming the existence of a particular set, such as $\mathbb{N}$. Nonetheless, what is really interesting concerning the infinite sets is the fact that their "existence" has been "proved" by Dedekind. As we will turn to axiomatic set theory, the existence of such a set will be guaranteed by an axiom. In any case, the German mathematician stated that «there exist infinite systems» and tried to prove it in a very particular and non-strictly mathematical way:

> Proof. My own realm of thoughts, i.e. the totality $S$ of all things, which can be objects of my thought, is infinite. For if $s$ signifies an element of $S$, the thought $s'$, that $s$ can be object of my thought, is itself an element of $S$. If we regard this as an image $\phi(s)$ of the element $s$, then the mapping $\phi$ of $S$, thus determined, has the property that the image $S'$ is part of $S$; and $S'$ is certainly a proper part of $S$, because there are elements in $S$ (e.g., my own ego) which are different from such a thought $s'$ and therefore are not contained in $S'$. Finally, it is clear that if $a$, $b$ are different elements of $S$, their images $a'$, $b'$ are also different, and that therefore the mapping $\phi$ is a distinct (similar) mapping. Hence $S$ is infinite, which was to be proved[45].

Except the fact that Dedekind's proof is in some sense "mystical", he is introducing *the totality $S$ of all things which can be objects of my thought* – his "universal" system or set – with arbitrary subset of that totality, namely each $S'$. As it is clear, however, here we are not in front of numerical sets but we are dealing with any collection of objects that will consequently count as a set. So, Dedekind's contributions fall under the so-called naïve or early development of set theory. Indeed, consider that, like Cantor's sets or Frege's extensions, also Dedekind's systems fall under the so-called Russell's Paradox[46]. In any case, this antinomy does «not invalidate his other contributions to set theory», such as «the definition of being Dedekind-infinite, the formulation of the Dedekind-Peano axioms, the proof of their categoricity, the analysis of the natural numbers as finite ordinal numbers»[47], and so on.
Similarly to Frege, Dedekind pursued a sort of *logicist* project but, differently from Frege's syntactic treatment of natural numbers, he tended to focus on model-theoretical aspects. Also, unlike Frege's treatment of cardinals, Dedekind's analysis is about "ordinals":

> Thus, nothing like Frege's analysis of deductive inference, by means of his "Begriffsschrift", can be found in Dedekind's work. Dedekind, in turn, is much more explicit and clear than Frege about issues such as categoricity, completeness, independence, etc. This allows him to be seen as a precursor of the "formal axiomatic" approach championed later by Hilbert and Bernays[48].

Indeed, in Dedekind's 71. Definition, defines what is for a set to be *simply infinite* and

[45]Dedekind 1888b, pp. 806–807, Theorem 66. This notion will lead to the definition of an "inductive set" or, in dedekindian terms, of a *simply infinite* system.
[46]See paragraphs before on Frege's extensions and Cantor's aggregates.
[47]Reck 2016, p. 17.
[48]Reck 2016, p. 20.

how it can be helpful in constructing and characterizing arithmetic. Dedekind stated some laws concerning the natural numbers – as Peano was doing – as definitions within a larger theory, namely set theory. We'll state Dedekind original claim by allowing us to put it in contemporary notation:

**Definition 17** (Simply − Infinite)**.** A set $N$ is said to be **simply − infinite** when there exist a one-to-one map of $N$ into itself such that $N$ appears as the chain[49] of an element not contained in $\phi(N)$. We call this element **base − element** and denote it with 1.
Additionally, we say that $N$ is **ordered** by $\phi$. The notation $\phi(n)$ can be abbreviated with $n'$.
Finally, if $N$ is a simply-infinite set, with a mapping $\phi(N)$ and an element 1, then it satisfies the 4 following conditions:

(i) $N' \subseteq N$

(ii) The chain of 1 is in $N$. Formally, $N = 1_0$

(iii) The element 1 is not contained in $N'$. In symbols, $1 \notin N'$

(iv) $\phi$ is one-to-one.

By (i), (iii), and (iv) it follows that $N$ is an infinite set, in the sense of Dedekind-infinite.

The characterization of $N$ – (i)-(iv) – given here by Dedekind has become to be known as the Dedekind/Peano axioms for the natural numbers. In other words, Dedekind's reasoning can be put as follows. Consider a set $S$ and a subset $N \subseteq S$. $N$ is called *simply-infinite* if there is a function $\phi : S \to S$ and the *base-element* $1 \in N$, such that:

1. $\phi$ maps $N$ into itself, $\phi : N \to N$.
2. $N$ is the chain of $\{1\} \subseteq S$, closed under $\phi$.[50]
3. $1 \notin \phi(N)$.
4. $\phi$ is one-to-one.

So, a simply-infinite system, such as $N$, will consist of a first element 1, a second element $\phi(1)$, a third element $\phi(\phi(1))$, and so on. Intuitively, any simply-infinite system $N$, with a base-element 1, and a "similar" map $\phi$, play the role of the arithmetical notions of $\mathbb{N}$, 1 and $S$, that is the "successor function". This means that Dedekind's construction is a "model" for the natural number sequence. (Consider the graphical representation of Dedekind's $N \subseteq S$ sequence, Figure 2.2.) Indeed, Dedekind's work is related to that of Peano for the following reasons:

[49]Dedekind defines a *chain* or *Kette* as follows: Given a map $\phi : S \to S$, we say that $K \subseteq S$ is a **chain** if $\phi$ maps $K$ to $K' \subseteq K$. See Dedekind 1888b, p. 803, Definition 37.

[50]This is the dedekindian version of the mathematical induction principle.

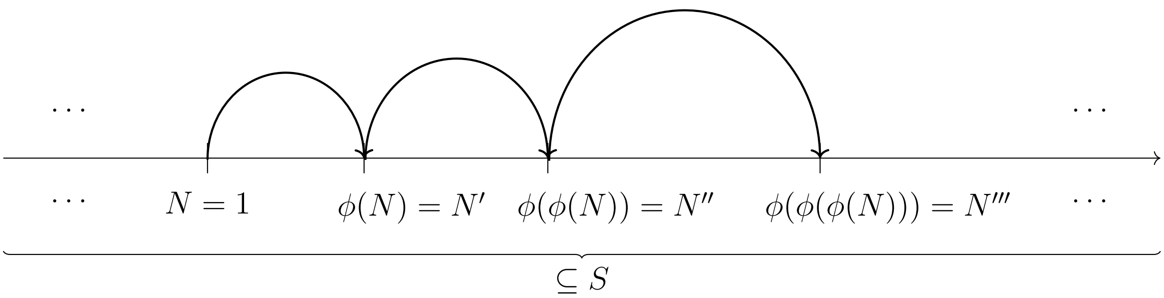



Figure 2.2: $N \subseteq S$ closed under $\phi$

- $N = 1$ equals exactly with Peano's axiom $\mathbb{N}(0)$
- $1 \notin N'$ is equivalent to $S(x) \neq 0$
- $\phi$ is one-to-one means what Peano states as follows: $S(x) = S(y) \to x = y$
- $N' \subseteq N$ corresponds to:

$$\forall X\Big(X(0) \wedge \forall y[X(y) \to X(S(x))] \to \forall y[\mathbb{N}(y) \to X(y)]\Big).$$

Equipped with these definitions – some paragraphs later – Dedekind proved his famous "categoricity" theorem, which states that any two simply-infinite systems are "isomorphic"[51]:

> Theorem. Every system which is similar to a simply infinite system and therefore to the number sequence $N$ is simply infinite[52].

In current language we're allowed to say that "every set which is in one-to-one correspondence to a simply infinite system is simply infinite". The proof relies on a simply idea. Let $[A, a, \phi]$ and $[B, b, \psi]$ denote two simply infinite systems, where $a$ and $b$ indicate their base-elements and $\phi$, $\psi$ two "similar" mappings. Define, then, an isomorphism $\pi : A \to B$, by mapping one intial object to the other, $\pi(a) = b$. Extend $\pi$, by putting $\pi(\phi(x))$ be $\psi(\pi(x))$, for any $x \in A$. Using induction it is possible to show that $\pi$ is defined on the totality of $A$ and is an isomorphism.[53]
As we have seen, Dedekind's results in set theory are important and elegant as those of Cantor and Frege. His way of constructing the natural numbers by introducing a simply-infinite set as their model, is very innovative and *abstract* – indeed, with his categoricity theorem he showed that *any* simply-infinite system is isomorphic to the simple-infinite system representing the natural numbers. Despite the importance of Dedekind's strategies considered from a technical and mathematical point of view, his work contains also some deep metaphysical and epistemological issues concerning mathematics itself. In particular, our next inquiries will be devoted to the notion of

[51]That is: it exists a bijective function (onto and one-to-one) from the domain of one set to the codomain of the other. Recall that in naïve set theory, the condition of being either an injection or a bijection are equivalent.

[52]See Dedekind 1888b, p. 822, Theorem 133.

[53]For commentary see Linnebo 2011, pp. 157–159 and Button and Walsh 2011, pp. 154–155.

"Dedekindian abstraction" or "creation" – concept that we have already implicitly encountered in the proof regarding the existence of an infinite system.

#### 2.4.1.2 Philosophical Revaluation of Dedekind's Notion of Logical Abstraction

In his essay Dedekind tried to explain how his foundations for arithmetic should be thought of and the passage we will quote contains one of the clearest declaration of Dedekind's conception of "abstraction"

> Definition. If in the consideration of a simply infinite system $N$ ordered by a mapping $\phi$ we entirely neglect the special character of the elements, simply retaining their distinguishability and taking into account only the relations to one another in which they are placed by the ordering mapping $\phi$, then these elements are called *natural numbers* or *ordinal numbers* or simply *numbers*, and the base-element 1 is called the *base-number* of the *number-series N*. With reference to this liberation of the elements from every content (abstraction) we are justified in calling the numbers a free creation of the human mind[54].

This passage has often been quoted by philosophers and logicians, since it is not clear whether Dedekind is employing a sort of "psychologistic" approach towards the philosophy of mathematics. According to M. Dummett[55], indeed,

> One of the operations most frequently credited with creative powers was that of abstracting from particular features of some object or system of objects, that is, ceasing to take any account of them.[...]
> It was to this operation that Dedekind appealed in order to explain what the natural numbers are.[56]

For Dummett, therefore, Dedekind's philosophy of mathematics and, especially his operation of abstraction or creation, is to be understood as "psychologistic". Recall that, in Dedekind's perspective, by starting from a triple $\mathbf{X} = [X, a, \phi]$, where $a$ is the *base-element* (which is not in the range of $\phi$ and $X$ is the least system containing $a$ closed under $\phi$) and $\phi$ is a one-to-one function $\phi : X \to X$. $\phi$ is said to order $X$ and $\mathbf{X} = [X, a, \phi]$ denotes any simply infinite system. By "abstracting" from the elements of $\mathbf{X}$ we may obtain a simply-infinite system $\mathbf{N} = [N, 1, S]$, representing the numbers. Thus – according to Dedekind's result – arithmetic depends only on the axioms of the second-order theory of simply infinite sets, and not on the choice of any particular theory. When we consider just a simple collection of elements, equipped with a "base" member, and a "similar" mapping, that is $\mathbf{X} = [X, a, \phi]$, we

[54]Dedekind 1888b, p. 809.

[55]Michael Dummett (1925 – 2011) has been a philosopher and logician, widely known for his contributions to the philosophy of language, of mathematics and to logic. His work concerned in particular intuitionistic logic and Frege's logicism. He held a view for which the Frege-Russell logicist project could have had success if it has had an underlying intuitionistic, and not classical, logic.

[56]Dummett 1995, p. 50.

consider it as a good explication and reduction of $\mathbf{N} = [N, 1, S]$. Indeed, $[N, 1, S]$ play the same roles of $[X, a, \phi]$ – the latter is technically called a "model" for the first one by representing it in a more *abstract* and *general* way. Dummett holds that Dedekind's account of abstraction in not tenable since it would lead to a *solipsistic* and psychologistic conception of arithmetic. If the way in which Dedekind conceives abstraction leads us to the sole consideration «that mathematical objects are "free creations of human mind"» then,

> [...] it is implicit [...] that each subject is entitled to feel assured that what he creates by means of his own mental operations will coincide, at least in its properties, with what other have created by means of analogous operations. [...] Such an assurance would be without foundation[57].

But, is it Dummett's argument that precise? Dedekind – in the preface – doesn't say that "mathematical objects" are free creations of human mind, but that this status concerns, in the context of his monograph, natural numbers. W. Tait correctly points out that:

> [...] it is too hasty to reduce [Dedekind's] "philosophy of mathematics" to a psychologistic reading of this metaphor. [...] It is reasonable to conclude that Dedekind's conception is psychologistic only if that is the only way to understand the abstraction that is involved. And we shall see that it is not. The difficulty with abstraction as a psychological operation would that what is abstracted is mental, that what I abstract is mine and what you abstract is yours.[58].

As it can be seen from our previous example concerning $\mathbf{N} = [N, 1, S]$ and $\mathbf{X} = [X, a, \phi]$, the notion of abstraction that Dedekind involves is not a psychologistic operation in the sense that, the system I obtain by abstracting from $\mathbf{X} = [X, a, \phi]$ is not a system different from the one another person obtains. The notion of abstraction can, indeed, be regarded as "logical" and, according to Tait and Reck, this latter consideration can be useful in revaluating Dedekindian abstraction.
First of all, Dedekind's work aims to treat «arithmetic (algebra, analysis) as merely a part of logic»[59] and, moreover, as «an immediate product of the pure laws of thought»[60]. His theory of numbers, indeed, has been introduced after having developed a theory of systems composed by objects that are related by mappings. These three elements – *System* (set), *Ding* (object) and *Abbildung* (map) – are primitive and irreducible, that is they cannot be further clarified. At most, according to Dedekind, we can have a better understanding of them in their usage and application. By taking the simply infinite system $\mathbf{X} = [X, a, \phi]$ and by abstracting from it $\mathbf{N} = [N, 1, S]$, we are able to understand the truth of the propositions of $\mathbf{N}$ just in terms of the

[57] Dummett 1995, p. 49.
[58] Tait 1996, p. 11.
[59] Dedekind 1888b, p. 790.
[60] Dedekind 1888b, p. 791.

truths of **X**. Hence, the truth of **N** is founded upon the truth of **X**[61].So, having understood how $[X, a, \phi]$ (system, base-element, and similar mapping) behave, then it is possible to understand how $[N, 1, S]$ (natural numbers, base-element 1, and successor function) behave.
This treatment clearly depends much on infinite systems and, maybe, this is the reason for which Dedekind's tried to prove its existence and did not assume it, as it is done in standard contemporary set theory:

> [...] does such a system *exist* at all in the realm of our ideas? Without a logical proof of existence it would always remain doubtful whether the notion of such a system might not perhaps contain internal contradictions[62].

This last consideration, contained in Dedekind's letter to Keferstein (1890), is in precise connection with the methodology he develops and briefly explains in the preface of his essay of 1888:

> In science nothing is capable of proof ought to be believed without proof. Though this demand seems reasonable, I cannot regard it as having been met even in the most recent methods of laying the foundations of the simplest science; viz., the part of logic which deals with the theory of numbers[63].

So, once we've understood some elementary logical notions such a set, object and mapping, we are able to found in a precise way the "simplest science", namely, arithmetic. By abstraction, then, it is possible to "extract", from the general theory of systems, a "model" for the natural number sequence. Hence, logical notions are needed in order to explain number theoretical concepts:

> [...] it would seem that logical abstraction, as it is described here, does play a role, not in proofs, but in that it fixes grammar, the domain of meaningful propositions, concerning the objects in question, and so determines the appropriate subject matter of proof[64].

Let's analyse for a moment how Dedekind's abstraction may be conceived of[65]:

> Dedekind simply posits, for each system $S$, a pure structure $[S]$, such that $S$ satisfies *Instantiation*, *Purity* and *Uniqueness*[66].

Let now our **N** be the system $S$ and **X** be the pure structure $[S]$.

1. *Instantiation*: **N** is isomorphic to **X**, which we abbreviate with $\mathbf{N} \simeq \mathbf{X}$.
2. *Purity*: $a$ is an element[67] in **N** and $P$ is a property. If $P$ is a property of $a$,

[61] See Tait 1996, p. 15.
[62] Dedekind 1890a, p. 101.
[63] Dedekind 1888b, p. 790.
[64] Tait 1996, p. 16.
[65] We will follow Linnebo and Pettigrew 2014.
[66] Linnebo and Pettigrew 2014, p. 278.
[67] Linnebo and Pettigrew were discussing mathematical-eliminative structuralism, that is the view for which mathematics studies formal "structures" containing just "positions" and no abstract

then for each system **X**, such that $\phi : \mathbf{X} \simeq \mathbf{N}$, $P$ is a property of $\phi(a)$.

3. *Uniqueness*: **X** is categorical.

Hence, by considering the number sequence **N** we can abstract its fundamental structure, namely **X**, show that $\mathbf{N} \simeq \mathbf{X}$ and show that any $\mathbf{N} \simeq \mathbf{X}$ is categorical. Therefore, despite the "psychologistic" metaphor of Dedekind's preface, there is a logical understanding of his notion of abstraction. Indeed,

> [...] our applications of abstraction do not *create* the abstract objects; rather, they pick out a certain realm of objects that already exist (and has always existed) and give us semantic and epistemic access to that realm[68].

This understanding of logical abstraction showed us how that Dedekind – in similar fashion to Frege – considered logic as, in some sense, preliminary and much useful in order to gain a better comprehension of arithmetic and, indeed, also his contributions to the foundations of mathematics can be thought of as a sort of *logicism* (based upon precise *abstraction principles*).

#### 2.4.1.3 Logicist Attitudes: Frege *vs* Dedekind

Logicism is the thesis that arithmetic is reducible to logic alone and, as we have seen, even if differently, both, Dedekind and Frege, advocated this view. If we pay attention to our characterization of logicism, we can see that it is composed by two sub-claims:

($L_1$) All arithmetical concepts are definable in terms of and thanks to the presence of logical concepts.

($L_2$) All arithmetical truths are derivable from logical truths.

Frege's project culminated in the logical system he established in the *Grundgesetze* (1884), while Dedekind pursued his logicist attitude towards his whole mathematical work. We focused our analysis on *Was sind die Zahlen und was sollen?* (1888) where Dedekind developed a natural numbers theory within a more larger set theory. In any case, both, Frege and Dedekind – we argued – based their considerations on particular operations of "abstraction" and

> This brings us back to Frege's main criticism of Dedekind: that he does not spell out the fundamental laws for his logical system explicitly, much less investigate their epistemological source further. Dedekind's approach would have to include laws governing all the existence assumptions in his constructions, concerning both classes and functions. (As he does not reduce functions to classes, separate laws will be needed

entities. What we are calling "element" has renamed the original word invoked by the authors: "position". Linnebo and Pettigrew aimed to show how Dedekind's and Frege's abstractionist accounts could not be useful in defence of an eliminative (structuralist) position. For clear introductions to "mathematical structuralism", see, among others, Linnebo 2011, pp. 154–169, and Button and Walsh 2011, §2.4, §5.2.

[68] Linnebo and Pettigrew 2014, p. 279.

> for them) [...] In Dedekind's case, all corresponding laws remain implicit, although they can be partly gleaned from his constructions[69].

Dedekind's work provides the elements in order to construct a theory of abstraction, which will correspond to a form of neo-logicism and will result different from the fregean and neo-fregean proposals. Consider, indeed, that even if there are some similarities, Frege and Dedekind disagreed upon some fundamental insights of their respective theories. In Frege's case we are presented with a "theory" (rules, symbols, deductive system . . . ), while Dedekind left the underlying theory implicit[70]. Hence, Frege provided a proof theoretical system based upon a determinate and precise formalism, while

> [...] Dedekind's logicism consists precisely in the attempt to derive arithmetic and analysis from core concepts, as opposed to relying either on geometric evidences or on empty formalisms. In the case of analysis, he makes the concepts of field, ordering, and continuity (line-completeness) basic; for arithmetic, those of infinity and simple infinity play the same role. This amounts to a form of logicism insofar as Dedekind's key concepts are defined solely in terms of "logical" notions – those of object, set and function[71].

Many contemporary philosophers developed several neo-logicist and neo-fregean proposals – many of them are based upon the elimination of Basic Law V from the axioms of the *Grundgesetze* and the adoption of Hume's Principle as sole non-logical law[72]. This latter principle, indeed, has been defined adequate since it is not "inflationary", in the sense of raising the cardinality of the domain of objects. In the case of Dedekind, the situation is slightly more difficult:

> Recall that Frege also complained, rightly, about the lack of principles for "Dedekindian abstraction". But again, the observation that Dedekind did not provide the latter does not establish, in itself, that they cannot be supplied for him retroactively. [...] This suggests a general program for developing a Dedekindian form of neo-logicism – on a par with neo-Fregean forms – namely: spell out all the needed basic principles, together with motivations for them (their epistemic sources, connections to mathematical practice, etc.)[73].

Anyway, here, our aim is not to reconstruct a Dedekindian approach to abstraction, therefore, we limit ourselves by observing that in the historical development of

[69] Reck 2013a, p. 261.

[70] In any case, for both authors, "logic" includes considerations on extensions or systems. See, for instance, Reck 2013b, p. 153.

[71] Reck 2013a, p. 255.

[72] «[...] C. Parsons (1965) was the first to note that Hume's Principle was powerful enough for the derivation of the Dedekind/Peano axioms. Though Wright (1983) actually carried out most of the derivation, Heck (1993) showed that although Frege did use Basic Law V to derive Hume's principle, his (Frege's) subsequent derivations of the Dedekind/Peano axioms of number theory from Hume's Principle never made an essential appeal to Basic Law V» (Zalta 2018, p. 29).

[73] Reck 2013a, p. 262.

logicism and set theory has, in some sense, left Dedekind's figure at a second stage. We believe that, if Dedekind's mathematical and conceptual work would be put at the same level of Frege's and Cantor's, there could be a more deep understanding of the early foundations of mathematics. In particular, it is noteworthy that, thanks to the axiomatization of set theory, Dedekind's systems and Cantor's collections will become equivalent. Additionally, we believe that an abstractionist point of view might be helpful in developing philosophical considerations concerning mathematical objects and theories. From this point of view, the contemporary philosophical logic and philosophy of mathematics have been rigidly influenced by the leading ideas of Frege's works. What it might be suggested by our investigations is that also Dedekind's work merits to be revaluated and reconsidered.
Finally, let me consider just one last point. Our discussion started by considering Frege's ontological view of logical objects, but we did not say anything for what concerns Dedekind's philosophical assumptions. In contemporary discussions there is a general agreement upon the thesis for which Dedekind has been the first "non eliminative" or "eliminative" structuralist. Indeed, if we get a look back to his treatment of sets and numbers, it is easy to see that, what is fundamental is the existence of the "simply infinite" system **N**, representing a system reduced to the general structure **X**. What generally is in quest is the fact of whether the elements of **N** or of **X** are abstract objects or should be considered simply as "places" or "positions" of the structure. Indeed, much of the criticism deserved to Dedekind's abstract and structuralist approach is that in his characterization of arithmetic as a branch of a more larger systems theory there are no elements such as the natural numbers. Hence,

> Unlike Frege's, Dedekind's natural numbers have no properties other than their positions in the ordering determined by their generating operation, and those derivable from them; the question is whether such a conception is coherent[74].[...]
> He thinks, further, that if these numbers are to be specific objects, they must possess properties other than the purely structural ones they have in virtue of their positions in the sequence; but that is just what Dedekind would deny. He believed that the magical operation of abstraction can provide us with specific objects having only structural properties[75].

Probably, Dedekind would have agreed with Dummett for saying that mathematical "entities" have only pure structural properties, but he would have pointed out that (i) we "create" the system (or the cut, in the case of analysis) starting from logical notions, and (ii) that the natural numbers (or the irrational numbers) shouldn't be identified with such new "creations". Therefore, our **N** is not *the* natural number sequence itself, but it is what mathematicians call its "faithful representation"[76].

[74]Dummett 1995, p. 51.
[75]Dummett 1995, p. 52.
[76]We will return at length on this concept once the axiomatic setup of set theory has been established. See Chapter 3, section "Reconsidering Benacerraf's thesis II", paragraph "Dedekind-

Therefore, Dummett's criticism is not that precise, since it leaves out from the discussion, very deep and profound intuitions of Dedekind. Let's read Dedekind's own words:

> (Letter to Lipschitz 1876) I show, without bringing in any foreign notions, that in the realm of the rational numbers a phenomenon can be identified (the cut) that can be used, by a single creation of new, irrational numbers, to complete the realm[77].

> (Letter to Weber 1888) [...] you say that the irrational number is nothing else than the cut itself; whereas I prefer to create something new (distinct from the cut), something that corresponds to the cut, and of which I say that it produces the cut[78].

From these passages it is not possible to draw solely a "psychologistic" reading of Dedekind's abstraction and, indeed, the notion of "creation" can be understood as a metaphor: the operation of abstraction is not concerned with human and particular thoughts, but becomes engaged with a "realm" of logical objects that can be picked out exactly by abstraction principles. In this sense, the "eliminative" structuralist perspective does not hold so strongly and

> It might seem that Dedekind abstraction must give us what we want, since it functions by asserting that there is something that satisfies our desiderata. And indeed, Dedekind abstraction is closer to what most non-eliminative structuralists appear to have in mind than Frege abstraction.[79].

## 2.5 Lightened or Heavy Platonism?

In the subsection "Ontological remarks" we've said that in some context it could be useful to invoke the "existence" of mathematical entities. In the paragraphs devoted to Frege and Dedekind we've noticed that mathematical objects can be introduced by well-defined "abstraction principles" and, to be precise, "Abstractionism" – i.e., the adoption of abstraction principles – does not immediately imply a definite philosophical reading. For a first characterization of what is generally understood under the term "Abstractionism", let's consider Ebert and Rossberg description:

> Abstractionism in the philosophy of mathematics has its origins in Gottlob Frege's logicism – a position Frege developed in the late nineteenth and early twentieth century.[80] [...]
>
> So understood, we can regard neo-Fregeans among the main proponents of *Abstractionism*: the view that abstraction principles play a crucial role in the proper foundation of arithmetic, analysis, and possibly other areas of mathematics. Abstractionism therefore has two main aspects, a *mathematical* and a *philosophical* one. The main aim of the mathematical aspect of any abstractionist program is the *mathematics of*

style reflections on 'mathematical representations'", of the present work.

[77]Quoted in Linnebo and Pettigrew 2014, p. 279.

[78]Dedekind 1888a, p. 835.

[79]Linnebo and Pettigrew 2014, p. 279.

[80]Ebert and Rossberg 2017, p. 3.

> *abstraction* – simply put: proving mathematical theorems about abstraction principles or taking abstraction principles as basic axioms and investigating the resulting theories. A primary aim is to capture various mathematical theories, such as arithmetic, analysis, complex analysis, or set theory as deriving from a few basic abstraction principles and (versions of) higher-order logic. Frege's Theorem is one of the most import- ant results for a mathematical Abstractionist and numerous other interesting results have been discovered since. Philosophical Abstractionism covers, broadly speaking, three philosophical topics: semantics, epistemology, and ontology.[81]

We believe, indeed, that – just by introducing the debate here – the defence (of a form) of Platonism fits with the usage of abstraction principles. We expose, for the moment, three points which sketches the main differences between two forms of Platonism and that should already indicate to the reader in which sense we think that the presence of abstraction principles, combined with a specific Platonistic reading, might render noteworthy considerations in the philosophy of mathematics.

**Two types of Platonism?** In this paragraph we will pause for a moment on the main differences between a *lightened* or a *heavy* Platonism – from an ontological, metaphysical[82] and epistemological point of view.

| | **Heavy Platonism** | **Lightened Platonism** |
|---|---|---|
| **Ontology** | Abstract objects | Abstract objects |
| **Metaphysics** | "Robust" sense of existence | "Thin" sense of existence |
| **Epistemology** | Mathematical intuition | Abstraction principles |

As it should result from the tabular above, both – light and heavy Platonist philosophers – have an ontology which includes "mathematical objects", for instance, both, Gödel and Frege, explicitly admitted the existence of abstract entities in their ontologies of mathematics. The first main difference lies, in any case, in the metaphysical characterization of the "entities" that they invoke. What we call "robust" sense of existence is in some sense opposed to that sense of existence that we've named as "thin": a Gödel-like (heavy) Platonist, indeed, thinks that abstract mathematical objects exist as physical bodies but without spatio-temporal collocation. A light

[81] Ebert and Rossberg 2017, p. 5.

[82] Since terms such as "metaphysics" and "ontology" are overloaded of meanings, here we restrict our usage to the methodological suggestions contained in Quine 1948. In this spirit, ontology and metaphysics are strictly connected and, while the first should answer the question "what there is", the latter investigates "what is what there is". More simply, an ontology tells us which objects exist (if any), while a metaphysics tries to explain the main characters of the objects included in the ontology.

Platonist, instead, thinks that the only way that we have to characterize "abstract objects" is by saying that their exist in virtue of our mathematical practices and theories. For example, indeed, Frege thought that the "fact" that the number of knives on the table is equinumerous to the number of pens on the same table, guarantees the existence of the "logical objects" denoting the "natural numbers". Dedekind, on the same line of thought, retained that the notion of "object" is primitive and "logical", and that their existence is justified by the purpose of obtaining a clear notion of "counting" sequences. The metaphysical status that the two different philosophical point of views defend are the main responsible for the epistemologies that the two kind of Platonists endorse. Starting from the first chapter, we've said that while defending a "robust" sense of existence, a Platonist – as Gödel himself – will be forced in accepting a sort of epistemology by intuition. Recall, indeed, that for Gödel, sets can be "perceived" by human minds and that there are "strong" epistemological implications that can be drawn from the Incompleteness Theorems for arithmetic. The main criticism upon this position has been exposed following the so-called Benacerraf-Field dilemma. Differently, instead, a light Platonist will employ different principles which will devoted to pick out the main mathematical objects. So, for examples, Frege introduced a particular object – namely that of "cardinal" – by specifying what we are doing while "counting" and by formulating a formal and logical principle meant to govern and pick out the reference of each logical object – in this case, for instance, Hume's Principle.

**Light Existence and Abstraction Principles** Notice that Gödel complaint the absence of a correct and precise formalization of Platonism and suggested that this lack has origins in the absence of an explicit and formal clarification of metaphysical and ontological concepts. Contemporary forms of Abstractionism, instead, could provide useful tools to discuss the "existence" of abstract entities and this can recover the "lack" Gödel was complaining. For what concerns the metaphysical characterization of the entities we invoke in an ontology of mathematics, Gödel thought that they were identical to physical objects but without any "physical characterization" and defective of casual powers. Recall, that the way we get knowledge of them is by "perception" or "intuition". Hence, to elaborate a mathematical theory – that is, to explicate a new part of the mathematical realm – means to "perceive" some abstract entities with some tools that we already possess. In this sense, the mathematician's aim is, therefore, to capture the entities, by defining them or by showing their existence. Differently, for a light Platonist the abstract entities "exist" in the sense that their ontological introduction is helpful for fixing our reference to mathematical objects. What guarantees the existence and the "perception" of those entities are the so-called abstraction principles. For example, we see that while counting things, we are putting some objects into a precise relation with some other elements, such as the natural numbers (See Figure 2.3, next page). This allows us, then, to define a general law which introduces a new object, by considering the cases in which the

$$\begin{array}{cccccccccc} a & b & c & d & e & f & g & h & i & \dots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \dots \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \dots \end{array}$$

Figure 2.3: Counting as a one-to-one relation

relation considered obtains.
So, generally:

$$\forall \alpha,\, \beta\, (\S\alpha = \S\beta \longleftrightarrow R\, (\alpha, \beta))$$

That is, the identity relation between the elements – characterized by an abstraction operator § – in the left-hand side of the biconditional, holds iff the equivalence relation, $R$, stated in the right-hand side, holds. In this way, a new logical object is introduced[83].

**Light Existence and Reference** Finally, epistemology will be defined by the same abstraction principles and, hence, to know a logical object means to know the conditions under which its "thin" existence is granted. Therefore, our knowledge and reference of such logical objects is assured thanks to the abstraction principles. These latter, indeed, define also the "thin" sense of their existence (metaphysics): if an abstraction principle will turn out incoherent or inconsistent – consider the case of Basic Law V –, nothing assures any more that the objects it postulates exist.
So – summarizing our achievements – we could state that from an ontological point of view, logical objects are coherent with some mathematical practices – consider the cases of arithmetic or of set theory. Metaphysically, we are trying to argue that they exist in an abstract manner and that, perhaps, the parallelism between their existence and that of the physical bodies is a bit misleading. Their existence is not "robust" (spatio-temporal and physically characterized), but is to be considered as "thin" ("simply" subject to precise logical abstraction principles). We have argued that, for example, for the "directions" to exist, it suffice that it exists a line from which the logical objects themselves – the directions – have been abstracted. Additionally, these latter laws allow us to pick out the reference of the objects that our ontology postulates, without employing the problematic notion of "mathematical perception". In this way, metaphysics and epistemology become strictly related and their explications obtain in virtue of the same logical abstraction principles.

[83]By instantiating the variables involved in the previous formula, we recognize the general form of rules such as Hume's Principle, $\#F = \#G \longleftrightarrow F \sim G$, or the Law for Directions, $d(\ell_1) = d(\ell_2) \longleftrightarrow \ell_1 \,||\, \ell_2$.

**First conclusions** The comparison we've sketched – as clear – is not complete at all and, indeed, much of the considerations concerning Abstractionism and Light conceptions of existence will be developed at length within Chapters 4-5. Anyway, – before turning to that point definitely –, to undermine Benacerraf's philosophical conclusions and to attempt to re-introduce abstract objects into the ontology of mathematics, through abstraction principles, we will consider again Benacerraf's logical and mathematical premises, but, this time, shifting our attention to the axiomatic, rather than to the naïve, version of set theory, so as to finally encounter, from a closer perspective, Zermelo's and von Neumann's set theoretical representations of the natural numbers.

# Chapter 3

# Numbers, sets and objects II. Axiomatic Considerations

**Overview.** In this chapter, we are going to deepen our inquiry into set theory to get an always clearer insight for what concerns Benacerraf's ontological thesis. Recall that Benacerraf's main point was that of eliminating any sort of realism within the philosophy of mathematics by starting his considerations from the apparent irreducibility of numbers to sets. Already in the foregoing chapter we've encountered a way in which, also within a naïve framework, Benacerraf's argumentation does not seem so conclusive as it should be. In this part of the thesis, we will enforce our previous argumentation strategy by employing axioms for sets and by finally considering von Neumann's and Zermelo's representations for sets as reducible (in some determined way) to the sequence of natural numbers. Moreover, recall that Benacerraf's set-theoretical argument was directed to a more general aim: he wanted to show that the adoption of abstract objects, within an ontology of mathematics, is troublesome by, consequently, rejecting any form of mathematical Platonism. In this spirit, our aim will be at the opposite side: starting from set-theoretical considerations – as Benacerraf himself did – we will be led in considering the possibility of having abstract objects within our ontology of mathematics, trying, finally, to restore and defend a more Platonist approach to mathematics.

## 3.1 One Number, Different Sets!

**Summary.** If we reconsider Benacerraf's article we may easily see that his argument is not directly concerned with naïve set theory and, indeed, his main criticism focuses upon the axiomatic version(s) of the theory. Moreover, he considers two young adolescents that have learned the same logic fundamentals but have taken two different set theory classes. Additionally, in those classes – Benacerraf considers – both of them have learned the structure of the natural number sequence as a part of the set-theoretical universe. But, Ernie, the young girl, has been educated to von Neumann's set-theoretical proposal, while his friend, Johnny, has apprehended

Zermelo's sets. According to Benacerraf, while comparing their (different) theories, the two adolescents will face a situation like this:

> Comparing notes, they soon became aware that something was wrong, for a dispute immediately ensued about whether or not 3 belonged to 17. Ernie said that it did, Johnny that it did not.[1]

In order to have a better insight on the motivations for which both theories, Ernie's and Johnny's, are somehow different, let's consider the main distinctions between the set-theoretical constructions they learned. Indeed, before turning to the axiomatic approaches of Zermelo and von Neumann, we shall give an intuitive and introductory idea of the main differences lying between the two sets constructions.

Ernie and Johnny agreed that numbers, such as 1, 2, 3, 4, and so on, could be reduced to sets, but disagreed on which set-theoretical representation should be considered. Indeed, Ernie, was told that when we are saying that "$x$ is the successor of $y$", we are simply representing the relation $R$ that $x$ bears to $y$. In this sense, $y$ is the set consisting of $x$ and all the members of $x$. Hence, for Ernie, sets represent numbers as follows:

$$\begin{array}{rcc}
\varnothing & =_{\text{def}} & 0 \\
\{\varnothing\} & =_{\text{def}} & 1 \\
\{\varnothing, \{\varnothing\}\} & =_{\text{def}} & 2 \\
\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} & =_{\text{def}} & 3 \\
\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}\} & =_{\text{def}} & 4 \\
\dots & \dots & \dots \\
n \cup \{n\} & =_{\text{def}} & S(n) \\
\dots & \dots & \dots
\end{array}$$

Differently, for Johnny, saying that $x$ is successor of $y$ only means that for any set $y$ we take its unit set, namely $\{y\}$, to get its successor, i.e. $x$. Hence, Johnny will represent the natural numbers progression in the following manner:

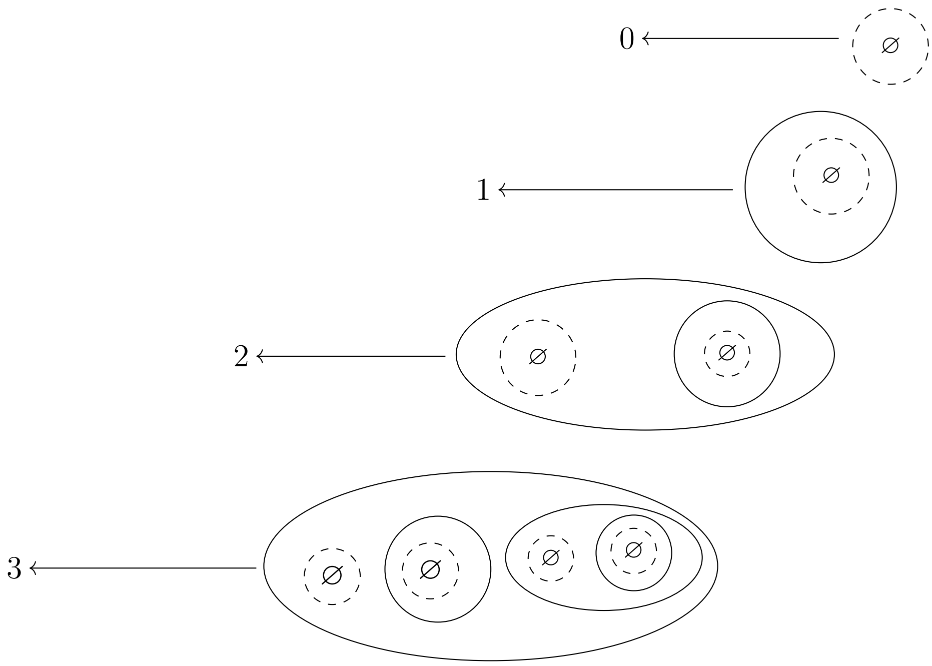
$$\begin{array}{rcc}
\varnothing & =_{\text{def}} & 0 \\
\{\varnothing\} & =_{\text{def}} & 1 \\
\{\{\varnothing\}\} & =_{\text{def}} & 2 \\
\{\{\{\varnothing\}\}\} & =_{\text{def}} & 3 \\
\{\{\{\{\varnothing\}\}\}\} & =_{\text{def}} & 4 \\
\dots & \dots & \dots \\
\{n\} = \underbrace{\{\dots}_{n+1 \text{ times}} \; \{\varnothing\} \; \underbrace{\dots\}}_{n+1 \text{ times}} & =_{\text{def}} & S(n) \\
\dots & \dots & \dots
\end{array}$$

[1] Benacerraf 1965, p. 278.

So, in this sense, one of the main differences between the two number-sets conceptions concerns the notion of "successor".
Another fundamental point of disagreement, Benacerraf writes, is the following. While Ernie thinks that any set has $n$ members just in case there is a biunivocal correspondence with the set $n$ itself, Johnny believes that every set is single membered. Indeed, the situation can be graphically understood as follows:

$$
\begin{array}{ccccc}
\varnothing, & \{\varnothing\}, & \{\varnothing,\{\varnothing\}\}, & \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}, & \dots \\
\downarrow & \downarrow & \downarrow \quad \downarrow & \downarrow \quad \downarrow \qquad \downarrow & \\
0 & 1 & \underbrace{0 \quad 1}_{2} & \underbrace{0 \quad 1 \qquad 2}_{3} &
\end{array}
$$

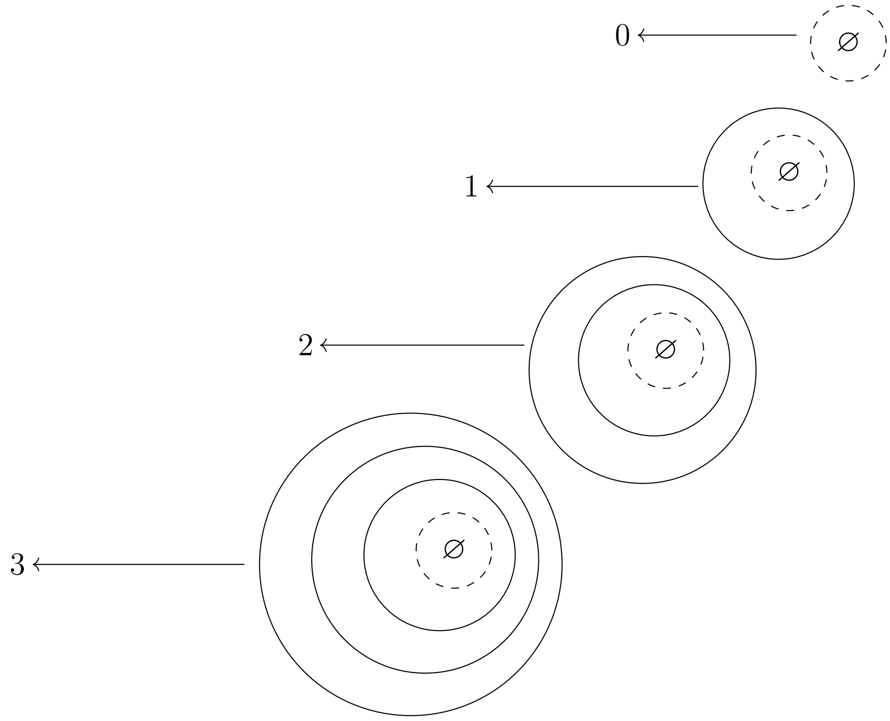
The same situation can be set-graphically presented in the following manner:




Differently, the other situation will be:

$$
\begin{array}{ccccc}
\varnothing, & \{\varnothing\}, & \{\{\varnothing\}\}, & \{\{\{\varnothing\}\}\}, & \dots \\
\downarrow & \downarrow & \downarrow & \downarrow & \\
0 & 1 & 2 & 3 &
\end{array}
$$

And, set-graphically, we may represent Zermelo's sets as follows:




So, in «short, their cardinality relations [are] different. For Ernie, 17 ha[s] 17 members, while for Johnny it ha[s] only one»[2].

Finally, this will be our investigative background: we will consider how to build Zermelo's and von Neumann's sets, discuss the axioms and follow the strategy of building equivalence classes between those sets. In this spirit, consider that the different notions of successor and cardinality are not the only problems that Benacerraf envisaged, and, indeed, what – according to the French philosopher – constitute the main philosophical disadvantage, in identifying numbers and sets, is the violation of Leibniz Law. Benacerraf's reasoning can be put in the following way:

**Remark.** For the sake of the argument, take the number-sign 2. Let's say that

$$\{\varnothing, \{\varnothing\}\} = 2 \qquad \{\{\varnothing\}\} = 2$$

By simple inspection of the properties of the two set-theoretical representations, it follows that:

$$\{\varnothing, \{\varnothing\}\} \neq \{\{\varnothing\}\}$$

Hence, for instance, consider

$$\{\varnothing, \{\varnothing\}\} = 2$$

Since the two sets representations are different, we may conclude:

$$\{\{\varnothing\}\} \neq 2$$

[2]Benacerraf 1965, p. 279.

We may also reverse the reasoning and start by considering

$$\{\{\varnothing\}\} = 2$$

As before, since the two sets are not identical:

$$\{\varnothing, \{\varnothing\}\} \neq 2.$$

Therefore, since 2, for instance, is identical to both sets – $\{\varnothing, \{\varnothing\}\}$ and $\{\{\varnothing\}\}$ – then also both sets should be identical one to another. But actually the two set-theoretical representations considered are not identical and hence, in violation of Leibniz Law, one and the same object (namely the number-sign 2) would be identical to two different and irreducible objects. Thus, according to Benacerraf, even if both theories allow the set-theoretical study of the natural numbers progression, the identification between sets and numbers has strictly to be rejected. Moreover, as we saw before, Benacerraf enforces his claim by concluding that, in an ontology of mathematics, no objects, of whatever sort, are needed.

This terminates our summary. The rest of this chapter will be devoted in seeing whether the notion of "identity" between sets and numbers, as invoked by Benacerraf, is the sole way for understanding the set-theoretical reduction of the natural numbers. We claim, indeed, that a different result can be established if the relation between sets and numbers is understood in terms of an equivalence relation, rather than the actual identity. Philosophically, this will lead us in (i) positively reconsidering some suggestions of Dedekind concerning what we usually call "mathematical representations" and (ii) open the door to the possibility of having abstract objects in our mathematical ontology.

## 3.2 Axioms for Set Theory

### 3.2.1 The Meaning of the Axiomatic Method

First of all, differently from Frege, Cantor and Dedekind, – after the appearance of Russell's Paradox –, logicians and mathematicians tried to save and secure set theory in different manners. In what follows, for our purposes, we will consider the most successful strategy, that is the method, employed during the first half of the twentieth century, of giving axioms that should govern every aspect of the set-theoretical universe. In the twentieth century, mathematicians

> [...] proposed to replace the direct *intuitions* of Cantor about sets which led us to the faulty General Comprehension Principle with some *axioms*, hypotheses about sets which we accept with little a priori justification, simply because they are necessary for the proofs of the fundamental results of the existing theory and seemingly free of contradiction.[3]

[3]Moschovakis 2006, p. 23.

Indeed, the great help of the axiomatic method was not only perceived by mathematicians and logicians, but also from philosophers. In this context, Russell's "regressive" conception of the importance of the axiomatic method, to systematize precise mathematical theories, can be conceived as a good path to understand its usefulness.

Before deepening the conception of "regressive methodology" in mathematics, as suggested by Russell, consider that axioms, within mathematics at least, have always been used since the Greeks. Historically, indeed, the main source for the axiomatic method has been Euclid and the rational tradition he started. From this early beginnings and along the entire medieval tradition, axioms have always been considered as (a) "self-evident" and (b) "epistemologically fundamental" propositions. The main troubles with (a) and (b) arise when Frege's revolutionary work, by changing the predominant conception of logic, rethought, consequently, also the traditional notion of "axiom". The paper Russell wrote in 1907, indeed, suggests to abandon the requirement for axioms to be (a) self-evident, while maintaining (b), that their content is epistemologically fundamental[4]. In this sense, Russell is suggesting not to consider axioms as self-evident, but to take them as less-evident than their consequences, and to evaluate their contents just on the basis of what they entail. So, it seems that Russell is completely reversing the Euclidean perspective: while anciently axioms were self-evident propositions, now, according to Russell, contemporary logic allows us to take them as non self-evident. In other words, logicians and mathematicians do not appeal any more just on the criteria of "self-evidence" to claim that a proposition is an axiom. But, does this mean that there are non self-evident propositions that are axioms? If yes, how do we justify their being considered axioms, if self-evidence is out of games? Let's read a very clear quote from Russell's 1907 paper:

> [...] in mathematics, except in the earliest parts, the propositions from which a given proposition is deduced generally give the reason why we believe the given proposition. But in dealing with the principles of mathematics, this relation is reversed. Our propositions are too simple to be easy, and thus their consequences are generally easier than they are. Hence we tend to believe the premises because we can see that their consequences are true, instead of believing the consequences because we know the premises to be true. But the inferring of premises from consequences is the essence of induction; thus the method in investigating the principles of mathematics is really an inductive method [...][5].

As clear, hence, «Russell claims, mathematical axioms can be justified by their ability to entail, explain and systematize more obvious mathematical propositions»[6]. Indeed,

[4]It is useful to notice that Gödel's notion of *extrinsic* form of intuition (see chapter 1, subsection "Epistemology and Mathematical Intuition") has been proposed while agreeing with Russell on the conception of axioms we are sketching in this section. For more on the dispute concerning the axiomatic method into the philosophy of mathematics see Linnebo 2011, pp. 170–182.

[5]Russell 1907, pp. 273–274.

[6]Linnebo 2011, p. 174.

in this general climate – as remarked – logicians, mathematicians and philosophers tried to give a solid axiomatization of Cantor's intuitions about sets. The first successful systematization of set theory has been proposed by E. Zermelo in his famous article *A new proof of the possibility of a well-ordering* (1908) and has been bettered by J. von Neumann in the paper *Introduction to transfinite numbers* (1923). Anyway, as reminded above, in both cases, the:

> [...] basic model for the axiomatization of set theory was Euclidean geometry, which for 2000 years had been considered the "perfect" example of a rigorous, mathematical theory. If nothing else, the axiomatic method clears the waters and makes it possible to separate what it might be confusing and self-contradictory in our intuitions about the objects we are studying, from simple errors in logic we might be making in our proofs[7].

In order to get a clearer idea on how the axiomatic framework of set theory escaped Russell's Paradox, we will first state and discuss some of the most important mentioned axioms.

### 3.2.2 Towards an Axiomatization. Open Problems: 1880-1930

***What is a set?*** Cantor's and Dedekind's set theories had both been declared affected by a contradiction and the fault of the inconsistency has been brought back to two fundamental features of what we've called naïve set theory, namely the fact that any collection could have determined a set and the lack of definite and precise conditions under which those collections could have been studied. Recall, indeed, that Cantor, among other mathematicians of the 19th century, defined sets in the following general manner: by «an "aggregate" (*Menge*) we are to understand any collection into a whole (*Zusammenfassung zu einem Ganzen*) $M$ of definite and separate objects of our thought. These objects are called the "elements" of $M$»[8]. From this informal definition, clearly, the faulty General Comprehension Principle for sets can be derived and, hence, forms of Russell's Paradox can be obtained within our set theory. In order to avoid inconsistencies, hence, mathematicians after Cantor, worked along a more precise and, in some sense, restricted notion of collection, beginning thus the "heavy" mathematization of sets:

> Set theory is an autonomous field of mathematics, enormously successful not only in its continuing development of its historical heritage but also at analysing mathematical propositions and gauging their consistency strength. But set theory is also distinguished by having begun intertwined with pronounced metaphysical attitudes, and these have been regarded as crucial by some of its great developers[9].

But, with the mathematization, «set theory has proceeded in the opposite direction,

[7]Moschovakis 2006, p. 23.
[8]Cantor 1915, p. 85.
[9]Kanamori 2009, p. 1.

from a web of intensions to a theory of extension *par excellence*, and like other fields of mathematics its vitality and progress have depended on a steadily growing core of mathematical proofs and methods, problems and results»[10].

***Are sets well-ordered?*** So, – except the task of finding a precise definition of sets – mathematicians sought also for a theory able to encapsulate (at least) three fundamental questions conjectured by Cantor. It is useful to point out that, while proposing his theory of collections, Cantor was trying to solve a problem he envisaged in 1883:

> The concept of *well-ordered set* turns out to be fundamental for the entire theory of manifolds. [...] I shall discuss the law of thought that says that it is always possible to bring any *well-defined* set into the *form* of a *well-ordered* set – a law which seems to me fundamental and momentous and quite astonishing by reason of its general validity[11].

Hence, for Cantor, already in 1883, one of the fundamental problems concerning sets was that of finding a proof for the conjecture that "every set has a well-ordering". Even if the meaning of "well-ordering" for sets of elements, that Cantor suggested, is very close to the intuitive ideas concerning orderings, it is useful to report the mathematician's idea:

> A *well-ordered* set is a well-defined set in which the elements are bound to one another by a determinate given succession such that (i) there is a *first* element of the set; (ii) every single element (provided it is not the last in the succession) is followed by another determinate element; and (iii) for any desired finite or infinite set of elements there exists a determinate element which is *their immediate successor* in the succession (unless there is absolutely nothing in the succession following them all).[12]

Several years after Cantor's publishing, at the International Congress for Mathematicians in Paris, D. Hilbert announced to the world the importance for mathematics to find a proof for the well-ordering theorem. Things began to change during the year 1904 when, at the Third International Congress of Mathematicians in Heidelberg, J. König presented what he considered to be a proof of the fact that the continuum has no well-ordering. Anyway, less than two months later, König himself withdrew his own proof and Zermelo completed and presented his proof of the existence of the conjectured well-ordering of sets[13]. We will focus on Zermelo's work in the next paragraphs.

***What are cardinal and ordinal numbers?*** Cantor formulated his theory of *Menge* in order to capture the most fundamental mathematical notions, such as the

[10]Kanamori 2009, p. 1.
[11]Cantor 1883, p. 886.
[12]Cantor 1883, p. 884.
[13]For the historical development of set theory after Zermelo's papers, see Kanamori 2009, pp. 29–47.

one of number. In the foregoing chapter, we've seen how it is actually possible to define a natural number, such as 2, just by considering collections of things. Cantor's interest, however, has been that of defining not just the notion of "number", but also the two more complicated notions of *ordinal* and *cardinal* numbers.
For what concerns sets and their cardinality, in 1895, the German mathematician wrote:

> Every aggregate $M$ has a definite "power", which we will also call its "cardinal number". We will call by the name "power" or "cardinal number" of $M$ the general concept which, by means of our active faculty of thought, arises from the aggregate $M$ when we make abstraction of the nature of its various elements $m$ and of the order in which they are given. We denote the result of this double act of abstraction, the cardinal number or power of $M$, by
>
> $$|M|.$$
>
> Since every single element $m$, if we abstract from its nature, becomes a "unit", the cardinal number $|M|$ is a definite aggregate composed of units, and this has existence in our mind as an intellectual image or projection of the given aggregate $M$.[14]

And, secondly, with respect of the ordinals, which Cantor used to call "ordinal type", he wrote:

> Every ordered aggregate $M$ has a definite "ordinal type", or more shortly a definite "type", which we will denote by
>
> $$\overline{M}.$$
>
> By this we understand the general concept which results from $M$ if we only abstract from the nature of the elements $m$, and retain the order of precedence among them. Thus the ordinal type $\overline{M}$ is itself an ordered aggregate whose elements are units which have the same order of precedence amongst one another as the corresponding elements of $M$, from which they are derived by abstraction. [...]
> A simple consideration shows that two ordered aggregates have the same ordinal type if, and only if, they are similar, so that, of the two formulae
>
> $$\overline{M} = \overline{N},\ M =_o N$$
>
> one is always a consequence of the other[15].

Informally, therefore, we may understand a cardinal number as telling us the "size" of a set, that is the number of its elements, and the ordinal numbers as informing us on the "positions" elements have in sets.

Consider that other problems affected naïve set theory, but, for our purposes, considering just these few, even if demanding, concerns is enough. Indeed – as done in the

[14]Cantor's notation for the cardinal number of $M$ was $\overline{\overline{M}}$. In what follows, we adopt the more conventional notation as given in the quote. See, Cantor 1915, p. 86.

[15]The notation we used preserves just in part Cantor's notation: the formula $M =_o N$ was not expressed in this way by Cantor, but it expresses the same proposition, that is $M$ is identical, with respect to the ordinal type, to $N$. See, Cantor 1915, pp. 111–112.

chapter before –, we restrict our attention on how the situation has been faced up, after the discover of Russell's Paradox, by mathematicians. Treating, among others, the questions we've mentioned – defining a set, proving the well-ordering theorem and submitting to inspection cardinals and ordinals –, within an axiomatic framework, will give us a clear idea of *why* sets – as is standardly accepted in mathematics – can accomplish the mission of representing the natural numbers sequence.
Additionally, it is useful to point out that, while we are going to focus the attention on how mathematicians, such as Zermelo and von Neumann, thought possible to "save" set theory, at the same time, we are also accomplishing our mission of "faithfully representing" the natural numbers sequence within a set-theoretical framework. Hence, – as in the previous chapter – the only formalism and mathematical tools we are going to introduce are those that will help us in shedding more light on the doubts we've raised with respect to Benacerraf's criticism.

### 3.2.3 Zermelo's Axioms

#### 3.2.3.1 Becoming Precise! (Mathematization)

We begin by describing how E. Zermelo began his axiomatic foundation of sets. The ingredients of the receipt are:

(i) A **domain** or **universe** $\mathcal{Z}$ of objects;

(ii) Some **definite conditions**, among them:

1. **Identity**: $x = y \Longleftrightarrow x$ is the same object as $y$.
2. **Sethood**: $\text{Set}(x) \Longleftrightarrow x$ is a set.
3. **Membership**: $x \in y \Longleftrightarrow \text{Set}(y)$ and $x$ is an element of $y$.

(iii) Some **definite operations**.

We allow all objects to be sets and, if some objects are not sets, then we call them **atoms** of $\mathcal{Z}$. We point out that definite conditions and operations are neither sets nor atoms. At this point, one of the fundamental distinctions between Zermelo's axiomatic approach and the traditional Greek one, emerges: «Euclidean geometry is quite complex: there are several types of objects and a long list of intricate axioms about them. By contrast, Zermelo's set theory is quite austere: we just have sets and atoms and only seven fairly simple axioms relating them»[16].
Finally, let's take care about the axioms Zermelo proposed:

**Axiom** (**I**. Axiom of Extensionality)**.** «If every element of a set $A$ is also an element of $B$ and viceversa, if, therefore, both $M \subset B$ and $B \subset A$, then always $A = B$; or, more briefly: Every set is determined by its elements»[17]. Formally:

$$\forall A, B[A = B \longleftrightarrow \forall x(x \in A \longleftrightarrow x \in B)]$$

[16]Moschovakis 2006, p. 23.
[17]Zermelo 1908b, p. 201.

**Axiom** (**II**. Emptyset, Singleton and Pairset Axiom)**.** (a) «There exists a (fictitious) set, the *null set*, $\varnothing$, that contains no element at all»[18], i.e.:

$$\exists A \forall x (x \notin A)$$

(b) «If $x$ is any object of the domain, there exists a set $\{x\}$ containing $x$ and only $x$ as element»[19] and (c) «if $x$ and $y$ are any two objects of the domain, there always exists a set $\{x, y\}$ containing as elements $x$ and $y$ but no object $u$ distinct from both»[20], namely:

$$\forall x, y \ \exists A \ \forall u \ (u \in A \longleftrightarrow u = x \lor u = y)$$

**Remark.** Every single membered set, such as $\{x\}$, is called **singleton**. Furthermore, the axiom of extensionality, **I**, implies that there is only one empty set and, moreover, that for any two object $x, y$ there is just one set that can satisfy the paring axiom. As Zermelo remarked, we denote it by $\{x, y\}$ and call it a **doubleton**. Let's say that $a$ and $b$ are two any object of our domain. Suppose additionally that $a = b$. By the paring axiom, **II**, the doubleton that collects $a$ and $b$ – denoted by $\{a, b\}$ –is the singleton $\{a\}$ of the object $a$. Finally, as last part of this remark, consider that now we are able to begin constructing sets such as:

$$\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\{\varnothing\}\}\}, \ldots$$

But, notice, each of them must have at most two members and, in order to construct different sets, with more than two elements, we will need the entire Zermelian axiomatization plus another fundamental notion we will encounter soon.

**Axiom** (**III**. Separation Axiom)**.** «Whenever the propositional function $P(x)$ is definite for all elements of a set $A$, $A$ possess a subset $B_{P(x)}$ containing as elements precisely those $x$ of $A$ for which $P(x)$ is true»[21]. In more contemporary terms, for any set $A$ and each unary, definite condition $P$, there exists a set $B$ which collects all objects that are in $A$ and that satisfy $P$, namely:

$$\forall A \ \forall P \ \exists B [x \in B \longleftrightarrow x \in A \land P(x)]$$

**Remark.** As before, by axiom **I**, it follows that only one set $B$ can satisfy the axiom of separation. Let's denote the sets, such as $B$, as follows:

$$B = \{x \in B \mid P(x)\}.$$

This fundamental axiom, separation **III**, is of much help in our theory since it helps us to escape from Russell's Paradox and to banish any so-called universe set of sets:

[18]Zermelo 1908b, p. 202.
[19]Zermelo 1908b, p. 202.
[20]Zermelo 1908b, p. 202.
[21]Zermelo 1908b, p. 202.

**Theorem 3.2.1.** (a) For each set $A$, its Russell set:

$$R(A) =_{\text{def}} \{x \in A \mid x \notin A\}$$

is not a member of $A$.
(b) The collection of all sets is not a set, i.e.:

$$\neg\exists V[(x \in V) \longleftrightarrow \text{Set}(x)]$$

*Proof.* Consider $A$ and its Russell set, namely $R$

$$R(A) =_{\text{def}} \{x \in A \mid x \notin A\}.$$

$R(A)$ exists by the separation axiom, **III**. Now, assume, $R(A) \in A$. As in all other cases we obtain:

$$R(A) \in R(A) \longleftrightarrow R(A) \notin R(A)$$

Which is, another time, inconsistent. ■

Indeed, while considering his axioms, at the beginning of the 1908 paper, Zermelo announced that something, that assured the restriction of the faulty Comprehension Principle, was needed:

> [...] in view of the "Russell antinomy" of the sets of all sets that do not contain themselves as elements, it no longer seems admissible today to assign an arbitrary logically definable notion a set, or class, as its extension. Cantor's original definition of a set [...] therefore certainly requires some restriction.[22]

With this precise aim – immediately after having proposed the axiom of separation, **III** – Zermelo sums up the advantages of its adoption:

> [...] Axiom III in a sense furnishes a substitute for the general definition of set that was cited in the introduction and rejected as untenable [see above]. It differs from that definition in that it contains the following restrictions. In the first place, sets may never be *independently defined* by means of this axiom but must always be *separated* as subsets from sets already given; thus contradictory notions such as "the set of all sets" or "the set of all ordinal numbers" [...] are excluded. In the second place, moreover, the defining cirterion must always be definite [...] (that is, for each single element $x$ of $A$ the fundamental relations of the domain must determine whether it holds or not). [...] But it also follows that we must, prior to the application of our Axiom III, prove the criterion $P(x)$ in question to be definite, if we wish to be rigorous [...].[23]

**Axiom (IV**. Powerset Axiom**).** «To every set $A$ there corresponds another set $\wp(A)$, that contains precisely all subsets of $A$»[24], namely:

$$\forall A\, \exists B\Big[C \in B \longleftrightarrow [\text{Set}(C) \wedge \forall u(u \in C \to u \in A)]\Big]$$

[22] Zermelo 1908b, p. 200.
[23] Zermelo 1908b, p. 202.
[24] Zermelo 1908b, p. 203.

As usual, we call the set of all subsets of a set $A$ the **powerset** of $A$ and we denote it as follows:

$$\wp(A) =_{\text{def}} \{C \mid \text{Set}(C) \wedge C \subseteq A\},$$

where the notation $C \subseteq A$ abbreviates the formula $\forall u(u \in C \to u \in A)$.

**Axiom** (**V**. Unionset Axiom)**.** «To every set $A$ there corresponds a set $\bigcup(A)$, the union of $A$, that contains as elements precisely all elements of $A$»[25], formally:

$$\forall A, \exists B\Big[x \in B \longleftrightarrow (\exists C \in A)[x \in C]\Big].$$

We call any set $B$ the **unionset** of $A$ and we denote it by

$$\bigcup(A) =_{\text{def}} \{x \mid (\exists C \in A)[x \in C]\}$$

As always, axiom **I** implies that any unionset of a set is uniquely determined. Moreover, it is useful to notice that with **V**, the unionset axiom, we are able to define the binary, union operation between sets. Anyway, in order to complete Zermelo's axiomatic approach, we do need to state the two most controversial axioms the mathematician proposed and see how their presence has been justified.

**Axiom** (**VI**. Axiom of Choice)**.** «If $A$ is a set whose elements are all sets that are different from $\varnothing$ and mutually disjoint, its union $\bigcup(A)$ includes at least one subset $B_1$ having one and only one element in common with each element of $A$». Moreover, we «can also express this axiom by saying that it is always possible to *choose* a single element from each $C, D, E, \ldots$ of $A$ and to combine all the chosen elements, $c, d, e, \ldots$, into a set $B_1$»[26]. Formally,

$$\forall A\Big[\varnothing \notin A \longrightarrow \exists f : A \to \bigcup(A), \forall B \in A\ (f(A) \in A)\Big]$$

In other terms, for any set $A$ of nonempty sets, there exists a **choice function** $f$ defined on $A$.

**Remark.** The main usage and apllication of the axiom of choice, **VI.**, has that of permitting, among other fundamental things, the proof of the well-ordering theorem. We will discuss the main criticism deserved to axiom **VI.**, and to its application within a proof, in the next subsection. For the moment it is important to understand the meaning of the axiom itself and to relate it to the context of the general axiomatization of set theory which Zermelo was seeking for. The example we will use here is taken from Russell's work. Consider $A$ as our set of pairs of "shoes" (both, left and right) and let $B = \bigcup(A)$. Additionally, let $R(a, b)$ hold just in case

[25]Zermelo 1908b, p. 203.
[26]Zermelo 1908b, p. 204.

$a \in b$. We next define a function $f(a) =_{\text{def}}$ "the left shoe in $a$", for $a \in A$. In this sense, $f$ *selects* a shoe from each pair, namely $\forall a \in A\ R(a, f(a))$. We will call such a function $f$, a **choice function**.
However, if we understand $A$, for instance, as a set of "socks" (assumed to have no distinguishing features), we have no way to define a function $f : A \rightarrow \bigcup(A)$, which selects one $f(a) \in a$ from each pair, apart invoking the axiom of choice. In Russell's own words:

> Among boots we can distinguish right and left, and therefore we can make a selection of one out of each pair, namely, we can choose all the right boots or all the left boots; but with socks no such principle of selection suggests itself, and we cannot be sure, unless we assume the multiplicative axiom, that there is any class consisting of one sock out of each pair.[27]

First of all, notice that, with the name "multiplicative axiom", Russell called what here we've named axiom of choice. Anyway, by pursuing his analysis of this axiom, Russell wrote, with respect to the possibility of an ordering of an infinite set of pairs of objects:

> There is no difficulty in doing this with the boots. The pairs are given as forming an $\aleph_0$, and therefore as the field of a progression. Within each pair, take the left boot first and the right second, keeping the order of the pair unchanged; in this way we obtain a progression of all the boots. But, with the socks we shall have to choose arbitrarily, with each pair, which to put first; and an infinite number of arbitrary choices is an impossibility. Unless we can find a rule for selecting [...].[28]

And the rule – Russell thinks as appropriate to avoid a «number of arbitrary choices» – is, indeed, the axiom of choice.

Finally one last axiom is somehow needed:

**Axiom (VII.** Axiom of Infinity**).** «There exists in the domain at least one set $\mathbb{I}$ that contains the null set as an element and so is constituted that to each of its elements $a$ there corresponds a further element of the form $\{a\}$, in other words, that with each of its elements $a$ it also contains the corresponding $\{a\}$ as an element»[29]. Formally:

$$\exists \mathbb{I}\ [\varnothing \in \mathbb{I} \land \forall x\ (x \in \mathbb{I} \longleftrightarrow \{x\} \in \mathbb{I})]$$

This means that, there is a set $\mathbb{I}$ that collects the empty set and the singleton of each of its members.

**Remark.** Even if we do not possess yet a precise and rigorous definition of "infinite", it is not difficult to see that axiom **VII.** implies that $\mathbb{I}$ is infinite: it should be clear

[27] Russell 1919, p. 126.
[28] Russell 1919, p. 126.
[29] Zermelo 1908b, p. 204.

that any set with the properties of $\mathbb{I}$, as expressed by the axiom, must be infinite, since:

$$\varnothing \in \mathbb{I},\ \{\varnothing\} \in \mathbb{I},\ \{\{\varnothing\}\} \in \mathbb{I},\ \{\{\{\varnothing\}\}\} \in \mathbb{I}, \ldots$$

Moreover, notice that, it follows that, any of the sets $\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \ldots$ is, by axiom **I.** – namely Extensionality – distinguished by any other set. In this context, it's important to notice that, the postulation of something such as the "infinite set" of axiom **VII.**, could lead our Zermelian foundations of set theory into deep (philosophical) troubles. Indeed, Zermelo himself thought that axioms **I-VI** «suffice, as we shall see, for the derivation of all essential theorems of general set theory. But in order to secure the existence of infinite sets we still require the [preceding] axiom, which is essentially due to Dedekind». Moreover, Zermelo – whose intent was that of treating rigorously Cantor's and Dedekind's intuitions – with this axiomatic warranty, namely the existence of infinite sets, could accomplish the mission of representing the natural number progression set-theoretically. Consider that,

> [t]he set $\mathbb{I}$ contains the elements $\varnothing, \{\varnothing\}, \{\{\varnothing\}\}$, and so forth, and it may be called the *number sequence*, because its elements can take the place of numerals. It is the simplest example of a denumerably infinite set.[30]

With all the necessary background, indeed, Zermelo – at the end of his 1908 paper – proved the theorem for which the number sequence, as identified in the previous quote, is an infinite set. We will return back to this point closely in the next subsection.

#### 3.2.3.2 Ordered Pairs: from Frege to Kuratowski

A necessary historical and theoretical note, at this point of the discussion, must be given. Consider, for instance, that, usually, the notation $(x, y)$ indicates the *ordered pair* of $x$ and $y$. This means that $(x, y)$ represents the set that has $x$ as "first" member and $y$ as "second" element. Recall, now, that the Pairing axiom, **II**., enabled us to form (unordered) pairs such as, $\{x, y\}$.But, since, unordered pairs are given to us from axiom **II**., how should we introduce ordered pairs in our set-theoretical framework? Clearly, logicians and mathematicians – at the beginning of the century – tried to determine an operation, between sets, which would helped them in forming ordered sequences of elements. As clear form the informal characterization of pairs, hence, the main difference between ordered and unordered ones is that $\{x, y\} = \{y, x\}$, while $(x, y) \neq (y, x)$. Therefore, formally, the desired operation must respect the following condition:

(**C1**). $(x, y) = (z, u) \longleftrightarrow x = z \ \wedge \ y = u$

Moreover, we have to ensure that the ordered pair $(x, y)$ we have defined, is a set. This fact allows us to define the notion of "Cartesian product" of sets, which is a basic and very useful notion of set theory. Indeed, for every two sets $A$ and $B$, we define their Cartesian product as follows:

[30] Zermelo 1908b, p. 205.

(**C2**). $A \times B =_{\text{def}} \{(x, y) \mid x \in A \ \wedge \ y \in B\}$

As Moschovakis remarked, «the problem of representing the notion of "pair" in set theory takes the following precise form: we must define a definite operation $(x, y)$ such that [(**C1**)] and [(**C2**)] follow from the axioms of Zermelo»[31]. Anyway, – before turning to the now widely accepted solution to the problem – we will briefly give some indications on how the debate surrounding ordered pairs was understood by some of the first-line exponents of mathematical logic at the beginning of the 20th century. The treatment of the ordered pairs, indeed, is of great interests for us, since, the way in which set theorists encapsulated the operation for $(x, y)$, is of great difference from the strategies Frege and Russell-Whitehead applied. So, in order to have a more clear idea on the debate concerning ordered pairs, we will sketch the main ideas who were circulating at those time[32] Notice that, even if – from a philosophical point of view – it might be seen as "not necessary" to inquire the debate concerning ordered pairs, it actually is. The investigation on pairs will, indeed, require a previous investigation on the notion of *relation*, as understood by naïve set theorists (especially, Frege and Russell) and, more importantly, as refined by post-fregean mathematicians (starting from Peano). *Relations* have always played, and still play, a crucial role in mathematics and so, – if our aim is to philosophically understand maths –, then we have also to focus on what relations are and what are they for. Moreover, in «modern mathematics, the ordered pair is basic, of course, and is introduced early in the curriculum of study of analytic geometry. [...] the historical development proceeded on the mathematical side, in the work of Peano and Hausdorff. However, the development on the logical side, [has been buttressed] in the work of Frege and Russell [...]»[33].

**Frege on *relations*.** Frege's philosophical conceptions, recall, comprise two fundamental categories, the one of *function* and the other of *object*. In fregean terms, a function is an "unsaturated" entity, that is, something that has to be supplemented with something different. In this sense, objects, which are "saturated" entities, constitute the completion of functions, becoming their *arguments*. In such a context, as previously said, *concepts* are special cases of functions, i.e. a concept is a function that has two objects – the True and the False – as arguments. For instance, the concept "being a student in Venice" is something that has not to be confused with the objects that (truly or not) fall under its extension. Indeed, objects are considered saturated individuals, while concepts – defined as functions – receive their "saturation" just in case some object falls under them, allowing thus their mapping either to the True or to the False. In this sense, *relations* are special cases of concepts and, therefore, of

[31] Moschovakis 2006, p. 34.

[32] In what follows, we will give a very brief introduction to the above mentioned debate, and our reconstruction will chronologically consider just the major results of the years between 1880–1921. The exposition will end with the now commonly accepted solution, devoted to K. Kuratowsi. See, among others, Kanamori 2009, pp. 21–24. and, especially, Kanamori 2003, pp. 288–293.

[33] Kanamori 2003, p. 288.

functions too. In *Grundgestze*, we're told that a relation is a concept that takes two arguments and that its course-of-values is "double". Differently, a one-place relation, or simply, a concept, has as extension just its single course-of-values. As clear, hence, relations, for Frege, are to be assimilated to concepts, thus not belonging to the category of objects, being, henceforth, characterized as unsaturated individuals. With repsect to ordered pairs, then, Frege, equipped with this sharp distinctions, in §144 of his *Grundgesetze*, wrote:

> We always pair one member of the cardinal number series with one member of the $q$-series and form a series out of these pairs. The series-forming relation is determined thus: one pair stands in it to a second if the first member of the first pair stands in the $f$-relation to the first member of the second pair and the second member of the first pair stands in the $q$-relation to the second member of the second pair.[34]

In more common terms, it can be said that Frege defined the ordered pair of two objects as the class of all relations that hold between these objects, assuming that the notion of relation is primitive. Now, if $x$, $y$ represent objects and $R$ a relation, then Frege's definition can be formally rendered as follows:

$$(x, y) = \{R \mid xRy\}.$$

In *Grundgesetze* Frege introduced this notion for two main technical reasons. Firstly, he wanted to define the "coupling" of two relations and, secondly, he wanted to have a notion useful while treating the one-to-one correspondence between the natural numbers and the extensions of concepts. However, this definition is not admissible and, indeed, consider that:

> [u]nfortunately, Frege's definition of ordered pairs is, as George Boolos once put it, extravagant and can not be consistently reconstructed, either in second-order logic or in set-theory. According to Frege's definition, the ordered pair $(a; b)$ is the class to which all and only the extensions of relations in which $a$ stands to $b$ belong[35].

Anyway, fortunately, meanwhile Frege's logicist attempt, other mathematicians or logicians were working upon a different conception of relations and pairs. Consider, indeed, that – as it is nowadays usual practice – «[...] Peirce, Schröder and Peano essentially regarded a relation from the outset as just a collection of ordered pairs»[36].

**From the logic to the mathematics of pairs.** Consider that in «logic the ordered pair is fundamental to the logic of relations and epitomized, at least for Russell, metaphysical preoccupations with time and direction»[37]. Indeed, from a logical point of view, Frege's and Russell's treatment of relations considered

[34] Frege 1893/1903, p. 179.
[35] Heck 1995, p. 301.
[36] Kanamori 2009, p. 22.
[37] Kanamori 2003, p. 288.

functions as basic or primitive, by trying to define the concept of ordered pair just after having provided the notion of function. Differently, if the ordered pair is understood as a class, then the it is possible to start from it as fundamental and, just afterwords, accommodate the discussion concerning functions and relations. This latter strategy is exactly the one mathematicians, unlike Frege and Russell, applied to the mathematical study of relations and pairs:

> Given the formulation of relations and functions in modern set theory, one might presume that the analysis of the ordered pair emerged from the development of the logic of relations, and that relations and functions were analysed in terms of the ordered pair.[38]

Indeed, at the very beginning of set theory, logicians were still strictly concerned with metaphysical preoccupations, while mathematicians were trying to separate philosophy and logic from mathematics, by directing their approaches increasingly, and almost exclusively, towards extensional features. Indeed, next to Frege's *Begriffschrift*, during the last decades of the 19th century, it was proposed, by the Italian G. Peano, another very interesting formalized language for mathematics. This symbolic logic has been the one that Russell and Whitehead adopted to write their *Principia Mathematica* (1910-13) and that Russell alone applied in all of his formal-philosophical writings. Note, however, that whereas «Frege was attempting an analysis of thought, Peano was mainly concerned about recasting ongoing mathematics in economical and flexible symbolism and made many reductions, e.g., construing a *sequence* in analysis as a *function* on the natural numbers»[39]. In such a context, Peano, indeed, by using his notation, has been the first in stating the fundamental properties of ordered pairs. His analysis highlighted, not only the fact that the concept of ordered pair is of great importance, but also and especially that, if a couple has to be determined as an ordered pair, then condition **C1**[40] must hold:

$$(x, y) = (z, u) \text{ iff } x = z \text{ and } y = u.$$

Indeed, it has been exactly B. Russell who – in different parts of his writings – took position against Peano's suggestion, among others, to consider the ordered couple as a class. For instance, Russell's criticism – as expressed in the *Principles of mathematics*, with respect to Schröder's and Peirce's logics – points out that:

> their method suffers technically (whether philosophically or not I do not at present discuss) from the fact that they regard a relation essentially as a class of couple, thus requiring elaborate formulae of summation for dealing with single relations. This view is derived, I think, probably unconsciously, from a philosophical error: it has always been customary to suppose relational propositions less ultimate than class

[38] Kanamori 2003, p. 288.

[39] Kanamori 2003, p. 289.

[40] See beginning of this section, i.e., "Ordered Pairs: From Frege to Russell", for some further specifications on condition **C1**.

> propositions (or subject-predicate propositions, with which class-propositions are habitually confounded), and this has led to a desire to treat relations as a kind of classes. However this may be, it was certainly from the opposite philosophical belief, [...] that I was led to a different formal treatment of relations. This treatment, whether more philosophically correct or not, is certainly far more convenient and far more powerful as an engine of discovery in actual mathematics.[41]

Aside Russell's criticism and philosophical tastes, contemporary set theory has taken the opposite path and, as remarked by Kanamori, the pioneer researches of Schröder, Peirce and Peano, has led to the idea of intending couples of ordered elements as sets. Russell and Whitehead, whose main intent was that of reducing all branches of mathematics to their logical type theory, buttressed the *Principles* idea of deriving – from other fundamental notions, such as the one of relation – the concept of ordered pair. Indeed, in their *Principia Mathematica*, Russell and Whitehead adopted the following definitional strategy: by defining the Cartesian product between elements, they defined the ordered pair as follows[42]:

$$x \downarrow y =_{\text{def}} \{x\} \times \{y\}.$$

As clear, their work has been contrary to that of contemporary set theorists, who have to define the ordered pair to obtain the desired notion of Cartesian product and not *viceversa.* Indeed, Peano, in his *Formulario Mathematico* (1905-8), aware of the conception developed by Russell and Whitehead, «simply reaffirmed an ordered pair as basic, defined a relation as a class of ordered pairs, and a function extensionally as a kind of relation»[43]. For example, already in 1901, when Russell sent to Peano's review his article, titled *Sur la logique de relations avec des applications à la théorie des séries*, the Italian mathematician continued in pursuing his objective of understanding relations as ordered couples. Indeed, in the same years Peano was arguing that the «idea of a pair is fundamental, [but], we do not know how to express it using the preceding symbols», Norbert Wiener elaborated the first mathematical successfully strategy of characterization of ordered pairs. More specifically, his work provided us with a «definition of the ordered pair in terms of unordered pairs of classes only, thereby reducing relations to classes»[44], and accomplishing, thus, the inversion of trend, which contemporary set theory has inherited. Wiener's work has been conducted within Russell's type theory and, – apart defining the notion of ordered pair – as result, he made it possible to define "types" as sets. His definition can be given as follows:

$$(x, y) =_{\text{def}} \{\{\{x\}, \varnothing\}, \{\{y\}\}\}.$$

[41]Russell 1903, p. 24.

[42]We follow Russell's and Whitehead's notation. The "down arrow", put between two elements, symbolized that the two elements constituted an ordered pair.

[43]Kanamori 2003, p. 290.

[44]Kanamori 2003, p. 290.

Exactly the fact that Wiener was working in type theory, justifies the fact that $\{\{y\}\}$ is inserted in place of simply $\{y\}$: in *Principia*, Russell and Whitehead had asserted that every element of a class must be of the same type of any other element of the same class. Almost at the same time, another mathematician – Felix Hausdorff – arrived to a very similar solution to Wiener's:

$$(x, y) =_{\text{def}} \{\{x, \mathbf{1}\}, \{y, \mathbf{2}\}\},$$

where **1** and **2** are two distinct objects, different from $x$ and $y$ and, moreover, alien to the situation. He, thus, defined the ordered pair in terms of unordered pairs, formulated functions in terms of ordered pairs, and ordering relations as collections of ordered pairs, and, in doing so, he «made both, the Peano and Wiener moves *in* mathematical practice, completing the reduction of functions to sets»[45]. Indeed, his definition verifies the condition Peano first formulate, namely **C1**. Anyway, before arriving to the now widely accepted solution, it is worthy to point out that, Russell had rejected Wiener's and Hausdorff's analysis, and that he continued in thinking of relations as primitive and, especially, as intensional entities:

> After his break with neo-Hegelian idealism, Russell insisted in taking relations to have genuine metaphysical reality, external to the mind yet intensional in character. On this view, order had to have a primordial reality, and this was part and parcel of the metaphysical force of intension[46].

As remarked above, mathematicians – especially thanks to Hausdorff's work and unlike Russell – pursued the extensional understanding of relations and, finally, in 1921, the logician Kazimierz Kuratowski exposed his solution, that is, the result – concerning pairs – he obtained by analysing Zermelo's Well-ordering theorem[47]. Kuratowski's ordered pair is defined as follows:

$$(x, y) =_{\text{def}} \{\{x\}, \{x, y\}\}.$$

Two fundamental facts concerning the previous definition can be immediately seen. Both consequences will, as expected, shed light on the definitional adequateness of Kuratowski's pairs. First, consider a couple where the first and the second coordinate are identical, i.e. $(x, x)$. According to Kuratowski's definition, therefore, we can write:

$$(x, x) =_{\text{def}} \{\{x\}, \{x, x\}\},$$

By inspection, $\{\{x\}, \{x, x\}\}$ is equivalent to $\{\{x\}, \{x\}\}$ and, hence, to $\{\{x\}\}$. This simply means that Kuratowski's definition is adequate to define an ordered couple even if the elements it contains are identical. More importantly, now, we can easily prove that Kuratowski's definition satisfies **C1**. and **C2**., being thus the desired notion of coupling for our set-theoretical framework.

[45]Kanamori 2003, p. 291.
[46]Kanamori 2003, p. 290.
[47]The next section is devoted to this main result of set theory. See Zermelo 1908a.

**Lemma 1.** The Kuratowski pair operation

$$(x, y) =_{\text{def}} \{\{x\}, \{x, y\}\},$$

satisfies condition **C1**, i.e., formally,

$$(x, y) = (z, u) \longleftrightarrow x = z \ \wedge \ y = u.$$

*Proof.* ($\leftarrow$) We have $x = z$ and $y = u$. By Kuratowski's definition we have that $\{\{x\}, \{x, y\}\} = \{\{z\}, \{z, u\}\}$. But, this means $(x, y) = (z, u)$.

($\rightarrow$) Assume $(x, y) = (z, u)$. We have to distinguish two cases:

**1**. $x = y$

**2**. $x \neq y$.

**1**. If $x = y$:

Consider $(x, y) = \{\{x\}, \{x, y\}\} = \{\{x\}, \{x, x\}\} = \{\{x\}\}$. Since the set $(z, u)$ is assumed to be equal to $(x, y)$, we have $z = u$ and $(z, u) = \{\{z\}\}$. Moreover, since by $\{\{x\}\} = \{\{z\}\}$ we have that $x = z$, consequently, we obtain $y = x = z = u$.

**2**. If $x \neq y$:

The members of $(x, y)$ in this case are $\{x\}$ and $\{x, y\}$. By assumption they must be equal to those of $(z, u)$, namely to $\{z\}$ and $\{z, u\}$. In this sense, from $\{x\} = \{z\}$ and $\{x, y\} = \{z, u\}$, we immediately get $x = z$ and $y = u$.

■

So, this completes our research towards an appropriate set theoretic notion of ordered couple. As sketched, the research of this characterization has had different difficulties, in particular, arisen from the mistaken assumption of relations as primitive and intensional entities. By developing a pure extensional notion of couple, the mission of reducing relations and, consequently, functions to sets can be said accomplished by the leading work of Kuratowski. Notably, the

> general adoption of the Kuratowski pair proceeded through the developments of mathematical logic: von Neumann initially took the ordered pair as primitive but later noted the reduction via the Kuratowski reduction. Gödel in his incompleteness papers also pointed out the reduction. Tarski, seminal for its precise, set-theoretic formulation of first-order definable set of reals, pointed out the reduction and acknowledged his compatriot Kuratowski. In his recasting of von Neumann's system, Bernays also acknowledged Kuratowski and began with its definition for the ordered pair. It is remarkable that Nicolas Bourbaki in his treatise (1954) on set theory still took the ordered pair as primitive, only later providing the Kuratowski reduction [...][48]

[48]Kanamori 2003, p. 292.

### 3.2.4 Zermelian Considerations on the Axioms

Before publishing the whole collection of axioms, Zermelo published two different, but related, works. As said before, in 1904, Zermelo proved the fundamental claim that any set can be well-ordered. In the paper where he presents his brief proof, much of set theory is presupposed and no comments, for what concerns the usage of the controversial axiom of choice, are offered. Differently, some years later, and precisely in 1908, Zermelo published another proof of the well-ordering theorem, this time, by using less set-theoretical presuppositions and by adding, at the end of the paper, a careful analysis of the criticism deserved to his axiom of choice. Finally – always in the year 1908 – Zermelo accomplished the mission of giving an explicitly formulated axiomatization of set theory. This happened exactly by means of the axioms we've just given, i.e., **I.-VII.**. The history of set theory, anyway, does not terminate here and other works merit to be mentioned. Indeed, for our purposes, we have to arrive to the last non Zermelian axiom, that is the so-called "Axiom of Replacement" as announced by A. Fraenkel and T. Skolem around 1922[49], and as axiomatized, in the late 30s, by J. von Neumann. Indeed, in order to arrive to the addition of this new axiom, let's consider both, what our Zermelian framework is able to do and which tasks it cannot absolve. Additionally, the philosophical import of Zermelo's axioms should be analysed at length, but – since it would take us away from our main investigative path –, we restrict our considerations to two fundamental points: (i) the axiom of choice and its Zermelian defence and (ii) the axiom of infinity and its applicability within number theory.

#### 3.2.4.1 Well-Orderings, Choices and Axiomatic Method

Zermelo worked in several European universities, including Göttingen, where he arrived at the beginning of the century. There, moved by the great mathematician D. Hilbert, Zermelo began his work in set theory, trying to do justice to Cantor's and Dedekind's intuitions. His earliest contribution, as said, concerns the well-ordering theorem (1904)[50]. The background of such a proof was that given always by Cantor and its naïve understanding of sets. In particular, within a letter sent to Dedekind, Cantor explained what he meant with the concept of "well-ordering" for sets and, moreover, he conjectured the existence of a proof of the well-ordering of any set. Since Zermelo aimed to find a proof exactly of the well-ordering conjecture, let's read Cantor's own words:

> A multiplicity is said to be *well-ordered* if it satisfies the the condition that every *submultiplicity* of it has a *first* element; I call such a multiplicity "sequence" for short[51].

Apart having this informal definition, it would be better to possess a more conventional representation of what is meant for a set to be ordered and, moreover,

[49]See, in particular, Skolem 1922, pp. 290–301 and Fraenkel 1922, pp. 284–289. For clear commentaries see Kanamori 2009, pp. 30–31.

[50]See, in particular, Zermelo 1904, 1908a,b. For technical and historical commentaries see Moschovakis 2006, pp. 21–28 and, especially, Kanamori 2009, pp. 11–16.

[51]Cantor 1899, p. 114.

well-ordered. So, starting from Cantor's intuition, contemporary set theorists furnish the following definitions:

**Definition 18** (Ordering of a Set)**.** A binary relation $\leq$ on a set $A$ is a **partial ordering** if the following conditions hold:

1. **Reflexivity**: $x \leq x$;
2. **Antisymmetry**: $x \leq y \ \wedge \ y \leq x \longrightarrow x = y$;
3. **Transitivity**: $x \leq y \ \wedge \ y \leq z \longrightarrow x \leq z$.

We define the binary relation $<$ as follows:

$$x < y =_{\text{def}} \ x \leq y \ \wedge \ x \neq y.$$

The partial order $\leq$ is **total** or **linear**, if, additionally to clauses 1-3, any two elements of $A$ are **comparable** with respect to $\leq$. Formally,

$$(\forall x, y \in A) \ [x < y \ \vee \ x = y \ \vee \ x \leq y]$$

Now, with this preliminary definition of what does it mean for a set being ordered, we may state the definition of "well-ordered" set logically:

**Definition 19** (Well-Ordering of a Set)**.** The binary relation $\leq$ on $A$ is a **well-ordering** of $A$, if the following two conditions hold:

(i) $\leq$ is a total order;

(ii) every non-empty subset of $A$ has a **least element**, i.e.:

$$(\forall C \subseteq A) \ \Big[C \neq \varnothing \longrightarrow (\exists x \in C)(\forall y \in C)[x \leq y]\Big].$$

Now, consider that $\leq$ can be also defined on the set of natural numbers as the relation "less or equal to", between natural numbers. But, to do this job we have to prove the existence of such a set – namely of $\mathbb{N}$ –, and this, the required proof, will be given in the next section devoted to the axiom of infinity.
As remarked in the quote, Cantor's work, among other things, distinguished between "consistent" and "inconsistent" multiplicities. The latter ones, by admitting a contradictory and logically inadmissible notion, let mathematicians seek for another proof, which does not rely on the introduction of "inconsistent" sets. The proof, presented by Zermelo, differently, is brief and relies on the axiom of choice – a rule that Cantor, for instance, had used implicitly and instinctively in order to gain the possibility of a *simultaneous* choice. Differently, equipped with his axiom of choice **VI.**, Zermelo «shifted the notion of set away from the implicit assumption of Cantor's principle that every well-defined set is well-ordered and replaced it by an explicit axiom about a wider notion of set, incipiently unstructured but soon to be given

form by axioms»[52]. Zermelo's proof has been, indeed, pivotal to the development of the axiomatic version of set theory and it is useful to summarize his main proof-theoretical strategy. Arrived at this point, it is worthy to notice that, while Zermelo has been very careful in mathematizing naïve set theory, he almost did not care on the logical features which his mathematization was involving. In fact, consider that Zermelo, while trying to clarify the terms "relation" and "definite relation", invokes the generality and universal validity of the laws of logic. The difference between a simple relation and a definite one is that, in presence of the axiom of separation, the latter "separates" a subset from an already given set, while the first ones no. This characterization is, somehow, too mysterious for mathematicians and it has been argued that Zermelo did not pay «attention at all to the underlying logic, [the] laws are left unspecified, and the notion of definite property remains hazy»[53]. So, in order to proceed towards the well-ordering theorem, let's state what is for a relation to be "definite". We've specified, already at the beginning of the construction of $\mathcal{Z}$, that **identity**, **sethood** and **membership** are definite conditions[54]. Furthermore, let's establish that:

1. For any object $a$ and each $n$, the constant $n$-ary operation $F$
$$F(x_1, \ldots, x_n) = a$$
is definite.
2. Any projection operation,
$$F(x_1, \ldots, x_n) = x_i \quad (1 \leq i \leq n)$$
is definite.
3. If $P$ is a definite condition for $n+1$ arguments and for each $x_1, \ldots, x_n$ there is exactly one $u$, such that $P(x_1, \ldots, x_n, u)$ is true, then the formula
$$F(x_1, \ldots, x_n) = \text{the unique } u \text{ such that } P(x_1, \ldots, x_n, u)$$
is definite.
4. If $S$ is an $m$-ary definite condition, any $F_i$ is an $n$-ary definite formula for $i = 1, \ldots, m$ and
$$P(x_1, \ldots, x_n) \Longleftrightarrow S(F_1(x_1), \ldots, F_m(x_n)),$$

[52] Kanamori 2009, p. 1.

[53] In the *Introduction* to Zermelo 1908b, p. 199. Further, note that we are assuming that set tehory has always be done by taking a first-order language plus a sole non-logical constant, i.e., $\in$, but, consider that the pioneer work of "assembling" the language of set theory within that of first-order logic has been a success of – another time – A. Fraenkel and T. Skolem. See, indeed, Fraenkel 1922 and Skolem 1922.

[54] Consider that Skolem's and Fraenkel's considerations on the lack of a precise notion of "definite condition" in Zermelo's paper were presented in the early 20s. Zermelo himself, along his 1930 article, provided more precise characterizations of definiteness. In what follows, anyway, we will deal with the more usual and conventional practice.

then $P$ is also definite.

5. Let $\overline{x} =_{\text{def}} x_1, \ldots, x_n$. If $Q, R, S$ are definite conditions on a number of arguments, then so are the following conditions obtained by simple logical manipulations:

$$P(\overline{x}) \iff \neg P(\overline{x}) \iff P(\overline{x}) \text{ is false.}$$

$$P(\overline{x}) \iff Q(\overline{x}) \wedge S(\overline{x}) \iff \text{ both, } Q(\overline{x}) \text{ and } S(\overline{x}), \text{ are true.}$$

$$P(\overline{x}) \iff Q(\overline{x}) \vee S(\overline{x}) \iff \text{ either } Q(\overline{x}) \text{ or } S(\overline{x}) \text{ is true.}$$

$$P(\overline{x}) \iff Q(\overline{x}) \to S(\overline{x}) \iff \text{ if } Q(\overline{x}) \text{ is true, then } S(\overline{x}) \text{ is true.}$$

$$P(\overline{x}) \iff \exists u(S[(\overline{x}), u]) \iff \text{ for some } u, \; S[(\overline{x}), u] \text{ is true.}$$

$$P(\overline{x}) \iff \forall u(S[(\overline{x}), u]) \iff \text{ for any } u, \; S[(\overline{x}), u] \text{ is true.}$$

That's it for what concerns the definition of what Zermelo had in mind with "definiteness" for conditions and relations. With this in our heads, let's state the proof, of Cantor's conjecture, of well-ordered sets:

**Theorem 7** (Zermelo's Well-Ordering Theorem)**.** Every set can be well-ordered.

*Proof Sketch.* Suppose that $b$ is a set to be well-ordered. Assume, by the axiom of choice, **VI.**, that $\wp(b) = \{z \mid z \subseteq b\}$ has a choice function $\gamma$, such that for every non-empty element $z$ of $\wp(b)$, it holds $\gamma(z) \in z$. We say that $z$ is a $\gamma$-set if there is a well-ordering $R$ of $z$ such that for any $c \in z$ it holds that:

$$\gamma\Big(\{x \mid x \notin z \vee \neg R(z, c)\}\Big) = c$$

Hence, each member of $z$ is what $\gamma$ *chooses* from what does not already precede that member according to $R$. ■

Nice results immediately follow from Zermelo's proof and, with this theorem at hand, already at the end of his 1904, Zermelo noticed Cantor's conjecture concerning "cardinal comparability" could be finally secured. Even if also this latter point should be treated at length, it would take us away from our main purposes. What we actually are most interested in is the fact, that, already in 1904, Zermelo came up near to what set theorists would have called – after von Neumman's accurate explications – "Recursion theorem" and "Axiom of Replacement". Anyway, Zemelo did not formulate completely neither the theorem nor the axiom, since his main objective while axiomatizing Cantor's intuitions, «was not for the formulation and solution of a *problem* like the Continuum Problem, but rather to clarify a specific *proof*»[55], namely the demonstration of the well-ordering conjecture. In this specific context, the work of Dedekind – we've sketched in the foregoing chapter – has played a crucial rule for Zermelo's work within the axiomatization of sets (at least) in two senses. Firstly, in the proof the mathematician presented in 1908, titled *A new proof of the well-ordering*, he explicitly recognises the merits of Dedekinds *chains*, by applying them into his reformulated proof. Secondly:

> By Dedekind's time proof had become basic for mathematics, and, indeed, his work did a great deal to enshrine proof as the vehicle for algebraic abstraction and generalization. Like algebraic constructs, sets were new to mathematics and would be incorporated by setting down the rules for their proofs. Just as "calculations" are part of the sense of numbers, so proofs would be part of the sense of sets, as their "calculations". Just as Euclid's axioms for geometry had set out the permissible geometric constructions, the axioms of set theory would set out the specific rules for set generation and manipulation. [...] For Dedekind it had sufficed to work with sets by merely giving few definitions and properties, those foreshadowing Extensionality, Union, and Infinity. Zermelo provided more rules: Separation, Power Set, and Choice[56].

So, like Dedekind[57], also Zermelo thought that *proofs* and rigorous definitions for the construction and characterization of mathematical objects was a fertile way to do mathematics. Indeed, for instance, with respect to its axiom of choice Zermelo wrote:

> This logical principle cannot, to be sure, be reduced to a still simpler one, but it is applied without hesitation everywhere in mathematical deduction. For example, the validity of the proposition that the number of parts into which a set decomposes is less than or equal to the number of all of its elements cannot be proved except by associating with each of the parts in question one of its elements.[58]

Indeed, some couples of years later, Zermelo defended again his axiom of choice from the criticism mathematicians and logicians deserved to it:

> [...] the question that can be objectively decided, whether the principle is *necessary for science*, I should now like to submit to judgement a number of elementary theorems and problems that, in my opinion, could not be dealt with at all without the principle of choice. [...] Cantor's theory of cardinalities [...] certainly requires our postulate,

[55]Kanamori 2009, p. 13.

[56]Kanamori 2009, pp. 14–15.

[57]We've already explained Dedekind's conception of mathematics in the previous chapter. For the sake of the argument, anyway, that Dedekind himself began his monograph on natural number by asserting:«In science nothing capable of proof ought to be accepted without proof» Dedekind 1888b, p. 790.

[58]Zermelo 1904, p. 141.

> and so does Dedekind's theory of sets which forms the foundations of arithmetic[59].

Hence, for Zermelo, the axiom of choice, or better, axioms for in general, are to be understood as fundamental, when their applications result necessaries for the development of mathematics and science. Indeed, while some may agree that the choice principle, for instance, is not intuitive as the other axioms, and consequently try to argue that it should be rejected, they do not consider at all, according to Zermelo's perspective, its fruitfulness in applications and consequences. Anyway:

> [...] as long as [...], the principle of choice cannot be definitely refuted, no one has the right to prevent the representatives of productive science from continuing to use this "hypothesis" – as one may call it for all I care – and developing its consequences to the greatest extent, especially since any possible contradiction inherent in a given point of view can be discovered only in that way[60].

So, if we do not possess a refutation of inherent contradiction of some principle and its usefulness is evident in the sense that, nice results follow from its application, according to Zermelo, we should be convinced that its truth is, in this precise sense, *self-evident*:

> [...] That this axiom, even though it was never formulated in textbook styles, has been frequently used, and successfully at that, in the most diverse fields of mathematics, especially in set theory, by Dedekind, Cantor, F. Bernstein, Schoenflies, J. König, and others is an indisputable fact [...] Such an extensive use of a principle can be explained only by its *self-evidence*, which, of course, must not be confused with its provability.[61]

So, as remarked by Zermelo himself, mathematics should not be done without proofs and rigorous definitions, and, indeed, in order to give a clear and precise systematization of Cantor's and Dedekind's intuitions, he provided, not only proofs of some of the major claims and conjectures, but he furnishes us, especially, with precise and explicitly formulated rules for the construction and manipulation of sets. Moreover, he was attracted – as especially Dedekind has been some years before him – by «the reduction of mathematical concepts and arguments to set-theoretic concepts and arguments from axioms, based on sets doing the work of mathematical objects». Linked to this, as it should be now clear, Zermelo agreed with the general convictions concerning the axiomatic method and, *a posteriori*, his name can be well signed – among those of Hilbert and Russell – within the defender of the necessity of the axiomatic methodology for the development of mathematics.
Having a bit clarified Zermelo's conception and use of the axioms for sets, we can introduce how he thought to apply (and, consequently, justify) his other controversial axiom, namely the axiom of infinity. In the context of that analysis, finally, – we anticipate – something as the natural numbers sequence (as thought by Zermelo)

[59]Zermelo 1908a, pp. 187–189.
[60]Zermelo 1908a, p. 189.
[61]Zermelo 1908a, p. 187.

will appear.

#### 3.2.4.2 Infinite Sets, Natural Numbers and Axioms

Recall that one of the main important features of Dedekind's work has that of (trying) to prove – after having provided a definition of infinity– the claim for which infinite sets exist. Differently, set theorists, during and after Zermelo's pioneer work, adopted an axiom to assert the existence of such sets and, as in the case of the axiom of choice, some criticism has emerged. Roughly, it can be argued that being convinced that the existence of infinite sets, as asserted by the axiom of infinity, is much demanding – philosophically speaking. Now, we've seen that one way to defend the misunderstandings coming from the axioms of our mathematical theories is to appeal to their *self-evidence* in the Zermelian sense, that is by considering their usefulness and consequences. We will come to one of the most fundamental proofs for both, set and number theory, which requires the application of the axiom of infinity and which will shed more light on Zermelo's idea of *self-evidence* soon. For the moment, we will sketch another important way of defending the presence of some particular or "extravagant" axioms within mathematical theories. This defence strategy is not directly to be found in Zermelo's work, but it is one the most appealing consequences of the whole of his work in set theory. For the sake of the argument, consider again the axiom of infinity and that it's immediate reading seems to commit us immediately to the existence of, at least, one infinite set. But, actually, does the axiom of infinity really commit us to the existence of such a set? For some set theorists, Zermelo's axiom does not have this kind of commitment and, indeed, it is generally accepted that, one of the major contributions, lying along the entire Zeremelian work, is that of purging mathematics, ans especially set theory, from philosophical and existential doubts:

> The intuitive understanding of the axiom [of infinity] is that it demands precisely the existence of the set
>
> $$\mathbb{I} = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \dots\},$$
>
> but it is simpler (and sufficient) to assume of $\mathbb{I}$ only the stated properties, which imply that it contains all these complex singletons.
>
> It was a commonplace belief that among philosophers and mathematicians of the 19th century that the existence of infinite sets could be proved, and in particular the set of natural numbers could be "constructed" out of thin air, "by logic alone". All the proposed "proofs" involved the faulty General Comprehension Principle in some form or another. We know better now: logic can codify the valid forms of reasoning but it cannot prove the existence of anything, let alone infinite sets.[62]

But, notice, that this does not mean or imply that logical and philosophical concerns about set theory should be completely rejected, but, rather, that the philosophy of set theory should be separated from the mathematics of sets. In other words,

[62] Moschovakis 2006, p. 30.

while elaborating mathematical theories, no ontological concerns must be considered by mathematicians, since this kind of problems is exactly what a philosopher of mathematics is called to investigate and to answer. Indeed,

> By taking account of this fact[s] cleanly and explicitly in the formulations of his axioms, Zermelo made a substantial contribution to the process of purging logic of ontological concerns and separating the mathematical development of the theory of sets from logic, to the benefit of both disciplines.[63]

Thus, Zermelo's work can be seen as a step further in the development of mathematics, as an always more autonomous discipline, thanks to his work in set theory: Cantor's collections, Dedekind's systems and, especially, Frege's extensions were born as sons of both, mathematics and philosophy, but now, differently, Zermelo's sets are just the product of mathematics. In any case, as specified, this feature does not preclude the possibility of the philosophical investigation of sets or, more importantly, their fruitful application out of the range of pure mathematics.
Having, thus, seen how the axiom of infinity has been justified on the basis of a division – not explicitly stated by Zermelo, but directly connected to his work – between the ontology and the mathematics of sets, we are now going to analyse the sense in which Zermelo himslef thought the axiom of infinity being *self-evident*. Recall, another time, that – resounding Hilbert's and Russell's position – an axiom is *self-evident*, according to Zermelo, when nice conclusions, concerning a specific field of mathematics, can be proved while involving the axiom itself. So, to see in which sense the postulation of infinite sets is necessary for the development of mathematics, reconsider the natural numbers sequence $0, 1, 2, 3, \ldots$ and the fundamental characters the Dedekind-Peano axioms ascribe to it. The progression is usually structured as follows: we begin with a least number, 0, and we continue by characterizing any member $n$ of the progression by applying the operation of "successor of" to 0 as many times as necessary, i.e.:

$$0,\ \underbrace{S(0)}_{1},\ \underbrace{SS(0)}_{2},\ \underbrace{SSS(0)}_{3},\ldots,\ \underbrace{S\ldots S(0)}_{n},\ldots$$

As said several times, for instance, Frege, among others, tried to reduce the truths concerning the natural numbers to "more basic" truths, namely to those concerning the extensions. The theory of collections he developed was inherently inconsistent and, therefore, as remarked, the reduction he had planned – i.e., the derivation of the Dedekind-Peano axioms for number theory from the theory of extensions, as developed in the *Grundgesetze* – failed. But now, having a more mathematized and rich framework for sets that escapes the raise of several inconsistencies, it is possible to characterize set-theoretically the axiomatization of Dedekind and Peano. Let's start this task.

**Definition 20.** A **Peano system** or **system of natural numbers** is any ordered

[63] Moschovakis 2006, p. 30.

triple:

$$(\mathbb{N}, 0, S)$$

which satisfies the following conditions:

1. $0 \in \mathbb{N}$, i.e. the set $\mathbb{N}$ contains 0 as element;
2. $S : \mathbb{N} \to \mathbb{N}$, such that $n \mapsto S(n)$. That is, $S$ is a function defined on $\mathbb{N}$, in a way such that each $n \in \mathbb{N}$ is mapped to its successor, $S(n)$;
3. $S$ is an injective function, i.e. $S(m) = S(n) \to m = n$;
4. Any successor number is different from 0, namely, for each $n \in \mathbb{N}$, $S(n) \neq 0$;
5. **Induction Principle**. For any subset $A$ of $\mathbb{N}$, if 0 is contained in $A$, and, if a number is in $A$, then its successor is in $A$, then $A$ is identical to $\mathbb{N}$. Formally,

$$\forall A \subseteq \mathbb{N} \left[0 \in A \ \wedge \ (\forall n \in \mathbb{N})[n \in A \to S(n) \in A]\right] \to A = \mathbb{N}.$$

It is not difficult to see the similarities between these set-theoretical formulation of the Dedekind-Peano axioms and the way we've expressed them in the previous chapter. What we actually have to notice is that our main interests are not historical or exegetical and, therefore, it is better to point out that we are simply working in a Zermelian-style framework and, indeed, the definitions and the proofs we will give are stated in a more conventional and contemporary fashion. With this in mind, notice that our definition invokes a notion with which we are not so familiar yet, i.e. the notion of "ordered triple". The idea behind ordered triples can, roughly, be given as follows:

**Definition 21.** The **ordered triple** $(x, y, z)$ can be defined either as the ordered pair:

$$((x, y), z)$$

or as the ordered pair:

$$(x, (y, z)),$$

where $(x, y)$ and $(y, z)$ are themselves ordered pairs.

So, returning to our definition, a doubt may arise: how do we assure that every element of the $n$-sequence is a number different from any other? In more profane words, what does it grant us that, at a certain point, we will not encounter a number $n$ equal to its successor $S(n)$? The answer's clearly "by a proof". Let's clarify our doubt by proving the next lemma and by paying attention to the application of the set-theoretical version of the induction principle:

**Lemma 8.** In a Peano system $(\mathbb{N}, 0, S)$, any $n \neq 0$ is a successor. Formally:

$$\forall n \in \mathbb{N} \left[ n \neq 0 \longrightarrow (\exists m \in \mathbb{N})[n = S(m)] \right].$$

Moreover, each number is different from its successor, namely:

$$\forall n \in \mathbb{N} \; [S(n) \neq n].$$

*Proof.* Consider first that we have to prove that,

$$A = \{n \in \mathbb{N} \mid n \neq 0 \longrightarrow (\exists m \in \mathbb{N})[n = S(m)]\}, \tag{3.1}$$

which is logically equivalent to:

$$A = \{n \in \mathbb{N} \mid n = 0 \vee (\exists m \in \mathbb{N})[n = S(m)]\},$$

satisfies the following conditions:

(i) $0 \in A$;

(ii) $(\forall n \in \mathbb{N})[n \in A \longrightarrow S(n) \in A]$

Both – (i) and (ii) – immediately follows from the definition of $A$.
Secondly we prove that:

$$\forall n \in \mathbb{N} \; [S(n) \neq n]$$

satisfies the following conditions:

(i) $S0 \neq 0$;

(ii) $S(n) \neq n \longrightarrow SS(n) \neq S(n)$.

For (i) consider that, by axiom 4, for any $n$, $S(n) \neq 0$ and, hence, that also $S(0) \neq 0$. For the second point, consider that $S$ is an injection and, therefore, (ii) holds since $SS(n) = S(n) \longrightarrow S(n) = n$. ■

So, as it emerges from the proof, the set-theoretical version of the Induction principle has the fundamental feature that it concedes us to demonstrate that if some conditions are satisfied from a generic set $A$, namely that it contains 0, $n$ and its successor $S(n)$, then $A$ gets identified with set $\mathbb{N}$. In other words, we are able to prove – and that's the intuitive meaning of inductive proofs – determinate characters of the natural numbers along the entire sequence to which they belong. Indeed, without specifying any number-sign, we've been able to see that – no matter how long the sequence is – any number $n$ is different from any other and, additionally, that each $n$ is a successor of another number $m$, starting by 0.
So, with all these background, we can finally meet our natural numbers. Recall that, «[i]f number theory can be developed from the Peano axioms, then to give a faithful representation of the natural numbers in set theory, it is enough to prove from the axioms»[64] that, (i) natural numbers exist, and, (ii) that their existence is uniquely determined.

[64]Moschovakis 2006, p. 52.

**Theorem 9** (Existence of the natural numbers)**.** There exists at least one Peano system $(\mathbb{N}, 0, S)$.

This theorem relies on a very simple and elegant proof that involves the usage of the axiom of infinity. Let's see how it works.

*Proof.* With our axiom of infinity we have the warranty that there exists a set $\mathbb{I}$ which satisfies the following conditions:

(i) $\varnothing \in \mathbb{I}$

(ii) $\forall n\ [n \in \mathbb{I} \longrightarrow \{n\} \in \mathbb{I}]$

Now, using our $\mathbb{I}$ we define a set:

$$I = \{A \subseteq \mathbb{I} \mid \varnothing \in A\ \wedge\ \forall n\ (n \in A \longrightarrow \{n\} \in A)\}.$$

As obvious $\mathbb{I} \in I$. Now, let,

(i) $\mathbb{N} = \bigcap(I)$;

(ii) $0 = \varnothing$;

(iii) $S = \{(n, m) \in \mathbb{N} \times \mathbb{N} \mid m = \{n\}\}$

What we have to show is that the triple $(\mathbb{N}, 0, S)$ is a Peano system for arithmetic. First, consider that if some set $A \in I$, then it follows that $\varnothing \in A$. This implies that $\mathbb{N} \in I$. Additionally, for $\varnothing \in \bigcap(I) = \mathbb{N}$. Now, suppose that $n \in \mathbb{N}$. We know that $\mathbb{N} \in I$ and, so, that it holds that $\{n\} \in \mathbb{N}$. This implies directly the first two Peano axioms.
The third axiom holds because: since $\{n\} \neq n$, if $n \neq m$, then $\{n\} \neq \{m\}$. By stating the contrapositive, axiom 3. arises, i.e. $\forall n, m\ (\{n\} = \{m\} \longrightarrow n = m)$. Axiom 4 is obvious since, for all $n$ it holds that $\{n\} \neq \varnothing$.
Last point. We've defined $\mathbb{N}$ as an intersection, namely $\bigcap(I)$. Consider that this means that:

$$\bigcap_{n=0}^{\infty}(I_n) = I_0 \cap I_1 \cap \cdots = \{x \mid x \in I_n\} = \mathbb{N}.$$

That is, by definition of $I$, $\varnothing \in \bigcap(I)$ and for any $x$, $x \in \bigcap(I) \longrightarrow \{x\} \in \bigcap(I)$. Hence, by letting $\mathbb{N}$ be the intersection of $I$, we've finally confirmed Peano's last axiom, i.e. the Induction principle. ■

As clear, therefore, set-theoretical axioms has allowed us to prove the existence of a Peano system. Moreover it would be nice to have the following result:

**Theorem 10** (Uniqueness of the natural numbers)**.** For any two Peano systems $(\mathbb{N}_1, 0_1, S_1)$ and $(\mathbb{N}_2, 0_2, S_2)$ there exists exactly one bijective function,

$$\phi : \mathbb{N}_1 \to \mathbb{N}_2,$$

such that the following conditions hold:

$$\phi(0_1) = 0_2$$
$$\phi\Big(S_1(n)\Big) = S_2\Big(\phi(n)\Big).$$

■

Some remarks are, at this point needed. First, we will not prove the uniqueness theorem – other technicalities would be required, such as some versions of the Recursion theorem -, since, for what concerns our philosophical interest, it is sufficient. Indeed, we will describe soon how Zermelo's sets faithfully represent the $n$-sequence and, then, roughly turn to von Neumann's work.
Consider the whole of the work we've done in this section devoted to the axiom of infinity and to its application to natural numbers.

1. Thanks to our two above theorems (Existence and Uniqueness) we may now fix a specific Peano system $(\mathbb{N}, 0, S)$ and let its members be called the natural numbers.
2. We can, at this point, consider that the set, we constructed in the Existence proof, gets identified with $\mathbb{N}$. In this sense, the set that we have obtained from the application of the axiom of infinity is the following:

   $$\mathbb{N} = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \dots\}$$

   where $0 =_{\text{def}} \varnothing$ and $S(n) =_{\text{def}} \{n\}$.
3. As it should already be clear, one of the set-theoretical representation of natural numbers – Benacerraf attacked – has been now axiomatically set up. We have axioms which guarantee the existence of sets, in particular, of infinite sets, and, moreover, we have a theorem which tells us that one of those infinite sets may well be used to represent the natural numbers progression.
4. The notion of finiteness and infinity, with the axiomatic notions at our disposal, emerge again.

**Definition 22.** A set $A$ is **finite** if there is some natural number $n$, such that

$$A \sim \{k \in \mathbb{N} \mid k < n\} = \{0, 1, \dots, n-1\}.$$

**Infinite** if it is not finite; and, finally, **countable** if it is equinumerous with $\mathbb{N}$, namely:

$$A \text{ is countable iff } A \sim \mathbb{N}.$$

Least, the **finite cardinals** are the cardinal numbers of finite sets.

Finally:

**Theorem 3.2.2.** The set $\mathbb{N}$ of natural numbers is infinite.

*Proof.* The set $\mathbb{N}$ is characterized by the function $S$, which is defined as an injection $n \mapsto S(n)$ of $\mathbb{N}$ into $\mathbb{N}/\{0\}$. In other words, for given a number $n$, the function $S$ maps it to its successor, $S(n)$, in an open-ended process. ■

As we see, "$A$ is infinitely countable" means exactly "$A$ is equinumerous with $\mathbb{N}$", that is, there exists a bijection from $A$ to $\mathbb{N}$. In signs, $A \sim \mathbb{N}$ or, equivalently, $|A| = |\mathbb{N}| = \aleph_0$, where $\aleph_0$ standardly represents the cardinal number (or cardinality) of the set of natural numbers, i.e. of $\mathbb{N}$. Anyway, even if we've presented a way in which Cantor's intuitions on the set of natural numbers can be restored with the help of the axiomatic framework, Zermelo's interests were not primarily concerned with transfinite numbers:

> Zermelo's axiomatization also shifted the focus away from the transfinite numbers to an abstract view of sets structured solely by $\in$ and simple operations. For Cantor the transfinite numbers had become central to his investigation of definable sets of reals and the Continuum Problem [...]. Outgrowing Zermelo's pragmatic purposes axiomatic set theory could not long forestall the Cantorian initiative, as even $2^{\aleph_0} = \aleph_1$ could not be asserted directly, and in the 1920s John von Neumann was to fully incorporate the transfinite using Replacement.[65]

Indeed, it might be said that, while set theorists, after Zermelo's work, will try to secure transfinite numbers within axiomatic set theory, differently, «a substantial motive for Zermelo's axiomatizing set theory was to buttress his Well-Ordering Theorem by making explicit its underlying set existence assumptions»[66]. As remarked several times, much other technicalities and applications of the axioms **I.-VI.** could be interesting, but, since, in this context, our aim was that of simply giving an idea of how Zermelo accomplished the job of set-theoretically representing natural numbers, we stop our "Zermelian considerations" here.

## 3.2.5 von Neumann's contribution

### 3.2.5.1 Becoming Precise! (Consolidation)

In the general climate of the beginnings of axiomatic set theory, Zermelo had put aside the question of transfinite numbers – which was central to Cantor's project – in order to prove the Well-Ordering theorem by specifying the set-existence assumptions underlying his proof. Differently, John von Neumann, acknowledging Zermelo's and other results, accomplished the mission of giving a formal characterization of *ordinals*, i.e. of ordinal numbers. Von Neumann's analysis, indeed, carried at least two fundamental results concerning set theory, namely the establishing of the Recursion theorem and, consequently, with recursion at hand, the reconstruction of ordinals as sets.

[65] Kanamori 2009, pp. 15–16.
[66] Kanamori 2009, p. 13.

> Von Neumann, and before him Dimitry Mirimanoff and Zermelo in unpublished 1915 work, isolated the now very familiar concept of ordinal, with the basic idea of taking precedence in a well-ordering simply to be membership.[67]

So, as clear, Zermelo's proofs[68] of the well-ordering theorem have lead to the reduction of transfinite numbers to sets. Anyway, the novelties von Neumann brought into set theory, that we will consider, are all to be connected with his Recursion theorem, i.e., the proof that allows us to define concepts through transfinite recursion. As sketched in the foregoing section devoted to Zermelo's axioms, even in his 1904 paper, Zermelo, anticipated von Neumann's recursion theorem, but – and here's the crucial point – it has been explicitly formulated and proved just a couple of years later. Mathematicians, in connection to Zermelo's axiomatization, tried to solve different concerns, in particular they observed that – in order to prove transfinite recursion – a "fundamental" axiom was missing. Indeed, before von Neumann's complete axiomatization of set theory, in the 20s, Fraenkel and Skolem independently had found that – for treating sets with large cardinalities – the following axiom is required:

**Axiom** (**VIII**. Axiom of Replacement)**.** For each set $A$ and any definite operation $\mathscr{F}$ on $A$, the image of $A$ under $\mathscr{F}$ is a set. Formally,

$$\mathscr{F}[A] =_{\text{def}} \{\mathscr{F}(x) \mid x \in A\}.$$

Likewise, if we let $\overline{z} =_{\text{def}} z_1, \ldots, z_n$, then we may state the Axiom of Replacement as a "schemata", as follows:

$$\forall \overline{z}, \forall A \Big( (\forall x \in A)(\exists! y)\ \varphi(x, y, \overline{z}, A) \longrightarrow (\exists B), (\forall y)\, [y \in B \longleftrightarrow (\exists x \in A)\ \varphi(x, y, \overline{z}, A)] \Big)$$

Notice, that while in first-order logic we're not able to quantify over the notion of "definite function", such as $F$, then the inclusion of the schema for replacement is somehow needed. Indeed, consider that the advantages of a schemata, like **VIII**., is that any first-order definable function can be substituted to $\varphi$. Anyway, another interesting formulation of **VIII**. is the following, which we provide for its immediate clarity:

$$\forall A, \exists B, \forall C\ \Big[ C \in B \longleftrightarrow \exists D\ (D \in A\ \wedge\ C = \mathscr{F}[D]) \Big]$$

That is, given any set $A$, it exists a set $B$ such that, for all sets $C$, $C$ is an element of $B$ if and only if it exists a set $D$ such that $D$ is in $A$ and $C$ is identical to $D$ under $\mathscr{F}$. Hence, in other words, the essence of replacement is always the following: the image of any set under a definite condition is a set.
Now, let's think for a moment: Why did set theorists add an axiom, such as

[67] Kanamori 2009, p. 29.
[68] Zermelo 1904, 1908a.

Replacement, to their frameworks? Which help could have given to mathematicians? In order to answer this question we have to go back, for a moment, to Zermelo's axiom of infinity. Consider the set:

$$E(x) = \{x, \wp(x), \wp(\wp(x)), \dots\}.$$

In order to ensure that $E(x)$ is set precisely when $x$ is a set of the form of $\mathbb{I} = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \dots\}$, additional care was requested. Indeed, the set $\mathbb{I}$ has been postulated by Zemerlo's axiom of infinity and to ensure that $E(x)$ was another infinite set, Replacement has been added and applied. In this specific context, consider that the task of Replacement was that of "replacing" the sets belonging to $\mathbb{I}$ – as given by the axiom of infinity –, with other sets, in order to admit new infinite sets, such as $E(x)$. Indeed, for instance, consider what Skolem wrote in his 1922 paper:

> It is easy to show that Zermelo's axiom system is not sufficient to provide a complete foundation for the usual theory of sets. I intend to show, for instance, that if $M$ is an arbitrary set, it cannot be proved that $M, \wp(M), \wp(\wp(M)), \dots$, and so forth ad infinitum form a "set". [...]
> In order to remove this deficiency of the axiom we could introduce the following axiom: "Let $U$ be a definite proposition that holds for certain pairs $(a, b)$, in the domain $B$; assume, further, that for every set $a$ there exists at most one $b$ such that $U$ is true. Then, as $a$ ranges over the elements of a set $M_a$, $b$ ranges over all elements of a set $M_b$.".[69]

The first statements of the role of the axiom of replacement have been part of the work that Skolem – and Fraenkel almost at the same time – accomplished within set theory. But, despite the discover of all these advances, the history of axiomatic set theory does not end here and, indeed, «the full exercise of Replacement is part and parcel of transfinite recursion, which is now used everywhere in modern set theory, and it was von Neumann's formal incorporation of this method into set theory, as necessitated by his proof, that brought in Replacement»[70].
Despite the importance of these discovers, von Neumann's work in set theory has given a solution to Cantor's problem of representing ordinals within set theory. The question can be put as follows: let $A$ be a well-ordered set and let $\overline{A}$ be its ordinal type. How may we, then, associate to each well-ordered set $A$ its corresponding ordinal type $\overline{A}$? Von Neumann's ingenious work has the main advantages that, by allowing transfinite recursion, he can define $\overline{A}$, by *recursively* substituting each member of any $A$ by the set of its predecessors:

> What we really wish to do is to take as basis of our consideration: "Every ordinal is the type of the set of all ordinals that precede it". But, in order to avoid the vague

[69] We've slightly modified Skolem's original notation. See Skolem 1922, pp. 296–297.
[70] Kanamori 2009, p. 30.

> notion of "type", we express it in this form: "Every ordinal is the set of the ordinals that precede it". This is not a proposition proved about ordinals; rather, it would be a definition of them if transfinite induction [recursion] had already been established.

For clarity, we reconstruct von Neumann's procedure and, as always, we put it in a more contemporary fashion. Let's begin.

**Definition 23.** The **von Neumann map** of a well-ordered set $A$ is defined as follows:

$$\nu_A(x) =_{\text{def}} \{\nu_A(z) \mid z <_A x\}, \qquad (x \in A).$$

We, henceforth, define the **ordinal number** of $A$ to be the image of $A$ under its von Neumann's map, i.e., $\nu_A$:

$$\mathbf{ord}(A) =_{\text{def}} \nu_A(A).$$

Let, finally:

$$\mathbf{ON}(\alpha) \;\; =_{\text{def}} \;\; \alpha \in \mathbf{ON} \;\; =_{\text{def}} \;\; (\exists \text{ well} - \text{ordered } A)\,[\alpha = \mathbf{ord}(A)].$$

In this way, for instance, if we set up that:

$$A: \; 0_A, \; 1_A, \; 2_A, \ldots, \omega_A, \; S_A(\omega_A)$$

is a well-ordered set with a least element, i.e., $0_A$, first limit point, i.e., $\omega_A$ and last largest point, namely $S_A(\omega_A)$. If we repeat the application of the von Neumann map, with respect to any member of $A$, by simple computation, we obtain the following structure:

$$\begin{aligned}
&\nu(0_A) = \{\nu(x) \mid x <_A 0_A\} = \varnothing = 0\\
&\nu(1_A) = \{\nu(x) \mid x <_A 1_A\} = \{\varnothing\} = 1\\
&\nu(2_A) = \{\nu(x) \mid x <_A 2_A\} = \{\varnothing, \{\varnothing\}\} = 2\\
&\nu(3_A) = \{\nu(x) \mid x <_A 3_A\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} = 3\\
&\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots\\
&\nu(\omega_A) = \{\nu(x) \mid x <_A \omega_A\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \ldots\} = \omega\\
&\nu[S_A(\omega_A)] = \{\nu(x) \mid x <_A S_A(\omega_A)\} = \omega \cup \{\omega\}
\end{aligned}$$

And, finally, the image of the whole of $A$ under $\nu_A$:

$$\nu[A] \; = \; \{0, 1, 2, 3, \ldots, \omega, \omega \cup \{\omega\}\}$$

So, as requested, von Neumann's elaborated structure gives us what we were searching for, namely an alternative set theoretical construction suitable to faithfully represent the natural numbers sequence. From this statement many interesting Cantorian properties, with respect to the ordinals, were proved and, exactly for this reason, von Neumann's contribution – after Zermelo's great beginning – brought the transfinite, once for all, within set theory. Indeed, after von Neumann's characterization, we «think of ordinal numbers as standing for lengths of well-ordered sets, or for places of points in a well-ordered set. The latter agrees more with the use of ordinals in ordinary language, where "first", "second", ..., customarily describe the place of objects in a sequence»[71]. In order to accomplish the task of this alternative representation, let us define the fundamental notions of "least point", "limit point" and "first limit point".

**Lemma 2** (Least Point)**.** (a) If $0_A$ is the **least point** in a well-ordered set $A$, the $\nu_A(0_A) = \varnothing$.

(b) If $S(x)$ is the successor of $x$ in $A$, then,

$$\nu_A(S(x)) \;=\; \nu_A(x) \cup \{\nu_A(x)\}.$$

*Proof.* (a) For a function $f$, the image of the empty-set, $f[\varnothing]$ is empty. Let now $\nu_A : A \longrightarrow \mathbf{ord}(A)$. Since $0_A$ has no member that precede it,

$$\nu_A(0_A) = \nu_A[\{z \in A \mid z < 0_A\}] = \nu_A[\varnothing] = \varnothing.$$

(b) Now, consider $z <_A S(x)$. By the basic property of $S$ on a set $A$, it follows:

$$z <_A S(x) \longleftrightarrow z <_A x \;\vee\; z = x.$$

Using this fact as starting point, we can derive the desired conclusion in the following way:

$$\begin{aligned}\nu_A(S(x)) &= \{\nu_A(z) \mid z <_A S(x)\}\\ &= \{\nu_A(z) \mid z <_A x\} \cup \{\nu_A(x)\}\\ &= \nu_A(x) \cup \{\nu_A(x)\}.\end{aligned}$$

■

**Lemma 3** (Limit point)**.** (a) If $x$ is a **limit point** in a well-ordered set $A$, then $\varnothing \in \nu_A(x)$ and, (b)

$$\alpha \in \nu_A(x) \longrightarrow \alpha \cup \{\alpha\} \in \nu_A(x).$$

*Proof.* (a) If $0_A$ is least in $A$, then $0_A < x$. Moreover, then $\varnothing = \nu_A(0_A) \in \nu_A(x)$.

(b) Consider that $\alpha \in \nu_A(x)$. This means that $\alpha = \nu_A(z)$, for some $z <_A x$; but $z$ has a successor too, namely $S(z) <_A x$. By definition, since $x$ is a limit point, then

[71] Moschovakis 2006, p. 178.

it follows that $v_A(S(z)) \in v_A(x)$. By applying similar passages to the proof above, we conclude:

$$v_A(S(z)) = v_A(z) \cup \{v_A(z)\} = \alpha \cup \{\alpha\} \in v_A(x).$$

■

**Lemma 4** (First Limit Point)**.** If $\omega_A$ is the **first limit point** in a well-ordered set $A$, then:

$$\omega = v_A(\omega_A) = \bigcap \Big\{ C \mid \varnothing \in C \ \wedge \ (\forall \alpha \in C)\, [\alpha \cup \{\alpha\} \in C] \Big\}.$$

Moreover, $v_A(\omega_A)$ is independent of the particular well-ordered set $A$ chosen.

*Proof.* Consider:

$$T = \bigcap \Big\{ C \mid \varnothing \in C \ \wedge \ (\forall \alpha \in C)\, [\alpha \cup \{\alpha\} \in C] \Big\}.$$

Since $\omega_A$ is a limit point, by the previous lemma, we have that $v_A(\omega_A) \in T$ and $T \neq \varnothing$.

**Proof towards a contradiction**:

Assume that $C \in T$, such that $v_A(\omega_A) \nsubseteq C$, and let $z$ be least in $A$ such that $v_A(z) \notin C$. Now, $z$ is not the least member of $A$, since $v_A(0_A) = \varnothing$ and, by the hypothesis on $C$, it follows that $\varnothing \in C$. Additionally, $z$ is not a limit point of $A$, since $z < \omega_A$. Therefore, $z$ must be equal to some $S(x)$, for some $x \in A$; this means that $v_A(x) \in C$. Hence, we are allowed to conclude, by the hypothesis on $C$, that $v_A(z) = v_A(x) \cup \{v_A(x)\} \in C$. So, as clear, this result contradicts our initial assumption on $z$, i.e., $v_A(z) \notin C$ – terminating, thus, the proof. ■

Hence, as before, – with this entire package of well-defined notions – we may fix a specific (and unique) Peano system – such as $(\mathbb{N}, 0, S)$ –, where $\mathbb{N} = \omega$, $0 = \varnothing$ and the successor function is determined as always:

**Theorem 11** (von Neumman's finite ordinals)**.** If $<_{\mathbb{N}}$ is the ordering relation on the set $\mathbb{N}$ of natural numbers, then

$$\mathbf{ord}(\mathbb{N}, <_{\mathbb{N}}) = \omega,$$

where $\omega$ is the first limit point. Moreover, by letting:

$$S_\omega(n) = n \cup \{n\} \quad (n \in \omega),$$

we define $(\omega, \varnothing, S)$ as a Peano system and $v_{\mathbb{N}} : \mathbb{N} \longrightarrow \omega$ as the unique isomorphism of $\mathbb{N}$ with $\omega$.

*Proof Sketch.* Let $A = \mathsf{Succ}((\mathbb{N}, <_{\mathbb{N}})) = \omega$ is the set "next" to $(\mathbb{N}, <_{\mathbb{N}})$ with $\mathsf{t}$ on the top, then we have that $\mathsf{t} = \omega_A$, $\mathbb{N}$ is identical to an initial segment of $A$, up to $\mathsf{t}$, and $\mathbf{ord}(\mathbb{N}, <_{\mathbb{N}}) = \nu_A(\mathsf{t}) = \omega$. The desired result, then, immediately follows from the proof of the Least Point lemma.[72] ■

This theorem is of particular importance since it allows us to secure ordinals and cardinals. Since, we've defined what is an ordinal – i.e., $\mathbf{ord}(A) =_{\text{def}} \nu_A(A)$ –, and by, additionally, having the precise definite conditions of "successor" and "limit" element, we may state the following notions.

**Definition 24.** (a) If $P(\alpha)$ holds, then

$$(\mu\alpha \in \mathbf{ON})\, P(\alpha) = \mathsf{min}\{\alpha \in \mathbf{ON} \mid P(\alpha)\},$$

where the greek letter "$\mu$" has to be read as "the least".

(b) For each ordinal number $\alpha$, there is a next one denoted by $S(\alpha)$:

$$S(\alpha) =_{\text{def}} (\mu\delta \in \mathbf{ON})[\alpha < \delta] = \alpha \cup \{\alpha\}.$$

(c) Each set $C$ of ordinal numbers has a **least upper bound**, denoted as:

$$\mathsf{sup}(C) =_{\text{def}} (\mu\delta \in \mathbf{ON})(\forall\alpha \in C)[\alpha \leq \delta] = \bigcup(C).$$

(d) A **limit ordinal** is an ordinal $\lambda$ which is not successor of 0 and so that, for any ordinal $\alpha$, it holds that $\alpha < \lambda \longrightarrow S(\alpha) < \lambda$.

(d') Formally, a limit ordinal $\lambda$ satisfies the following property:

$$\mathsf{Limit}(\lambda) \longleftrightarrow \lambda \neq 0 \ \wedge\ \lambda = \mathsf{sup}\{\alpha \mid \alpha < \lambda\}.$$

Anyway, this is not even a part of the work that von Neumann's theorem allows us to do. In particular, by defining ordinal multiplication and addition[73],we can, finally, encounter the successor of $\omega = (\mathbb{N}, \leq_{\mathbb{N}})$:

$$\omega + 1 = S(\omega),\ \omega + 2 = S(\omega + 1),\ \omega + 3 = S(\omega + 2), \ldots$$

and, especially, the second limit ordinal "immediately above" $\omega$, namely:

$$\omega + \omega = \mathsf{sup}\{\omega + n \mid n \in \omega\} = \omega \times 2.$$

Similarly, by adding $n$ times $\omega$ to itself, then we may obtain $\omega \times n$. After a bit, we may arrive at $\omega^2$, which is:

[72] We do not cover the entire proof, otherwise other technical notions would have to be proved. Anyway, for a complete proof and definition of the concepts involved, consider Moschovakis 2006, pp. 91, 177, 265, respectively.

[73] That is, $\alpha + \beta$ and $\alpha \times \beta$.

$$\omega \times \omega = \mathsf{sup}\{\omega \times n \mid n < \omega\} = \omega \times 2.$$

Generally, hence:

$$\omega^n = \omega^{n-1} \times \omega$$

Additionally to ordinal arithmetic, we've said several times that von Neumann's set theoretical construction secures also cardinal arithmetic. We, indeed, may get the notion of "cardinal assignment" for a well-ordered set:

**Definition 25** (von Neumann's Cardinals)**.** We define a **cardinal number** by setting:

$$|A| = \begin{cases} (\mu\xi \in \mathbf{ON})[A \sim \xi] & \text{(if } A \text{ is well} - \text{orderable),} \\ A, & \text{(otherwise).} \end{cases}$$

Thus, the value of $|A|$ for any well-orderable set $A$ is defined as follows:

$$\mathsf{Card}(\kappa) \longleftrightarrow \text{for some well} - \text{orderable } A,\ \kappa = |A|.$$

Moreover, what is important in von Neumann's approach is that cardinals are defined as the "initial ordinals", establishing thus, till to the end, the connections between ordinals and cardinals:

**Proposition 12.** That a cardinal $\kappa$ is an **initial cardinal** means that:

$$\mathsf{Card}(x) \longleftrightarrow \kappa \in \mathbf{ON}\ \wedge\ (\forall \alpha < \kappa)[\kappa \nsim \alpha]$$

Moreover, for every $\kappa \in \mathsf{Card}$, it holds that:

$$|\kappa| = \kappa.$$

■

So, this is the general way in which cardinal and ordinal arithmetic are secured within set theory. As before, some fundamental details are here left unspecified, anyway for our purposes this is enough.

As we have just sketched, the work of to defining and representing natural numbers with the help of sets, has not been that easy and, indeed, while philosophically reflecting on its major contributions, we have to consider, as much as possible, all the technical issues that the axiomatization of set theory has involved. Indeed, much many other important discoveries were done, after von Neumann's characterization of ordinals, and indeed, the same Zermelo, in the 30s, returned to sets and presented what we nowadays call the "Zermelo Universe"[74]. As briefly announced, von Neumann,

[74]Consider, now, that both – Zermelo and von Neumann – adopted an axiom of infinity, but, that von Neumman's version is different since it encapsulates his usage of the Kuratowski pair. Indeed, the axiom von Neumann used, can be stated as follows: $\exists \mathbb{I}\ [\varnothing \in \mathbb{I} \wedge \forall x\ (x \in \mathbb{I} \longleftrightarrow x \cup \{x\} \in \mathbb{I})]$.

thanks to the intuitions of Fraenkel and Skolem, consistently added an axiom, the so-called axiom of replacement, – implicit in many mathematical theorems – which allowed him to reconstruct a valid axiomatization of set theory. His article (1925)[75] is difficult to access, but thanks to the simplifications and clarifications P. Bernays (1937)[76] provided, it is now standard to take von Neumann's 1925 work as a pioneer strategy in order to get the division between "sets" and "classes". The work von Neumann established was based upon an intricate division between functions and arguments, which was efficiently adapted – thanks to Bernays – to the division between classes and sets and which, fundamentally, has been applied from Gödel (1940)[77] to provide the result that Cantor's Continuum Hypothesis is consistent if added to the axioms of Zermelo (plus Replacement). Hence, for clarity, consider that *classes* are different from sets in some respects otherwise Russell's Paradox may be derived again. Recall that we have given a precise characterization of the notion definite condition, which was absent from Zermelo's original work. Now, as sketched, $n-$relations and $n$-properties – from a set theoretical point of view – are regarded as subsets of given set. This means that an arbitrary condition $P$ on a set $A$, set theorists will regard $P$ as a set, "under which" the objects of the superset $A$, which satisfy condition $P$, fall. This little background will help us in defining a class in the following way:

**Definition 26** (Coextensionality and Classes)**.** (a) A unary definite condition $P$ is **coextensive** with a set $A$ if the objects which satisfy $P$ itself are precisely the members of $A$. Formally:

$$P =_e A =_{\text{def}} \forall x\,(P(x) \longleftrightarrow x \in A).$$

(b) A **class** is either a set or a unary definite condition which is not coextensive with a set. For any unary condition $P$, hence, we establish that the class:

$$\{x \mid P(x)\} = \begin{cases} \text{the unique set } A \text{ such that } P =_e A, & (\text{if } P =_e A \text{ for some set } A), \\ P, & (\text{otherwise}). \end{cases}$$

If we denote the class of $P$ as follows:

$$A = \{x \mid P(x)\},$$

then it holds that:

$$\forall x\,(x \in A \longleftrightarrow P(x)).$$

For example, pick $P(x) =_{\text{def}} x \neq x$. Consider that, from our definition it follows that there is a set $A = \{x \mid x \neq x\}$, which is coextensive with $P$. But, by inspection,

[75] See von Neumann 1925, pp. 393–413 and Moschovakis 2006, pp. 175–197, for technical introduction and commentaries.

[76] See Kanamori 2009, pp. 29–31.

[77] See Kanamori 2009, pp. 35–40.

$A = \varnothing$ and, hence $P$ is coextensive with the empty set, i.e., $P =_e \varnothing$. Now, for instance, let $A =_{\text{def}} \{x \mid P(x)\}$. According to our definition, then, (i) if $P$ is coextensive with a set, $P =_e A$, then, $x \in A \longleftrightarrow P(x)$; else, (ii) if $P \neq_e A$, then, $A = P$, and, hence,

$$x \in A \longleftrightarrow x \in P \longleftrightarrow P(x).$$

So, as it emerges from our definition, every set is a class, but – to avoid paradoxical situations – not every class determines a set. To see this, consider, for *reductio*, that any class is a set. Suppose that:

$$P(x) \longleftrightarrow \mathsf{Set}(x) \ \wedge \ x \notin x.$$

Now, from our definition we obtain the following set:

$$A = \{x \mid \mathsf{Set}(x) \ \wedge \ x \notin x\}.$$

But this means that $x \in A \longleftrightarrow P(x)$, that is, $x \in A \longleftrightarrow \mathsf{Set}(x) \ \wedge \ x \notin x$. Suppose $A$ to be such a set and conclude $A \in A \longleftrightarrow A \notin A$, which is absurd. Therefore, our initial assumption that any class is a set has to be refuted.

However – and not surprisingly –, indeed, the now standard name for the theory which encapsulates the division between classes and sets is "von Neumann - Bernays - Gödel", NGB, set theory. Anyway, returning to von Neumann's work, one last thing deserves to be mentioned, i.e. a brief reflection on the final considerations he presented in his 1929[78] paper devoted, another time, on axiomatic set theory. What emerged in the context of this new paper, is exactly the necessity for a set to be "well-founded" and, indeed, in order to accomplish this latter requirement, von Neumann established the so-called Axiom of Foundation[79]. In logical terms, the axiom can be stated by clarifying some preliminary concepts:

**Definition 27** (Ill- and Well-Founded Sets)**.** Let $E$ be a set and $g$ a function such that $E \longmapsto E$. Let $u \in E$ be such that it exists a unique function $f : \mathbb{N} \longrightarrow E$, satisfying the following two conditions:

$$f(0) = u$$
$$f(n+1) = g[f(n)].$$

(a') An object $x$ is **ill-founded** if it is the beginning of a descending $\in$-chains. This means, that $x$ is ill-founded if there is a function $f : \mathbb{N} \longrightarrow E$, such that

$$x = f(0) \ni f(1) \ni f(2) \ni \ldots.$$

(a") Equivalently, an object $x$ is ill-founded if

$$\exists A \Big[x \in A \ \wedge \ (\forall z \in A)(\exists u \in A)[z \ni u]\Big].$$

[78]See von Neumann 1929, pp. 227–241. As always, our notation is adapted to the contemporary one.

[79]In literature, this axiom is sometimes called the axiom of Regularity.

(b) Finally, objects which are not ill-founded are, thus, **well-founded** or **grounded**[80].

Thus, we include, within our axiomatization, this last following axiom:

**Axiom** (**IX**. Axiom of Foundation)**.** Every set is well-founded.

$$\forall x \Big( x \neq \varnothing \longrightarrow (\exists z \in x)[x \cap z = \varnothing] \Big).$$

## 3.3 Zermelo's *vs* von Neumann's Sets

One of the most elegant and successful applications of the axiom of foundation is within von Neumann's construction of the so-called *cumulative hierarchy*. His set-theoretical construction allowed him to prove the first consistency results in set theory through "inner models":

> In the axiomatic tradition Fraenkel, Skolem and von Neumann considered the salutary effects of restricting the universe of sets to the well-founded sets. Von Neumann formulated in his functional terms the Axiom of Foundation, that every set is well-founded [...].[81]

Actually, with the notion of "founded", von Neumann restricted the universe of sets to those that may be called "well-founded", via his method of transfinite recursion. In contemporary set theory, this, as it is widely known, entails that the universe $\mathcal{V}$ is "stratified" into cumulative ranks $\mathcal{V}_\alpha$:

**Definition 28** (Cumulative Hierarchy of Grounded Sets and von Neumann Universe)**.** For each ordinal $\alpha$ the set $\mathcal{V}_\alpha$ is defined by recursion on **ON**, as follows:

$$\mathcal{V}_0 = \varnothing$$

$$\mathcal{V}_{\alpha+1} = \wp(\mathcal{V}_\alpha)$$

$$\mathcal{V}_\lambda = \bigcup_{\alpha<\lambda} (\mathcal{V}_\alpha), \text{ for } \lambda \text{ limit ordinal.}$$

So, finally, the **von Neumann Universe** is the union of all sets $\mathcal{V}_\alpha$:

$$\mathcal{V} =_{\text{def}} \bigcup_{\alpha \in \mathbf{ON}} (\mathcal{V}_\alpha) = \{x \mid (\exists \alpha \in \mathbf{ON})\, [x \in \mathcal{V}_\alpha]\},$$

and the **rank** operation on it:

$$\mathsf{Rank}(x) =_{\text{def}} (\mu\alpha \in \mathbf{ON})[x \in \mathcal{V}_{\alpha+1}],\ \ (x \in \mathcal{V}).$$

[80]Notice that ill-foundedness is a notion strictly connected to self-membership, i.e., for instance, if $A \in A$, then $A \ni A \ni A \ni \ldots$.

[81]Kanamori 2009, p. 31.

With these formal setting, henceforth, von Neumann proved the consistency of the axiom of foundation with respect to Zermelo's axioms **I.-VII.**, plus Replacement, by developing always in a deeper way set theory.

As we have seen, von Neumann's work in set theory provides a useful construction to study natural numbers, ordinals and cardinal arithmetic. We've introduced von Neumann's hierarchical universe of sets, but, now, consider that – almost in the same years – that is, in the 30s, Zermelo came back to set theory by proposing a "new" axiomatization[82]. In this more recent study, Zermelo took in consideration all the axiomatic studies that were proposed up to the 30s, by – amongst other things – adopting Replacement and Foundation. The most interesting consequence, that is possible to draw from Zermelo's 1930 paper, is what set theorists nowadays call "least Zermelo Universe". In what follows, we will introduce how it is possible to "construct" it and sketch the main differences underlying between Zermelo's and von Neumann's Universes.

**Definition 29** (Transitivity)**.** (a) A class or set $T$ is **transitive** if:

$$\bigcup(T) \subseteq T.$$

(b) Equivalently:

$$(\forall x \in T)\,(\forall z \in x)\,[z \in T].$$

(c) Or, equivalently:

$$(\forall x)\,[x \in T \longrightarrow x \subseteq T].$$

And, through the notion of transitivity the definition of a Zermelo universe follows:

**Definition 30** (Zermelo Universe)**.** A transitive class $T$ is a **Zermelo Universe** if it is closed under the following operations:

(i) Pairing, $\{x, y\}$;

(ii) Unionset, $\bigcup(C)$;

(iii) Powerset, $\wp(C)$.

Additionally, $T$ contains:

$$\mathbb{N}_0 = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\{\{\varnothing\}\}\}, \{\{\{\{\varnothing\}\}\}\}, \dots\}.$$

The **Least Zermelo Universe** is $\mathcal{Z} = T(\mathbb{N}_0)$ determined by the following identity conditions:

[82]Zermelo 1930.

$$\mathcal{Z}_0 = \mathbb{N}_0;$$

$$\mathcal{Z}_{n+1} = \wp(\mathcal{Z}_n);$$

$$\mathcal{Z} = \bigcup_{n=0}^{\infty} (\mathcal{Z}_n).$$

So, this is the result of Zermelo's work in the 30s. Let's focus, for a moment, upon the differences lying between our two set-theoretical universes, $\mathcal{Z}$ and $\mathcal{V}$.

### 3.3.1 Reconsidering Benacerraf's thesis II

**Introductory remarks** For the moment, we are going to introduce an example that should clarify our main intent[83]. Consider, for instance, the set $S = \{1, 2\}$ and assume that it is partially ordered. This means that the elements in our set are ordered with respect to a reflexive, anti-symmetric and transitive relation. Moreover, notice that we may assume that the ordering relation between the elements of $S$ is $\leq$ and, hence, our set is ordered as follows: $1 < 2$. We notice, additionally, that any of the elements of $S$ is comparable, i.e., for all $x$, $y$, $x < y \vee x = y \vee y < x$. This brings us in calling the set $S$ as totally (or linearly) ordered. Our idea is now to consider – as in our foregoing remarks – the already familiar notion of "partition". We recall that a set is said to be a partition of a set $S$ if it is a family of non-empty disjoint subsets of $S$, such that their union is equal to $S$ itself. From a mathematical point of view, a partition, along with a total order, on the sets belonging to the partition, gives a structure called "ordered partition". With respect to our example, the ordered partitions of $\{1, 2\}$ are the following three:

$$\{1\}, \{2\}$$

$$\{2\}, \{1\}$$

$$\{1, 2\}.$$

Now, suppose that our $S = \{1, 2\}$ is a subset of $\mathbb{N}$, which is – as clear – partially ordered under $\leq$. In this sense, by considering $S \subseteq \mathbb{N}$, $S$ is considered as a **chain**. Indeed, we may define chains exactly as totally ordered subsets of partially ordered sets. In our example, by simple inspection, of course if $S \subseteq \mathbb{N}$ and $(\mathbb{N}, \leq)$, then also $(S, \leq)$. Thus, any chain of a partial ordered set, such as $\mathbb{N}$, keeps the same order relation to the superset to which it belongs. Clearly, therefore, since in $\mathbb{N}$, $1 < 2 < 3 < \ldots$ and since $S$ is a subset of $\mathbb{N}$, then $1 < 2$, i.e. the order relation is preserved. At this point a fundamental question must be asked: if any two elements $a$ and $b$ stand in a total order relation, $a < b$, is it possible to introduce a third element $c$ between them, such that $a < c < b$ without losing the total order? The answer is "no" and the notions of "saturated" and "maximal" chains may reveal useful. Firstly,

[83] The following remarks are to due, as always, to Moschovakis 2006, but, especially, to Harzheim 2006, pp. 11–70.

a chain such that no element can be added between two of its elements without losing the property of being totally ordered is said a saturated chain. But, similarly, a chain in a partially ordered set to which no element can be added without losing the property of being totally ordered and that excludes the existence of elements either less than all elements of the chain or greater than all its elements, is said to be maximal. A finite saturated chain is maximal if and only if it contains both a minimal and a maximal element of the partially ordered set to which the chain itself belongs. For instance, let's take the set $S_n = \{n \mid n \in \mathbb{N}\}$, that is the set of natural numbers "below" $S_n$. Moreover, we can easily see that $S_n$ is a chain, i.e. a totally ordered subset of $\mathbb{N}$. What we have to notice is that, by inspection, our $S_n$ has a minimal and maximal element belonging to the partially ordered set $\mathbb{N}$ to which it belong, being thus a "maximal chain". Indeed, for example, a set $S_n = \{1, 2\}$ ordered by $\leq$ is a maximal chain of $\mathbb{N}$ since (i) no third element $c$ may be added between 1 and 2, without losing the total order and (ii) it contains a minimal element, i.e. 1, and a maximal member, 2, both from the partially ordered set $\mathbb{N}$ of which $S$ is a chain. Additionally, for $S$ being a chain of $\mathbb{N}$, we may be interested to its "length". In other terms, we may interested in seeing how long a chain of a partially ordered set may be. Hence, if $S \subseteq \mathbb{N}$ is a chain such that $S$ is finite and different from the empty set, then its length is determined by the cardinality of $S$ itself minus one, i.e. $\mathsf{len}(S) = |S| - 1$. For example, if $S = \{a, b\}$, we may compute its length as follows: $\mathsf{len}(S) = 2 - 1 = 1$.

Now, by considering order relation we have always assumed $\leq$, but there are other order relations from which we may get a partial order. Consider a set of subsets $A = \{B, C, D\}$ and let $\subseteq$ be the relation of inclusion on the elements in $A$. We denote with $(A, \subseteq)$ the "relational structure" of $A$ as determined by $\subseteq$. To check that it is a partial order relation, we see that the inclusion relation is reflexive, anti-symmetric and transitive:

1. **Reflexivity**: $\forall B \in A \Big(B \subseteq B\Big)$;
2. **Anti-symmetry**: $\forall B, C \in A \Big(B \subseteq C \wedge C \subseteq B \longrightarrow B = C\Big)$;
3. **Transitivity**: $\forall B, C, D \in A \Big(B \subseteq C \wedge C \subseteq D \longrightarrow B \subseteq D\Big)$. ■

So, we conclude that two fundamental partial order relations are the "less than or equal" relation on a set of natural numbers and the "subset" relation on a set of sets. In this spirit, we introduce just a few other notions connected to chains and order relations. We have already clarified the usage of "lengths" with respect to chains, what we now have to sharply distinguish are the two fundamental notions of "height" and "rank". Let $(A, \leq)$ be a finite partially ordered set, and $a$ an element of $A$. The height of $a$ in $A$, denoted by $\mathsf{ht}(a)$, is the greatest natural number $n \in \mathbb{N}$, so that there exists a chain $\{a_0, \ldots, a_n = a\}$, where $a_0 < \cdots < a_n$. In this sense, the height indicates us the greatest number for which a chain of $A$ with $\mathsf{ht}(a) + 1$ elements and which ends in $a$ exists. Furthermore, For $n \in \mathbb{N}$, let $L_n$ denote the set of all elements

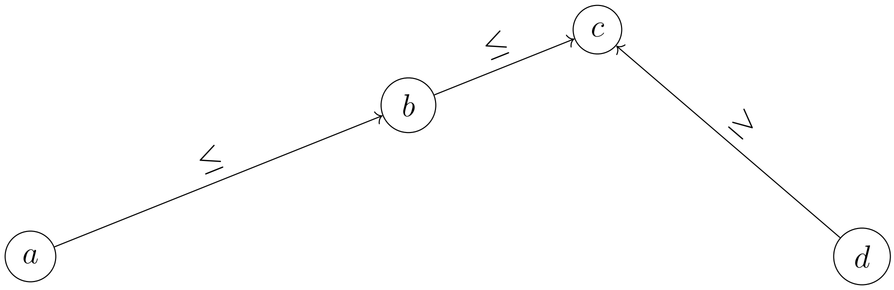



Figure 3.1: Graph of the set $(A, \leq)$

of $A$ which have height $n$, i.e. the *n-level* or *height-n-set* of $A$. If $A \neq \varnothing$, we say that $\mathsf{ht}(A)$ is the *height* of $A$, i.e., the maximum of all numbers $\mathsf{ht}(a) + 1$, for $a \in A$. Differently, but in a very similar fashion, we may define "ranks". Let $(A, <)$ be a partially ordered set in which all chains are finite. Then a mapping $\rho : A \longrightarrow \mathbb{N}$ is called a rank function, if for all $a, b \in A$ with $a < b$ there holds $\rho(b) = \rho(a) + 1$. So let's call $\rho(a)$ the rank of $a$. For $n \in \mathbb{N}$ we denote the set of elements of $A$ which have the rank $n$ by $\rho_n(A)$ and let's call it the $n^{th}$-rank of $(A, <, \rho)$. So, if $A$ is finite, and if the least number of $\{\rho(a) \mid a \in A\}$ is 0, then $\mathsf{max}\{\rho(a) \mid a \in A\}$ is called the rank of $A$. As it should be clear, height functions are defined for any finite partially ordered set, while, generally, rank functions do not apply to finite ones.

**Example.** In order to be clear, we will represent this example graphically. For instance, consider the partially ordered set $(A, \leq)$, such that $A = \{a, b, c, d\}$ and $a \leq b \leq c$ and $d \leq c$.

Consider Figure 3.1. By inspection, $\mathsf{ht}(a) = 0$, $\mathsf{ht}(b) = \mathsf{ht}(a) + 1 = 1$ and, therefore, the height of $c$, namely $\mathsf{ht}(c) = \mathsf{ht}(b) + 1 = 2$. Differently, the height of $d$ is $\mathsf{ht}(d) = 0$. Additionally, consider that $A$ has a rank function $\rho$, for which $\rho(a) = 0$, $\rho(b) = \rho(d) = 1$ and $\rho(c) = 2$. Notice that while the height of $d$ is 0, its rank is 1 and, hence, height and rank do not always coincide. Instead, for the other three elements of $A$ heights and ranks are identical. In our case, in informal terms, the height of $d$ is 0 since it occupies the first position with respect to the order relation $d \leq c$, but its rank is 1. The ranks of the 3-ordered relation $a \leq b \leq c$ are, $\rho(a) = 0$, $\rho(b) = 1$, and $\rho(c) = 2$. Now, consider that the 2-ordered relation $d \leq c$ has always $c$ as greatest element of the sequence and that this relation – with respect to the other – lacks the 0-ranked member, i.e., the element occupying the position before $d$. Indeed, since, in the other sequence, $b$ comes immediately before $c$ and its rank is 1 (by $a$ being of rank 0), it may be concluded that $d$ – by immediately preceding $c$ – has the same rank as $b$, i.e., 1. In more simple words, the height measures the greatest number of elements contained in a chain and a rank informs us about the "distance" between one element and another one of the same chain. In this sense, always with respect to our example, it could be said that $a$ and $d$ – by sharing the same height – belong to the same $n$-level set $L_0$ of the partially ordered set $(A, \leq)$.

In the same vein, $b$ and $d$ – by sharing their rank – belong to the same $n^{th}$ set $\rho_n$ of $(A, \leq)$, i.e., in this case, to $\rho_1(A, \leq)$.

In turn, we consider – with respect to this consideration – the two set-theoretical construction we've introduced in this long chapter, namely Zermelo's and von Neumann's. Let's establish that:

$$\mathcal{V} = \varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \ldots$$

and

$$\mathcal{Z} = \varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \ldots$$

are the two sets under inspection. We have seen that if we compare them, by taking the membership relation, then two representations show up as divergent. In particular, exactly from this divergence, as remarked several times, Benacerraf raised doubts against the possibility of identifying the two sets reductions. Indeed, if two objects $a$ and $b$ are both identical to a third one, call it $c$, then it is not admissible that $a$ and $b$ are different one with respect to the other. According, hence, to the French philosopher, the situation we will encounter is the following: let $a = \{\varnothing, \{\varnothing\}\}$, $b = \{\{\varnothing\}\}$ and $c = 2$. So, we know that $a = c$ and $b = c$, but that also – with respect to $\in$ – $a \neq b$ and, therefore, also $c \neq c$, which is, of course, absurd. We have said that a set can be investigated also by considering its partial order relations, such as $\leq$ or $\subseteq$. Indeed, if we get, for a moment, rid of the usage that Benacerraf made of the $\in$-relation, then a different conclusion may be drawn.
Let's consider finite subsets of both partially ordered sets, for example, $V_n \subseteq \mathcal{V}$, such that.

$$V_n = \{V \subseteq \mathcal{V} \mid V \text{ is finite.}\}, \text{ for some } n \in \mathbb{N},$$

and $Z_n \subseteq \mathcal{Z}$, such that:

$$Z_n = \{Z \subseteq \mathcal{Z} \mid Z \text{ is finite.}\}, \text{ for some } n \in \mathbb{N}.$$

Furthermore, the two collections $\mathcal{V}$ and $\mathcal{Z}$ can be considered as partially ordered under the relation $\subseteq$ and hence we may consider the two relational structures $(\mathcal{V}, \subseteq)$ and $(\mathcal{Z}, \subseteq)$. In this sense, also any of the subsets $V_n \subseteq \mathcal{V}$ and $Z_n \subseteq \mathcal{Z}$ is ordered by the subset relation and, moreover, any of them may be considered as a chain, which respects the maximality condition.

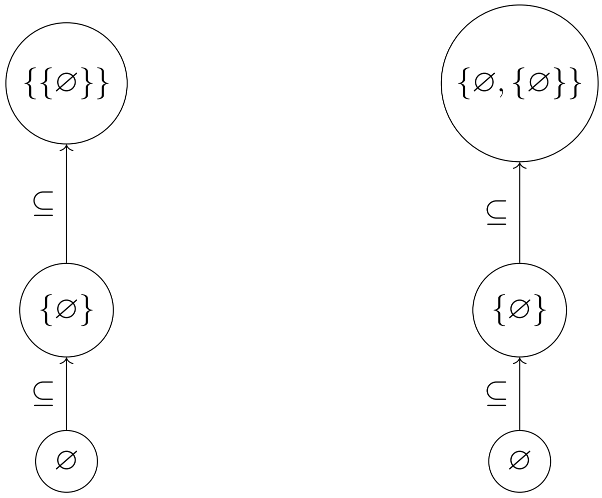
**Example.** Pick $V_2 = \{\{\varnothing\}\}$ and $Z_2 = \{\varnothing, \{\varnothing\}\}$. Consider that both are chains (totally ordered sets) of the partial ordered sets $\mathcal{V}$ and $\mathcal{Z}$. We assume that the relation which defines the order is, in both cases, the $\subseteq$ relation. We easily check that

$$\{\{\varnothing\}\} = \varnothing \subseteq \{\varnothing\} \subseteq \{\{\varnothing\}\}$$

and that

$$\{\varnothing, \{\varnothing\}\} = \varnothing \subseteq \{\varnothing\} \subseteq \{\varnothing, \{\varnothing\}\}.$$

Additionally, if any other element $c$ is inserted between each of the sets standing in the $\subseteq$ relation, then both representations loose their total order, so for example, nor $\{\varnothing, \{\varnothing\}, c\}$, nor $\{\{c, \varnothing\}\}$ are valid chains of $\mathcal{V}$ and $\mathcal{Z}$. Additionally, as clear, both of them have the least and maximal element taken from the partially ordered sets to which they belong. Thus, both, $V_2$ and $Z_2$, preserve the maximality condition. Furthermore, graphically, the two representations share sameness in height and ranks, always with respect to $\subseteq$ relation. For what concerns Zermelo's representations of natural numbers we obtain the directed graph as represented in the following figure:




With respect to the height of $Z_2$ we easily see that $\mathsf{ht}(\varnothing) = 0$, $\mathsf{ht}(\{\varnothing\}) = 1$ and, finally, $\mathsf{ht}(\{\{\varnothing\}\}) = 2$. Identically, the height of the chain $V_2$ we easily see that $\mathsf{ht}(\varnothing) = 0$, $\mathsf{ht}(\{\varnothing\}) = 1$ and, finally, $\mathsf{ht}(\{\varnothing, \{\varnothing\}\}) = 2$. Both sets, namely $Z_2$ and $V_2$, share identity between ranks and heights: $\rho(\varnothing) = 0$, $\rho(\{\varnothing\}) = 1$ and $\rho(\{\varnothing\}) = \rho(\{\varnothing, \{\varnothing\}\}) = 2$.

Now, for our purposes, we may consider, indeed, that $Z_2$ and $V_2$ represent the *same* maximal chain:

$$Z_2 = \varnothing \subseteq \{\varnothing\} \subseteq \{\{\varnothing\}\}$$

$$V_2 = \varnothing \subseteq \{\varnothing\} \subseteq \{\varnothing, \{\varnothing\}\}.$$

Indeed, what we mean with the term *same* is that $Z_2$ and $V_2$ are not identical, but just equivalent chains (ordered by $\subseteq$) under the condition of maximality, by additionally knowing that both set theoretical constructions allows us to define the natural number "2".

By considering these last reflections we will try to analyse again Benacerraf's criticism.

Our main inter is threefold in this section. We will, indeed, consider **(1)** the logical-mathematical strategy of Benacerraf, **(2)** his philosophical conclusions and **(3)** establish a Dedekind-style reading of mathematical representations.

#### 3.3.1.1 Logical-mathematical Development of Benacerraf's Paper

Already at a first reading of Benacerraf's article[84], it is easy to notice that the strength of the article is not obtained by carefully analysing axioms and Zermelo's or von Neumann's proposals. From a mathematical point of view, Benacerraf tried to show that the identification of sets and numbers is wrong, while, philosophically speaking, he tries to argue against abstract objects. Taking into account Benacerraf's objectives, his analysis – as developed in the article – is misleading for two main reasons: (i) it does not offer an overview concerning the foundational importance of the set theoretical study of the natural numbers and (ii) simplifies in a exceedingly exaggerate way the comparison between Zermelo's and von Neumann's sets. In the foregoing section,indeed, we have tried to show that by shifting the attention from the $\in$-relation to the $\subseteq$-relation, Zermelo's and von Neumann's sets become equivalent. Hence, as it emerges, Benacerraf's argumentative strategy is somehow *ad hoc* constructed in order to arrive to his desired philosophical conclusion and – apart not considering the extraordinary complexity of set theory – he, additionally, did not dedicate any part of his paper by distinguishing naïve sets from axiomatic ones. What Benacerraf identified as the main problem for the re-conduction of numbers to sets, can somehow be avoided by taking in considerations other relations between sets and, indeed, an implicit error in Benacerraf's strategy is, according to us, to consider the membership relation as the sole relation that allows to argue in favour or against set theoretical representations of mathematical objects. Therefore, Benacerraf's *ad hoc* path from the reflections concerning the $\in$-relation, to his philosophical conclusions, holds just in case no other relation between sets is introduced. In this sense, our path has been different by considering the $\subseteq$-relation and, by additionally, trying to furnish a clear insight on how sets (Zermelo's and von Neumann's) were supposed to accomplish the work of representing the natural numbers sequence. Indeed, from the early developments and, hence, from the first reflections of Frege, Cantor and Dedekind, set theory – by becoming always more "mathematized" and consolidated – has furnished the "arena" where a great amount of mathematics can be secured. Indeed, by not seeing the faithfulness of representing mathematical objects trough sets, Benacerraf seems, additionally, not to consider that – especially thanks to Zermelo's pioneer work – foundational issues and mathematical ones became always more separated. Clearly, this does not mean that mathematics and philosophy of sets, for instance, are deeply disconnected, but, more simply, that set theory represents a proper mathematical discipline with its own subject matter and that additionally can be helpful in discussing foundational, i.e., philosophical, problems concerning mathematics itself. Indeed, if we take care of this distinction, then – as we have tried to do – we have to consider, firstly, the mathematical theory itself, by focusing the attention, as much as possible, on its technical and historical exigences and, than, trying to understand its philosophical consequences and its foundational power. For example, during the $19^{th}$ century, the problem of finding the correct systematization

[84] We always refer to Benacerraf 1965.

of the negative integers $(\dots, -2, -1, 0, 1, 2 \dots)$, within the universe of numbers, has to be considered as a foundational concern of mathematicians. Differently, the fact, for instance, that an imaginary number $-i$, if multiplied to itself, gives as result always a negative number, – contradicting, thus, the basic rules for positive and negative signs –, has to be considered as a pure mathematical problem to solve. Here, as I am trying to argue, the same distinction applies, even if Benacerraf disregarded it: e.g., considering whether arithmetic can be secured within set theory means to be worried on foundational issues about a determined branch of mathematics; but, differently, considering problems – such as the presence of paradoxes, for instance – and, by trying to solve to them, means to focus the attention on the mathematics and logic of a determined discipline or mathematical branch. For our concerns, indeed, we are not aiming to show or to comment mathematical features of sets, but, we aim to use these latter properties of sets as fertile basis for the foundational discussion of arithmetic.

In conclusion, – I believe –, Benacerraf's path, that moves mixed philosophical and mathematical criticism to Zermelo's and von Neumann's set identification, by concluding against Platonism in general, is not that tenable. Moreover, we claim, Benacerraf's argumentation is, additionally, not so informative and precise about sets and their "identity" with the natural numbers sequence. Consider, indeed, that, of course, if we apply the equinumerosity of, e.g., either $Z_2$ or $V_2$, then two sets under consideration will show unavoidably "different". But, – even if Benacerraf did not consider alternatives to the identity as determined by the $\in$-relation – it is possible to consider another basic feature of sets, namely the $\subseteq$-relation. The French philosopher, in this context, additionally, wrote:

> I know that $2 \neq 3$ because I know, for example, that 3 is odd and 2 is not, yet it seems clearly wrong to argue that we know that $3 = \{\{\{\varnothing\}\}\}$ because, say, we know that 3 has no (or seventeen, or infinitely many) members, while $\{\{\{\varnothing\}\}\}$ has exactly one. We know no such thing. We do not know that it does[85].

In this respect, Benacerraf – apart not considering maximal chains – seems to be completely misunderstanding how sets were supposed to accomplish the mission of representing, for example, the natural number 3. Here's exactly the point upon which we will stress in the third point of the present subsection. We suspect that Benacerraf – for the sake of his purposes – completely avoided to consider the difference between the possible readings of the "identity", involved in statements of the form $3 = \{\{\{\varnothing\}\}\}$. Let's say that the general form of an identity statement is of the following form:

$$a = b.$$

With respect to the identity that the previous statement expresses, we may consider that the sign = indicates two distinguished features: either, the complete identification

[85]Benacerraf 1965, p. 288.

of the first term $a$ to the second one, $b$, or – more simply – we may introduce the identity $a = b$ by considering that the two terms "faithfully represent" each other. In different words, if $a$ represents $b$ in a way such that $a$ is able to gives us an idea of the main properties of $b$, then their identity – $a = b$ – follows. Anyway, some care is needed and we will return to this discussion soon.
Now, consider that, already, at the beginning of the following work, we have reserved much attention on the division between naïve and axiomatic set theory, since – by developing different frameworks – as clear, also objectives, methodologies and conclusions may be different. This consideration is probably – I think – one of the biggest lacks in Benacerraf's paper. Unlike naïve set theory, indeed, the axiomatic version allows us to prove well-defined properties of orderings of sets, permit us to focus the attention on chains, with the additional features of measuring length, height and rank and gives us a precise idea of why sets are sometimes understood as a foundations for, at least, number theory. Indeed, as I claim, philosophers of mathematics have to consider a (Zermelian) heuristic path, that takes into account the division between "foundational issues" and "pure mathematical ones", so as to avoid mixed and wrong conclusions.

Benacerraf, we remark, indeed, (i) does not explicitly distinguish naïve set theory from its axiomatic version; (ii) apart the $\in$-relation, he does not consider the complex of partial orderings relations to which sets are subject and, (iii) purely philosophically speaking, he does not consider a clear notion of identity and an explicit concept of faithful representation of mathematical objects. We will focus on this exact notion in the third part of this section.

#### 3.3.1.2 Towards a Philosophy of Sets

In this point I'll try to argue against the conclusions that Benacerraf established with respect to the existence of abstract objects. We will reserve a more deep discussion concerning their existence and possible knowledge in the next two chapters. Before trying to introduce them explicitly – as we will do in the fourth chapter –, here, we will just introduce the discussion.
When we remarked that it is important to take into account a clear and precise distinction between naïve and axiomatic set theory, we have done it since, Benacerraf mixed inappropriately the two theories. Indeed, by simple inspection of his paper, it immediately emerges that he criticises, for example, Frege's realism by employing an axiomatic approach to sets. If the distinction between naïve and axiomatic set theory would have been considered, then, Benaceraff would have seen that, while Frege was moved by strict philosophical convictions, Zermelo and von Neumann, for instance, were moved by more mathematical interests. Indeed, the last two mathematicians rarely reserved ontological or epistemological considerations concerning their works, while – oppositely – Frege started exactly his writings inspired in finding the "laws" of rational thought. It should be considered, indeed, that Zermelo and von Neumann gave a precise and powerful framework that could be discussed and applied by

philosophers coming from different schools of thought, it was somehow *neutral* philosophically speaking. In this respect, Frege's work has taken the opposite direction: he systematized his philosophical convictions by applying his concept of "collection of objects" and his newly proposed notation. Hence, any part of Frege's work is connected and linked to the philosophical ideas he wanted to develop and, trying to argue against one of its thesis, by employing a different framework – not explicitly distinguished from Frege's original one – suggests that something important is missing. In order to be clear, it should be always considered, indeed, that – for example – Frege could not use the Kuratowski pairing operation to form von Neumann's set means exactly that for Frege relations were not determined as by contemporary set theorists. Probably, also Frege would have agreed that if $|\{\{\varnothing\}\}| \neq |\{\varnothing, \{\varnothing\}\}|$, then $\{\{\varnothing\}\} \nsim \{\varnothing, \{\varnothing\}\}$. But this does not consequently mean that no other way to determine positively the identity between $\{\{\varnothing\}\}$ and $\{\varnothing, \{\varnothing\}\}$ can be find. Indeed, with respect to the notion of maximal chain (developable in the axiomatic framework), the two sets can be re-declared – if not identical – at least equivalent. For Frege recall that a number is identical to another just in case the object that fall under the first one are in a one-to-one correspondence with the objects that fall under the second one. In our example, Zermelo'sand von Neumann's set can rendered equivalent by their sharing the same maximal chain length:

$$\{\{\varnothing\}\} = \varnothing \subseteq \{\varnothing\} \subseteq \{\{\varnothing\}\},$$

where $\varnothing$ is the first subset in the chain, $\{\varnothing\}$ is the second one and, finally, $\{\{\varnothing\}\}$ is the third one. We compute the length of the chain – namely the number of subsets it includes – as follows:

$$\mathsf{len}(\{\{\varnothing\}\}) = |\{\{\varnothing\}\}| - 1 = 3 - 1 = 2.$$

Likewise,

$$\{\varnothing, \{\varnothing\}\} = \varnothing \subseteq \{\varnothing\} \subseteq \{\varnothing, \{\varnothing\}\},$$

where, as before, $\varnothing$ is the first subset in the chain, $\{\varnothing\}$ is the second one and, finally, $\{\varnothing, \{\varnothing\}\}$ is the third one. We may compute the length of this chain, i.e., the number of sets it includes, as follows:

$$\mathsf{len}(\{\varnothing, \{\varnothing\}\}) = |\{\varnothing, \{\varnothing\}\}| - 1 = 3 - 1 = 2.$$

Of course, it may be objected that things are anyway different from how Frege and logicists wanted to develop them, and, indeed, our aim – up to now – is to show that Benacerraf's philosophical abolition of abstract objects – as based upon his more mathematical discussion of sets – has not the desired strength. With a simple change in Benacerraf's argumentation, equivalence between sets can be re-settled. Anyway, we do not claim that thanks to this equivalence re-settling something as "strong identity" follows. We will develop this last consideration at length within the next point.

#### 3.3.1.3 Dedekindian Reflections on "Mathematical Representations"

We've insisted several times upon the importance of the notion of representation within a mathematical context. Here, we will offer, finally, a more detailed account of what we mean with this notion and how it is helpful in considering Benacerraf's argument. Firstly, we will quote the most appropriate characterization of the notion of "representation" we possess. Now, consider that, for example,:

> [...] the "identification" of the directed line $\Pi$ with the set $\mathbb{R}$ of real numbers, via the correspondence which "identifies" each point $P$ on the line with its coordinate $x(P)$ with respect to a fixed choice of an origin $O$. What is the precise meaning of this "identification"? *Certainly not that point are real numbers*[86].

And, henceforth:

> What we mean by this "identification" of $\Pi$ with $\mathbb{R}$ is that the correspondence $P \longmapsto x(P)$ gives a **faithful representation** of $\Pi$ in $\mathbb{R}$ which allows us to give arithmetic definitions for all the useful geometric notions and to study the mathematical properties of $\Pi$ **as if point were real numbers**[87].

So, in order to be clear, we can summarize the main points of Moschovakis characterization as follows:

1. The identification of two kinds of mathematical objects is subject to their possibility to faithfully represent each other;
2. That $a$ is a faithful representation of $b$ means that:
   (i) $a$ helps me in characterizing in a precise and determined way $b$;
   (ii) always more mathematical properties of $a$ may be discovered.

In this sense, we may reformulate the example concerning $\Pi$ and $\mathbb{R}$. Let's say that the identification of points of $\Pi$ with the real numbers respects the conditions at point (i) and (ii), then it might be concluded that:

(i) $\mathbb{R}$ helps me in characterizing arithmetically a geometric object such as $\Pi$;

(ii) gives us the opportunity of gaining an always deeper understanding concerning the mathematical structure of $\mathbb{R}$ itself.

Finally, the conjunction of this two features together gives us a way in which two divergent mathematical objects – such as $\Pi$ and $\mathbb{R}$ – may be identified. This should, already, warn us that "identity questions", in mathematics, are complex than it may seem and, probably, as subtle as their philosophical counterparts.

Previously, we've argued that this point of view is not new in mathematics, and, actually, that's true. We may trace back its origin to R. Dedekind's work on rational numbers. As said in the paragraph devoted to the German mathematician, Dedekind has had fundamental and brilliant intuitions not only with respect to arithmetic,

[86] Moschovakis 2006, p. 33.
[87] Moschovakis 2006, p. 33.

but also in other and different areas of mathematics. We will give an idea of how Dedekind arrived to his intuition and how it applies to Zermelo and von Neumann's sets. First, consider that a rational number is any number that can be expressed as the quotient or fraction $n/m$ of two integers, a numerator $n$ and a non-zero denominator $m$. Since $n$ may be equal to 1, every integer is hence a rational number, and their set is usually denoted by $\mathbb{Q}$. Furthermore, out from this set, Dedekind constructed the set of real numbers, i.e., of those numbers whose value is that of a continuous quantity that can represent a distance along a line such as $\Pi$. To be clear:

$$\mathbb{N} = \{0, 1, 2, \dots\} \subset \mathbb{Z}\{\dots, -2, -1, 0, 1, 2, \dots\} \subset \mathbb{Q} = \{\dots, \tfrac{1}{2}, \tfrac{-2}{7}, 2.14, \dots\} \subset$$

$$\mathbb{A}_{\mathbb{R}} = \{\dots, \sqrt{2}, \sqrt[6]{7}, \tfrac{1+\sqrt{5}}{9}, \dots\} \subset \mathbb{R} = \{\dots, \pi, e, -2\pi, \dots\}$$

Where, $\mathbb{Z}$ is the set of integers, $\mathbb{Q}$ the set of rational numbers, $\mathbb{A}_{\mathbb{R}}$ is the set of real algebraic ones and finally $\mathbb{R}$, the reals. Consider that the real algebraic numbers and the real ones are both called also irrational numbers, while only the members of $\mathbb{R}$ are characterized as transcendental numbers. The mathematical characterization of these sets does not affect out intent and, indeed, we will briefly present Dedekind's work and the commentary that himself moved with respect to his mathematical strategies, in order to adopt and adapt to our purposes his argumentative schema.

> Arrange the rational numbers in a row or a line in the usual way, increasing from negative to positive as you go from left to right. By a "cut" Dedekind means a separation of this row into two pieces, one on the left, one on the right. The row can be cut in infinitely many different places. Dedekind regards such a split or "cut" in the rationals as being a new kind of number! He shows in a natural way how to add, subtract, multiply, or divide any two cuts (not dividing by zero, of course). In an equally natural way, he defines the relation "less than" for cuts, and the limit of a sequence of cuts. Once these rules of calculation are laid out, the cuts are established as a number system[88].

Hence, consider that the

> [...] basic idea of Dedekind was that a real number $x$ is completely determined by (and hence can be "identified" with) the set
>
> $$(-\infty, x) \cap \mathbb{Q} =_{\text{def}} \{r \in \mathbb{Q} \mid r < x\}$$
>
> of all rationals preceding it, and that the set of the form $(-\infty, x) \cap \mathbb{Q}$ can be characterized by three simple conditions[89].

The "three simple conditions" to which we are referring are encapsulated within the next definition and are exactly the characterization that Dedekind offered us with respect to the reals:

[88]Hersh 1997, p. 274.
[89]Moschovakis 2006, p. 216.

**Definition 31** (Dedekind Cut)**.** A **Dedekind cut** is any set $C$ of rational numbers which satisfies these three conditions:

1. $C \neq \varnothing$, $(\mathbb{Q} \setminus C) \neq \varnothing$;
2. $r < q \land q \in C \longrightarrow r \in C$;
3. $q \in C \longrightarrow (\exists r)\,[q < r \land r \in C]$.

So, we set:

$$\mathscr{D} =_{\text{def}} \Big\{ C \subseteq \mathbb{Q} \mid C \text{ is a Dedekind cut} \Big\}.$$

What actually happens is that

> Every rational number $x$ defines an associated cut. The left piece is simply the set of rational numbers less or equal $x$, and the right piece is the set of rationals greater than $x$. By this association between cuts and rational numbers, we make the rational numbers a subsystem of the system of cuts. To identify Dedekind cuts as the sought-for "real number system", we must show that they include all the rationals and irrationals—all the numbers that can be approximated with arbitrary accuracy by rationals[90].

So, just to understand, using Dedekind's definition we may define $\sqrt{2}$ as follows:

> I must identify a left half-line and right half-line associated with $\sqrt{2}$ . What rationals are less than $\sqrt{2}$? Certainly all the negative ones, and also all those whose squares are less than 2. All numbers $x$ such that either $x < 0$ or $x^2 < 2$. That specifies the left piece of the cut, the left half-line associated to $\sqrt{2}$ . Its complement is the corresponding right half-line. It's easily verified that when this cut is multiplied by itself, it produces the cut identified with the rational number 2. Among Dedekind cuts 2 does have a square root![91]

So, let's say that:

$$A = \{x \in \mathbb{Q} \mid x^2 < 2 \lor x < 0\},$$
$$B = \{x \in \mathbb{Q} \mid x^2 \geq 2 \land x \geq 0\}.$$

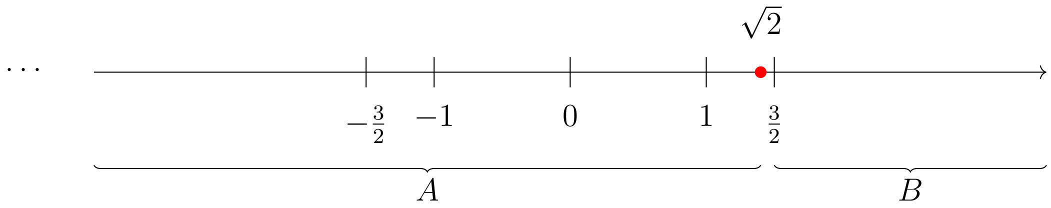
Now we check the following conditions:

1. Both $A$ and $B$ are non-empty;
2. Every member of $A$ is less than every member of $B$; and
3. The set $A$ contains no greatest element. In this case that means that given $a \in A$ there exists $c \in \mathbb{Q}$ such that $a < c$ and $c^2 < 2$.

Additionally, if $B$ contains a least element, the cut corresponds to that rational number. i.e. the rational number is a member of the real numbers. Differently, if $B$

[90] Hersh 1997, p. 274.
[91] Hersh 1997, p. 274.

does not contain a least element, then the cut defines a unique irrational number that "separates" the two sets. In this case the irrational number is $b \notin B$, such that $b^2 = 2$ (or $b = \sqrt{2}$). But, actually, in order to show that $b \notin B$, we must show that $b$ corresponds to an irrational number – in the case of $\sqrt{2}$ we have such a proof. To clarify this example, imagine an infinite right-oriented arrow that contains all rationals in $\mathbb{Q}$:




The number we are trying to insert into the oriented line is precisely $\sqrt{2}$. By inspection, the set $A$ – as defined – collects all rationals $q$, such that $q^2 < 2$, or $q$ is less than 0, i.e., all rationals (including positive and negative integers) whose square is less than 2. The set $B$ is $A$'s complement and, indeed, it collect all rationals, major than 0, whose square is greater (or equal) to 2. By simple inspection, $B$ has no least element – that is, there is no number representable as a ratio between two integers – and, hence, $\sqrt{2} \notin B$. In our oriented line, the space left between the two sets $A$ and $B$, metaphorically identified with the red point, is the cut that represents the irrational number $\sqrt{2}$. More mathematically speaking, the cut, given by the partition $(A, B)$, represents the irrational number $\sqrt{2}$. So, as it should be clear that, representing the reals (rationals and irrationals) within the geometric oriented line means – in Dedekind's work – to fill a gap where there should be a value. His idea has been that of "isolating" those gaps by describing them as "cuts" (i.e., as sets of rationals that share the least upper-bound property, but have no greatest member) of the geometric oriented line.

Anyway, without spending to much time on this elegant, but complex, construction let's ask, if what we are characterising as "cuts", *are* the real numbers themselves: in other words, cuts do just represent – in a mathematically appropriate way– the reals, or, *are* they completely identical to the real numbers? Now, consider that, for what concerns the answer, there's some general agreement and, probably, the first formulation of the commonly accepted answer has to be traced back to Dedekind's work. Let's, hence, focus for a moment on his reflections[92]:

> (Letter to Weber 1888)[...] where you say that the irrational number is nothing other than the cut itself, while I prefer to create something new (different from the cut) that corresponds to the cut and of which I say that it brings forth, creates the cut. We have the right to ascribe such a creative power to ourselves; and moreover, because of the similarity [GleichartigkeitJ of all numbers, it is more expedient to proceed in

[92]We have already quoted the following passages in the section devoted to Dedekind's abstraction. Here our main intent is different.

> this way. The rational numbers also produce cuts, but I would certainly not call the rational number identical with the cut it produces;[93].

So, as we might imagine, Dedekind is saying that the identity between a cut and the rational it should represent, is not more than an "expedient" for mathematical practice and that establishing equality between these two objects would be an error. What is instructive is that, in the letter quoted above, Dedekind is answering to a question that concerns his cuts, by comparing them to what happens within natural numbers when represented with collections:

> I hold the cardinal number (Anzahl) to be only an application of the ordinal number, and in our ["counting"] too one reaches the concept five only via the concept four. But if one were to take your route – and I would strongly urge that it be explored once to the end – then I would advise that by number (Anzahl, cardinal number) one understand not the *class* itself (the system of all finite systems that are similar to each other) but something *new* (corresponding to this class) which the mind creates[94].

What should we learn from this suggestions coming directly from Dedekind? Actually, we should carefully consider that sometimes some peculiar and well-defined mathematical objects (such as sets or the geometric line Π) are useful to "create" new mathematical objects (e.g., the set $\mathbb{N}$ or the reals $\mathbb{R}$). What actually Dedekind meant with creation is doubtful[95], but – it is clear – that, what he had in mind, at least in part, was connected to the work of giving systematizations of all possible branches mathematics. In this context, I believe, Dedekind's deep intuition should be considered: even if some objects are suitable to systematize determined portions of mathematics, this does not mean that the two branches get identified. Indeed, recall that our previous investigation on the concept of "identification" has brought us in characterizing the notion of "faithful representation", where, with this latter, we are meaning exactly what Dedekind – already at the end of the $20^{th}$ century – theorized. So, the Dedekind-style suggestion we were searching for can be put as follows: whenever two distinct mathematical objects are helpful in characterizing each the fundamental properties of the other (and viceversa), then their identification has just to be considered as a"representative" expedient. As it should already emerge, hence, the ontological argument, which tries to ascribe identity to sets and numbers, even if to negate it, is based upon a misunderstanding on the mechanism for which sets and numbers are supposed to represent each other. In this sense, we claim that Benacerraf – apart not considering the different versions of set theory, with the consequences we've exposed above –, did not, additionally, care about the notion of mathematical representation and, hence, cannot see the importance of, for instance, the set theoretical study of natural numbers. Generally speaking, if we are trying to do philosophy of mathematics, the "data" that we researcher has at our disposal, are

[93]Dedekind 1888a, p. 845.
[94]Dedekind 1888a, p. 835.
[95]See our philosophical valuation of Dedekind's notion of abstraction.

those coming directly from the successes and failures of the works of mathematicians. In this context, if we get rid of some parts of their works and did not care about their results – just to arrive to our desired philosophical conclusions – then, we did not render a good service to both disciplines, mathematics and its philosophy. Hence, in conclusion, it might be said that Benaceraff, with his clear ontological objective, disregarded and wrongly simplifies some fundamental features of the objects which he was using as examples. If sets are "good representatives" of natural numbers and if cuts are the same for real numbers, then this does not imply their reciprocal "ontological reducibility". If Benacerraf would have considered that in no case sets *are* natural numbers and cuts *are* reals – because of their just "good representatives" each one of the other –, then probably he could not have drawn his philosophical conclusions.

By criticizing Benacerraf's paper we've found out three main problems with respect to the analysis of set theory he establishes and that – according to us – undermine his ontological conclusion:

1. The superficiality in distinguishing naïve and axiomatic set theory and, consequently, the inability in seeing the divergences with respect to their frameworks, objectives, methodologies and, especially, motivations.

2. Implicitly admits that Zermelo's and von Neumann's sets would be "identical" just in case their equinumerosities were equal. He does not consider any further possibility for letting the two set theoretical constructions be, at least, equivalent.

3. Did not consider an identity relation, as based upon the notion of faithful mathematical representation. For him mathematical identity propositions can be identified with ontological identity statements, but, thinks – from an ontological point of view – are more general. So, for example, asking if abstract objects have identity conditions means asking something of *whatever* object in which we may be interested in and, hence, in any entity which respects determined, let's say, ontological conditions (e.g, being not spatio-temporal located). Differently, by trying to argue if natural numbers are representable by sets we are not worried about existential or epistemological problems, but we have mathematical and maybe logical concerns with our theories. But, reflect that, also sets and numbers are within the range of abstract objects, falling, thus, under the investigations of ontology, but – as clear – the systematization we are searching for, in this case, is rather different from the one mathematicians could be interested in. In conclusion, as repeated severally, it should not be the case to draw conclusions within the philosophy of a determined science, such as mathematics, by not considering the complexity of the science itself and its different and philosophically independent exigences.

# Part II
# Individuating abstract objects

# Chapter 4

# Mathematics by Abstraction

**Overview.** In the present section our aim is to discuss from a closer point of view some aspects of abstraction principles. In particular, recall that, by trying to undermine Benacerraf's logical and mathematical premises (Chapters 2-3), as said several times, our purpose was that of reconsidering his final philosophical observations, i.e., his conclusion for which no abstract objects are needed for a philosophical development of mathematics. Roughly, from its origins, Frege's logicism has brought the attention on the possibility of (i) reducing arithmetic to logic and (ii) of faithfully introducing logical objects (truth values, numbers, course-of-values, . . . ) as referents for the statements in which those objects themselves figure. As remarked, Frege's attempt (i) failed for his system was inconsistency entailing and the main fault was to be tracked back to one of his principles, i.e., Basic Law V which was meant to govern "extensions". Nevertheless,– as sketched at the end of Chapter 2 – the story of Fregean attempts towards the philosophy of mathematics did not end here and, indeed, since the early 60s, the philosophical debate has envisaged the possibility of rendering justice to Frege's original project, in particular thanks to the usage of abstraction principles different from BLV. In this context, in what follows, we will develop a formal model which, we think, will help us in speaking *about* some instances of abstraction principles. So, rather than to *use* specific and determined abstraction principles we will, in some sense to be specified at length, evaluate when their usage is "allowed". Therefore, the main territory of investigation of the present chapter deals with some semantic, truth-theoretic aspects of Abstractionism. For what concerns our ontological and epistemological interpretation of the model we will discuss soon and, more generally, for a philosophical evaluation of the whole of the work we've done up to now, consider our arguments in Chapter 5.

## 4.1 Three Problems for Abstraction Principles

Abstractionism is the philosophy of mathematics which is engaged with the arrangement and usage of abstraction principles for foundational purposes[1]. In this spirit, it is useful to point out that we will consider abstraction principles of the following form:

$$\S\alpha = \S\beta \longleftrightarrow R(\alpha, \beta) \tag{$\Sigma$}$$

As it is easy to check, $\Sigma$ is a biconditional that contains an equivalence relation $R$ on its right-hand side (RHS) and an identity statement on its left-hand side (LHS). It is to understand, that the *abstracta*, standing into the identity relation, are introduced thanks to the abstraction principle, given the equivalence relation. This means that the equivalence relation, in the RHS, of $\Sigma$ has *conceptual priority*, with respect to the objects that are abstracted from the equivalence relation itself, and which are to be considered as the *resulting items* of the abstraction operation. By saying to we will conduct a semantic inquiry, we mean that we will consider whether «our capacity to have singular thoughts about objects of certain type derives from and is constituted by an appreciation of the truth-conditions of identity judgements about objects of that type»[2] . In this context, we will sketch how – given an equivalence relation $R$ on a RHS – it is possibly to give truth-conditions of identity judgements on a LHS, by means of an abstraction principle. What we ideally should get, after an abstraction operation, is a *new* concept which grasps our comprehension of particular abstract objects. So, for example, in HP case, the equinumerosity relation, from which we start abstracting, should concede us to apprehend and grasp our understanding of the term "number".
Anyway, consider that abstraction principles, also in contemporary debates, are threatened by problems and objections and, indeed, in the following three paragraphs we will deepen some of them, by, consequently discussing how our model theoretical considerations might be helpful.

### 4.1.1 Symmetry/Asymmetry

By analysing the two sides of any abstraction principle expressed by biconditional, we see that the relation between the Left-Hand-Side (LHS) and Right-Hand-Side (RHS) is not that explicit[3]. This means that we are not told whether the objects postulated in the LHS are, in some sense to be explicated, *presupposed* by the entities involved in the RHS. In more technical terms, it could be asked whether *impredicative* abstraction principles – the LHS's objects are *presupposed* by the RHS's entities – are

[1] This chapter should be considered as a "work in progress" since much of the literature concerning semantic, mathematical and philosophical abstractionism has not been considered yet. Anyway, for the purposes of this present work, the literature considered here can be regarded as sufficient.

[2] Ebert and Rossberg 2017, p. 5.

[3] This paragraph owes much to Linnebo 2017, pp. 247–268.

more suitable than *predicative* ones. In order to have a clear idea of the distinction between predicativity and impredicativity, let's consider the following definition[4]:

**Definition 32** (Characterization of Abstraction Principles)**. (a)**. An **axiomatic abstraction principle** is a law of the following form:

$$\S(F) = \S(G) \longleftrightarrow \Phi[F, G], \tag{$\Sigma$ A}$$

where:

1a. § is the so-called **abstraction operator**, which takes second-order variables to first-order terms;

2a. $\Phi$ expresses the equivalence relation under consideration.

A **schematic abstraction principle** is a law of the following form:

$$\S x.\varphi(x) = \S x.\psi(x) \longleftrightarrow \Phi[{}^{\varphi}/_{F}, {}^{\psi}/_{G}], \tag{$\Sigma$ S}$$

where:

1b. § is the so-called **variable-binding** abstraction operator, that takes open formulas to first-order terms;

2b. $\Phi[{}^{\varphi}/_{F}, {}^{\psi}/_{G}]$ expresses the result of substituting, in $\Phi$, simultaneously $\varphi(t)$ and $\psi(t)$ for any occurrence of $F(t)$ and $G(t)$, respectively, with $t$ first-order term.

**(b)**. An abstraction principle is **impredicative** if the terms on its LHS refer to objects which would be included in the range of some quantifier occurring on its RHS. Otherwise, it is **predicative**.

We will return to the distinction between axiomatic or schematic principles very soon. Anyway, by considering our characterisation of "impredicativity", one requirement for abstraction principles could considered that of predicativity, in the sense that the RHS of an abstraction principles does not quantify over the sort of objects to which its LHS refers. Indeed, it may rightly asked: is it actually a good move to banish impredicative abstraction principles, in virtue of the fact that there is a sort of presupposition between the two sides of the biconditional? In order to answer this question, let's focus on the following argument[5]:

(P1) **Asymmetry**
The right-hand side of an abstraction principle must not presuppose any of the objects to which the left-hand side refers.

(P2) **Quantification incurs presupposition**
A quantified statement presupposes every object in the range of its quantifiers.

---

[4] We've slightly modified Linnebo's definition, but its main features remain always the same. Compare, Linnebo 2017, pp. 249–250.

[5] See Linnebo 2017, pp. 258–259.

(C) **Conclusion**
Impredicative abstraction principles are unacceptable.

(P1) is the claim that contains the "impredicative thesis" concerning abstraction principles, while (P2) states in terms of presupposition the "logical" behaviour and understanding of a quantified statement. If (P2) is applied to (P1), the latter can be reformulated by saying:

(P1*) **Asymmetry**
The quantification over the right-hand side of an abstraction principle must not involve any of the objects that the left-hand side "newly" introduces.

In such a spirit, the conclusion of the argument, (C), by unifying (P1) and (P2) (or (P1*) and (P2)), establishes the untenability of the "impredicative thesis" concerning abstraction principles. The "predicate" version of those principles, differently, escapes the negative argument we've just exposed and has very interesting features and consequences. By focusing upon a concrete example it should become intuitively clear why a "predicative" reading of such principles has been thought of as more suitable. Recall the abstraction principle governing directions:

$$d(\ell_1) = d(\ell_2) \longleftrightarrow \ell_1 \,||\, \ell_2. \tag{Dir}$$

According to our definition, (Dir) belongs clearly to those principles which we will call predicative. The LHS should contain entirely new objects, namely entities that were not available before their introduction, given the abstraction principle. Indeed, within (Dir), the quantifier ranges just over lines ($\ell_1$ and $\ell_2$) but not over directions, which, therefore, have to be understood as not presupposed by the RHS. In this sense, by having, for instance, the parallelism relation between two lines, we can introduce new concepts, namely the directions of those lines. Our new objects, $d(\ell_1), d(\ell_2)$, have not been derived from the objects that were available at the stage before, and this means that the new objects, before their introduction, were not *presupposed* by something already existing, namely $\ell_1 \,||\, \ell_2$ (the parallelism between the two lines). Differently, for example, two other laws – of Fregean reminiscence – are impredicative abstraction principles. Let's consider them closely:

$$\epsilon F = \epsilon G \longleftrightarrow \forall x(F(x) \equiv G(x)); \tag{BLV}$$

$$\#F = \#G \longleftrightarrow F \sim G. \tag{HP}$$

As clear, both RHSs quantify over some objects whose identity conditions are specified by their LHSs, being therefore impredicative. Differently from before, the "impredicative" abstractionists – without rejecting the notion of "presupposition" – try to understand the abstraction process by two fundamental means. The first point can be also common to most sustainer of the "impredicativity thesis", while the second is directly connected to the acceptance and assessment of the relation of presupposition between LHSs and RHSs. Schematically:

1. Abstraction can be understood as a well-behaved process by which more and more *abstracta* are "introduced".
2. Each stage of the process "presupposes" only what was available at the foregoing stage.

In what follows, the instances of abstraction principles we want to speak *about*, through the construction of our model, will consider exactly (i) impredicative abstraction principles and (ii) a strategy that encapsulates the notion of *presupposition* semantically.
Before turning to that point, consider that, abstractionism faces at least other two puzzling problems, that we will roughly introduce in order to give an idea of which concerns have brought us in considering and constructing the model we will develop soon. Furthermore, the distinction between axiomatic and schematic principles – just sketched – is of prime importance for our work and, indeed, will be adequately covered while introducing the first formal details of our model.

### 4.1.2 "Bad Company" or "Embarrassment of riches" problem

A question that abstractionist proponents are requested to answer is the following: is there a way to distinguish good *vs* bas, acceptable *vs* unacceptable, abstraction principles?[6] For some philosophers the untenability of abstractionist accounts of mathematics is strictly to be linked to the presence of "too many" principles, in the sense that, some of them are, for instance, individually consistent, correct or philosophically informative, but, – when compared to other abstraction principles – they turn out as inconsistent, incorrect or philosophically uninformative. In other words:

> The bad company problem is one of the most serious problems facing one of the most exciting philosophical approaches to mathematics. [...] The problem is that not every abstraction principle is acceptable: some are downright inconsistent, while others are unacceptable for more subtle reasons. Abstraction principles with desirable philosophical and technical properties are thus surrounded by "bad companions". Some philosophers claim that this gives rise to a devastating objection to the neo-Fregean programme, while others respond that the challenge it poses is perfectly surmountable[7].

Again, differently put:

> The embarrassment of riches objection is that there is a plurality of consistent but pairwise inconsistent abstraction principles, thus not all consistent abstractions can be true[8].

[6] For introductory remarks see Linnebo 2009b, pp. 321–329. For an attempt that tries directly to confront with the "Bad Company" problem see, among others, Linnebo 2009a, pp. 371–391. Both papers, in different respects, have inspired our research.

[7] Linnebo 2009b, p. 321.

[8] Weir 2003, p. 13.

Now, we draw our solution, as said, by constructing the model we have in mind, especially by giving a semantic characterization of what we mean with the notion of "well-founded process". In this sense, – I announce it preventively –, the strategy we have elaborated, finally, will suggest to "restrict" abstraction principles in a specific and well-determined way – thus, furnishing a "criteria", different from "acceptability", to evaluate the behaviour of abstraction principles. What we will discover, with respect to the "Bad company" problem, is that – according to our presentation – the dichotomy good *vs* bad abstraction principles is misguiding, so that, at the end, we will have no «good guys left, no bad company either»[9].

### 4.1.3 The "Julius Caesar" Problem

In *Grundlagen*, §55, Frege himself showed to be aware of the so-called "Julius Caesar Problem". Referring to HP, in particular, we are able to recognize that two concepts are identical by the fact that they share the same number of objects, but we are not told *what* exactly the "numbers" themselves *are*. In other terms, Frege is asking himself what does it mean for a logical object to be a "number" and, more generally, whether this answer can be given directly by HP. His answer within the *Grundlagen*, is negative and, indeed, Frege rejected HP as a definition for numbers and its usefulness as axiom in his logicist project. Consider his words:

> [...] we can never – to take a crude example – decide by means of our definitions whether any concept has the the number Julius Caesar belonging to it, or whether that same familiar conqueror of Gaul is a number or is not. Moreover we cannot 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 find it impossible in general to prove a numerical identity, since we should be quite unable to achieve a determinate number. It is an illusion that we have defined 0 ans 1; in reality we have only fixed the sense of the phrases
>
> "the number 0 belongs to"
> "the number 1 belongs to";
>
> but we have no authority to pick out the 0 and 1 as selfsubsistent objects that can be recognised as the same again[10].

In other terms, Frege is posing a very intricate question. Consider that statements such as:

"the number of $F$ is identical with $x$",

that is, formally,

$$\#F = x,$$

cannot be asserted since we are not given a definition – by HP – of the concept of "number" and, hence, this statement is neither true nor false, unless $x$ is stated in

[9]Leitgeb 2017, p. 270.
[10]Frege 1884, p. 68.

the form "the number of $G$". So, Frege concluded, that even something absurd such as

$$\#F = \text{Julius Caesar}$$

could be concluded. This Fregean example motivates the label "Julius Caesar" problem. Recall, that – starting from the 60s – some philosophers argued that Frege's logicism could be saved by appealing to HP as an axiom and by rejecting any application of BLV during the proofs. The problem is that, originally, Frege rejected HP as a definition and characterization of the concept of "number", since it was subject to the aforementioned problem. The "Julius Caesar" problem, clearly, poses not only a definitional limitation to HP, but also an epistemic one, since the only characters we know – given HP – are the identity conditions for two logical objects labelled numbers, but not further defined. In this spirit, moreover, Frege noticed that, if HP would have been considered as a definition of the concept "number", then, consequently, his famous *context principle* (CP)[11] would have been violated. According to CP, if our knowledge of logical objects, such as numbers, derives from the analysis of the content of those sentences, in which these objects figure as referents, then we should be able to recognize – through a general criterion – when some object $x$ is the logical object we are speaking about. But, given the "Caesar" problem, according to Frege, this does not seem to be given by HP again.
Contemporary debates, instead, by focusing the attention on the usefulness of HP as an axiom, argued in favour of its epistemic features in different manners, by trying to answer the "Caesar" problem. Now, putting aside HP, it is possible to consider any abstraction principle, such as $\Sigma$, and ask whether there is some way to give definite truth-conditions of mixed statements of the form:

$$\S\alpha = t, \text{ where } t \neq \S\beta.$$

So, finally, as clear, the "Julius Caesar" problem is intricate and ambitious to solve and, indeed, we believe that some suggestions – as developed in the last part of Chapter 5 – can just indicate the direction that we wish to take. Anyway, since our model describes exactly "grounded" passages to settle truth-conditions for identity statements, we have to primarily consider its development and, just in a second moment, its philosophical, in this case epistemological, interpretation. As said, given its high ambitions, there wont be a complete ans satisfactory answer to the "Julius Caesar" problem, but we think that the suggestions we will give could be expanded and deepened in an interesting way.

[11]Frege writes: «How, then, are number to be given to us, if we cannot have any ideas or intuitions of them? Since it is only in the context of a proposition that words have any meaning, our problem becomes this: to define the sense of a proposition in which a number word occurs. That, obviously, leaves us still a very wide choice. But we have already settled that number words are to be understood as standing for selfsubsistent objects. [...] 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 it proper name», Frege 1884, p. 73.

## 4.2 An Interesting Response: Groundedness and Abstraction

One way to answer, for instance, the "Bad Companion Objection" against abstraction principles, is to claim that the distinction acceptable vs unacceptable, good vs bad, abstraction principles is a mistake from the start.[12] Let's begin our reflections by considering that in the history of philosophy, two main forms of abstraction has been developed: Plato's theory of ideal forms and Aristotle's conception of forms as contained within sensory objects. The first one held that our whole world is just an imperfect "shadow" of the realm consisting in ideas or *abstracta.* The latter realm is, indeed, ontologically and epistemologically independent from our world. Aristotle, differently, held that abstract forms are just *extrapolate* from sensory objects. Indeed, by a mental process we are able to recognize the fundamental characters present in different objects and isolate them. Thus forms are existent just within the objects of our experience and, moreover, their existence is ontologically and epistemologically dependent on the latter. As it should clear, the two positions – Plato's and Aristotle's – are in contrast one with the other, and it may be correctly asked if there is a sort of midway. If there is, then some elements of the Platonic conception should be maintained, while other should be rejected and substituted with concepts deriving from the Aristotelian tradition; and viceversa.

> A standpoint intermediate between Plato and Aristotle can be adopted. One can hold with Aristotle against Plato that the existence of an abstraction somehow depends on the given, and at the same time hold with Plato and against Aristotle that an abstraction does not exist *in* the objects from which it is abstracted. In other words, abstract objects are not multiply instantiated – they do not exist in space and time. But their existence depends on the objects from which they are abstracted nonetheless. This is sometimes called a "light" conception of abstract objects[13].

We will focus on the "light" conception of existence at the end of the section. Up to now, we are seeking for an "abstraction process" for which (i) the abstractum depends on the objects from which it is abstracted and (ii) the abstractum does not exist before its introduction within the objects from which it is extrapolated. In order to be clearer, let's focus upon a concrete example:

$$\epsilon F = \epsilon G \longleftrightarrow \forall x(F(x) \equiv G(x))$$

By considering our mixed Platonic and Aristotelian ideas, it could be said that (i) the extensions on the LHS depend upon the equivalence between the object falling under $F$, and $G$, as expressed in RHS, while (ii) the existence of the extensions themselves is not contained in what RHS expresses. This latter point, consequently, means that what we encounter in a LHS of an abstraction principles is a new object,

[12]See especially Horsten and Leitgeb 2009 and Leitgeb 2017.

[13]Horsten and Leitgeb 2009, pp. 217–218.

that is an object simply construed from the objects already existing. In this spirit, extensions are not in the equivalence between the two courses of values, but instead they are something that is introduced, starting from the equivalence of course of values, at a more advanced stage. Moreover, a methodology of abstraction is very useful in both mathematics and philosophy (thanks to the contributions of Frege and Carnap) since it allows us to introduce new abstract objects in a very fruitful way: by considering "equivalence classes"[14]. By reconsidering the abstraction principle for directions and by taking equivalence classes explicitly, it could be read as follows: "$G$ is a collection of straight lines; $R$ is the relation of parallelism; $A$ is a collection of directions". Thus, $A$ is the new collection composed by each equivalence class determined by the equivalence relation $R$ on $G$. In other words, the members of $A$ – the directions – are considered as abstracted from the elements in $G$ – the straight lines – through the equivalence relation $R$, in this case the relation of parallelism. Other two famous examples we will consider soon are Frege's celebrated BLV and HP:

$$\epsilon F = \epsilon G \longleftrightarrow \forall x (F(x) \equiv G(x)) \tag{BLV}$$

$$\#F = \#G \longleftrightarrow F \sim G \tag{HP}$$

As it should already be evident, BLV and HP regulate the identities and differences of some presented abstract objects (numbers or extensions in this case).Note, moreover, that the abstracta in the LHS of the biconditional are settled once the identities and differences between the objects in the RHS are settled. Hence, for instance, once the sameness of cardinality between $F$ and $G$ has been established, then the identity of the number of the objects falling under both, $F$ and $G$, can be established. With the same reasoning, it could be said that, if the cardinality of $F$ and $G$ is different, then, likewise, the two numbers counting the objects of $F$ and $G$ are different. In this sense, the identity or difference between some presented abstracta involve the identity or difference between two other presented abstracta. Hence, the asymmetry/symmetry problem raises again: «*can we come up with a general method for generating abstracta when the equivalence relation itself involves the abstract entity already*?» How it is possible to answer to the circularity worry? By considering Boolos' description of Frege's abstraction, our starting point will emerge:

[14]Recall that in the previous chapters we were dealing with some aspects regarding equivalence classes of equipotent sets in order to consider Benacerraf's argument. It emerged somehow that the arguments against the identification between natural numbers and sets, invoked by the French philosopher, does not hold since it did not consider the identity notion but the simpler and innocent concept of "faithful mathematical representation". Moreover, Benacerraf directed his attack to abstract mathematical objects in general by saying that their truth and identity conditions can never be established correctly. This chapter aims to show that (i) there are ways to introduce abstract mathematical objects and (ii) that there is the possibility to establish clear identity conditions for abstracta.

> For how does Frege show that the number 0 is not identical with the number 1? Frege defines the number 0 as the number belonging to the concept *not identical with itself*. He then defines 1 as the number belonging to the concept *identical with 0*. Since no object falls under the latter, the two concepts are, by logic, not equinumeorus, and hence their numbers are, by Hume's Principle, not identical. [...] 2 arises in like manner: Now that 0 and 1 have been defined and shown different, from the concept *identical with 0 or 1*, take its number, call it 2, and observe that the new concept is coextensive with neither of these concepts *because the distinct objects 0 and 1 fall under it*. Conclude by Hume's Principle that 2 is distinct from both 0 and 1.[15]

Therefore, the identities (or differences) of some abstract objects depend upon the identities (or differences) of some other abstracta. The process of forming numbers, for instance the concept number 2, – according to Boolos – starts by setting all the identities or differences between what we want to introduce and what comes "before". Hence, the introduction of a new object, namely 2, depends upon the identities (or differences) settled between 0 and 1 at earlier stages. Since, the concept of *not identical with itself* and *identical with 0* are not equinumerous[16], then by HP, the numbers belonging to both concepts are different. Formally,

$$[x \mid x \neq x] \nsim [x \mid x = 0] \longrightarrow \#[x \mid x \neq x] \neq \#[x \mid x = 0]$$

Hence,

$$[x \mid x \neq x] \nsim [x \mid x = 0] \longrightarrow 0 \neq 1$$

Likewise for number 2: since *identical either with 0 or 1* is not equinumerous with the concepts *not identical with itself* and *identical with 0*, then its number will likewise be different from both, 0 and 1. Consider that this process describes an abstraction process which proceeds at different stages and may be preliminarily represented as follows:

$$\begin{array}{c}
\cdots \\
\Uparrow \\
\neq \\
\#[x \mid x = 0 \vee x = 1 \vee x = 2] =_{\text{def}} 3 \\
\Uparrow \\
\neq \\
\#[x \mid x = 0 \vee x = 1] =_{\text{def}} 2 \\
\Uparrow \\
\neq \\
\#[x \mid x = 0] =_{\text{def}} 1 \\
\Uparrow
\end{array}$$

[15] Boolos 1990, p. 272.

[16] We denote "non equinumeorusity" between two concepts as $F \nsim G$ – which is equivalent to $\neg(F \sim G)$.

$$\neq$$
$$\#[x \mid x \neq x] =_{\text{def}} 0$$

Therefore, it may be concluded that «there is an ontological dependence of identities and differences between concepted abstracta on identities and differences between other concepted abstracta. Especially the differences are important»[17]. Moreover, it emerges that the identities or differences for abstracta in LHS are not established at the same time when the equivalence relation for the objects in RHS is settled. Indeed, – as we will be shown soon –, the identities or differences of the abstract objects in the RHS are settled at an earlier stage than the identities or differences of the abstracta in the LHS. In this way, indeed, abstracta are not something already contained in the objects in the RHS and merely extrapolate, but, instead, abstracta are objects, whose identity conditions simply depend on the identity conditions of other abstract objects. In other words, by focusing our attention for a moment on a concrete example, the identity $\#F = \#G$ depends on the equivalence relation $F \sim G$. In particular, setting the equivalence relation $F \sim G$ gives us the conditions to state and introduce correctly, at a following stage, the identity $\#F = \#G$. As a crude example, take $A = \{a, b, c\}$ and $B = \{d, e, f\}$. Notice that

$$A = \{a, b, c\} \sim B = \{d, e, f\}$$

and conclude, by HP, that

$$\#A = \#B$$

As it should already be clear from the example, the equivalence relation between the two concepts, $\sim$, settled at the first passage, permits the introduction of the identity between the numbers of the two concepts. Likewise, but differently, take $C = \{g, h, i\}$ and $D = \{l, m\}$ and say that

$$C = \{g, h, i\} \nsim D = \{l, m\}$$

Apply HP and conclude

$$\#C \neq \#D.$$

Even in this context, the difference established by $\nsim$ between two concepts is responsible to the corresponding difference between the numbers of the concepts previously considered. In other terms, $\#C \neq \#D$ depends on $C \nsim D$ (in a way to be specified).

In order to clarify our ideas and to move from simple examples to concrete developments and applications of these preliminary remarks, it is better to get a look to some formal features we are going to employ and discuss.

[17]Horsten and Leitgeb 2009, p. 221.

# 4.3 Dependence and Supervenience

## 4.3.1 Introductory Remarks

First of all, in order to establish what we mean by claiming that there is "dependence" between identities (or differences) of some abstract objects and identities (or differences) of other presented abstracta, we need to specify the general framework thanks to which we are going to analyse abstraction processes. Let's deepen our starting point just a bit:

> The identities and differences between some presented abstracta do not depend on identities and differences between presented abstracta *at all*. These will be our Archimedean starting point. But many identities and differences will only be determined once certain other identities and differences are settled. Thus the identity and difference conditions of presented abstracta can depend on other identities and differences. Identities and differences are no long settled in one go, but are determined in stages. At some point, "settling process" gives out.
> The objecthood of an abstractum presupposes that the abstractum has been given determinate identity conditions In Quinean terms: *no entity without identity!* Thus the objecthood of abstracta can depend on the objecthood of other abstracta. (This is of course a thoroughly un-Aristotelian idea.)
> If at the end of the process all identities and differences have been settled, then all presentations present abstract objects with an associate determinate identity relation[18].

The formal framework was, in any case, implicit in Leitgeb 2005[19]. Indeed, the author suggests to extend and adapt the considerations of his previous article within the context of abstraction processes. This is exactly what we are going to do in the next sections. Anyway, before turning to the expansion of Leitgeb's proposal, we have to notice that in contemporary philosophical literature, the notions of "grounding", "dependence" are overloaded of meaning and, thus, in order to explain what is meant for facts to *depend* each on the other and that something *grounds* something else, we introduce the additional notion of "supervenience". This latter can be generally understood as the American philosopher D. Lewis defined it:

> To say that so-and-so supervenes on such-and-such is to say that there can be no difference in respect of so-and-so without difference in respect of such-and-such. Beauty of statues supervenes on their shape, size, and colour, for instance, if no two statures, in the same or different worlds, ever differ in beauty without also differing in shape or size or colour.[20]

More close to our purposes, Leitbeg gives a characterization of the notion of "su-

[18] Horsten and Leitgeb 2009, p. 221.

[19] Compare also Kripke 1975, Yablo 1982 and Leitgeb 2007.

[20] Lewis 1983, p. 358. D. Lewis (1941-2001) has been one of the most famous philosophers of the $20^{th}$ century and, indeed, many of his intuitions and works are, still nowadays, a great source of inspiration. His main contributions, among others, are within modal metaphysics, counterfactuals semantics and mereology.

pervenience", clearly linked to that of Lewis, and which is of great importance for our understanding of the formal work we are going to settle. Let $\mathscr{L}_{Tr}$ be the formal language to which we add the truth-predicate $Tr(\varphi)$, which indicates that $\varphi$ is true. Let's say that $\varphi$ is a statement and $\Phi$ is a set that contains $\varphi$ (and maybe some other statements). Now,

> Our notion of dependence m ay be circumscribed in the following ways: if $\varphi$ is a sentence of $\mathscr{L}_{Tr}$, and $\Phi$ is a subset of $\mathscr{L}_{Tr}$, $\varphi$ depends on $\Phi$ if and only if the truth value of $\varphi$ depends on the presence or absence of the sentences that are in $\Phi$ in/from the extension of the truth predicate; [...] These formulations show that the notion of dependence which we aim at is a kind of *supervenience*: the truth value of $\varphi$ supervenes on which members of $\Phi$ are to be found in the extension of *Tr*. [...] $\varphi$ does not really depend on anything "outside" of $\Phi$ then.[21].

So, starting from this kind of context, we aim also to study the behaviour of abstraction principles in similar terms. Indeed, what we should achieve is a first characterization of dependence between RHS and LHS of any abstraction principle, so that – roughly, for $X \subseteq \mathscr{L}$ – the identity of some individuals $x, y \in X$ will supervene on the identity of any other pair of individuals in $X$. In other terms, always given a set $X \subseteq \mathscr{L}$, any identity, between a pair of individuals $x, y \in X$, depends on any other identity between any other pair of individuals, always in $X$ – as determined by the abstraction principle considered. At the end, some or all identity/difference instances of a given abstraction principle will turn out – in a sense to be explicated – grounded in some other identity/difference instances. Anyway, to do this job precisely, let's move from this sparse and informal intuitions, to our formal and model-theoretical construction.

## 4.3.2 Formal Framework

### 4.3.2.1 Logical Preliminaries

First of all, in order to consider how it would be possible to "ground" abstraction principles, let's establish the following vocabulary[22]:

- variables letters[23]: $x_n, y_n, z_n, \ldots, x'_n, y'_n, z'_n, \ldots$
- function-variable letters: $f_i^n, g_i^n, h_i^n, \ldots$
- relation-variable letters: $F_i^n, G_i^n, H_i^n, \ldots$
- the identity sign: $=$
- unary or one-place sentential connective: $\neg$

[21] Leitgeb 2005, p. 160.

[22] For the present formal work consider Horsten and Leitgeb 2009 and Leitgeb 2017 as our primary resources. Other useful papers have helped us in developing our model-theoretical considerations, especially, Horsten and Linnebo 2016, Linnebo 2017 and Linnebo 2018. For what concerns our (semantic) conception of "groundedness" as applied in context different from abstractionism, see, in particular, Leitgeb 2005 and Leitgeb 2007. Additionally, some useful sources have been found in Linnebo 2004 and Heck 2011.

[23] Sub- and superscripts are allowed.

- binary or two-place sentential connectives: $\wedge, \vee, \rightarrow$
- quantifier: $\forall$
- auxiliary symbols: (, ), ,

We now offer a recursive definition of the syntax of the language:

**Definition 33.** The following, and nothing else, is a first-order $\mathscr{L}$-term:

- any variable from $\mathscr{L}$.

The following, and nothing else, are first-order $\mathscr{L}$-formulas:

- $t_1 = t_2$, for any $\mathscr{L}$-terms $t_1, t_2$.
- $f^n(t_1, \ldots, t_n)$, for any $\mathscr{L}$-terms $t_1, \ldots, t_n$ and a $n$-place function-variable $f^n$.
- $F^n(t_1, \ldots, t_n)$, for any $\mathscr{L}$-terms $t_1, \ldots, t_n$ and a $n$-place relation-variable $F^n$.
- $\neg\varphi$, for any $\mathscr{L}$-formula $\varphi$.
- $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \rightarrow \psi$, for any $\mathscr{L}$-formulas $\varphi, \psi$.
- $\forall F^n \varphi$, for any $n$-place relation-variable $F^n$ and for any $\mathscr{L}$-formula $\varphi$ which contains no instance of $\forall F^n$.

**Remark.** As always, recall that $p \longleftrightarrow q =_{\text{def}} (p \rightarrow q) \wedge (q \rightarrow p)$ and $\exists\varphi =_{\text{def}} \neg\forall\neg\varphi$.

We begin by building a $\mathscr{L}$-structure, $\mathcal{M}$[24]:

**Definition 34.** A Henkin $\mathscr{L}$-structure is a triple $\mathcal{M} = \langle D^1, D^2, \mathit{Fun}, \mathcal{N} \rangle$, with $D^2 \subseteq \wp(D^1)$. Let:

- $D^1$ be the first-order domain *Dom* of *abstract (pre)-objects*.
- $D^2$ be the subset $\mathit{Con} \subseteq \wp(\mathit{Dom})$ of *concepts* on the pre-objects.
- *Fun* is a set of $n-$place functions.
- $\mathcal{N} : \mathit{Con} \rightarrow \mathit{Dom}$ be the *abstraction map*.

Some specifications on our $\mathcal{M}$ may be useful:

- The non-empty domain *Dom* is regarded as the set of abstract **pre-objects**. We are speaking of *pre*-objects (or **labels**) since what we actually have in our domain are all the objects that – once our individuation process has ended – may "become" **proper abstract objects**. In order to distinguish, if $x \in \mathit{Dom}$ is a pre-object, then, we may rewrite the variable standing for the same, but this time proper, abstract object $\mathsf{x} \in \mathit{Dom}$.
- $\mathit{Con} \subseteq \wp(\mathit{Dom})$ is the set of **concepts** from which we abstract. It is our second

[24]See Button and Walsh 2011, pp. 24–26.

order domain composed by $\mathscr{L}$-first-order formulas $\varphi(x)$ with an appropriate number of free-variables. This means that if $F \in Con$, then $F = \{(x_1, \ldots, x_n) \in Dom \mid \mathcal{M} \vDash \varphi(x_1, \ldots, x_n)\}$.

- First of all, $\mathcal{N}$ is an injective map, which is not necessary onto. It applies to concepts and outputs one of the pre-objects of $Dom$: for some $F \in Con$, $\mathcal{N}(F) = x$. So, generally, if $F =_{\text{def}} \varphi(x)$, and $\mathcal{N}(F) = y$ is the abstraction map of $F$, then $\mathcal{N} : \varphi(x) \mapsto y$. Once, the mapping has output one of the pre-objects or labels $y \in Dom$, the proper abstract object will be, for clarity, rewritten as $\mathsf{y} \in Dom$. In other terms, $\mathcal{N}$ interprets §, namely the **abstraction operator**.

Consider now our Abstraction Principle:

$$(\S x.{}^{\varphi(x)}/_{F}) = (\S x.{}^{\psi(x)}/_{G}) \longleftrightarrow \Phi[{}^{\varphi(x)}/_{F}, {}^{\psi(x)}/_{G}].$$

We denote with $\Phi[F, G]$ a second-order formula (with exactly two free-variables) that encapsulate the equivalence relation occurring between the two concepts involved. An equivalence or congruence relation $R$ (i) determines the formula $\Phi[F, G]$ of an abstraction principle and (ii) settles the interpretation of $=$. Hence, since $\mathscr{L}$ contains a primitive identity symbol, by (ii), $\mathcal{M}$ "thinks" that it is an identity symbol, that we interpret as a congruence predicate rather then the actual identity. Additionally, denote with $Val_R(\varphi)$ the valuation function of $\varphi$ in our second order model. It follows that any $\varphi$ of $\mathscr{L}$, such that $\varphi[\forall F^n(F^n \ldots F^n \ldots)]$, is subject to the maximal constraint that the second-order quantifiers range over all concepts, with respect to the equivalence relation $R$. So, for $F \in Con$, if $R(x, y)$, then for all $Con \subseteq Dom$, $x \in F$ iff $y \in F$. Moreover, if $\neg R(x, y)$, then, for some $Con \subseteq Dom$, such that $F \in Con$, $x \in F$ iff $y \notin F$.

Anyway, before we give the main truth clauses of our account, let's focus on a moment on the notion of variable assignment. We consider a function $\sigma$ that (i) for each variable outputs an object $x \in Dom$, (ii) for each relation-variable outputs an element of $Con$, namely $\sigma(F^n) \in Con$ and (iii) for every function-variable outputs an element of $Fun$, i.e. $\sigma(f^n) \in Fun$:

**Definition 35.** We define satisfaction for an element $t$ of $Dom$ with free variables $x_1, \ldots, x_n$, as follows:

$$t^{\mathcal{M},\sigma} = \sigma(x_i),$$

if $t$ is the variable $x_i$. Now, we define satisfaction relative to variable-assignment $\sigma$:

$$\mathcal{M}, \sigma \vDash t_1 = t_2 \text{ iff } t_1^{\mathcal{M},\sigma} = t_2^{\mathcal{M},\sigma},$$

Furthermore,

$$\mathcal{M}, \sigma \vDash f^n(t_1, \ldots, t_2) \text{ iff } \mathcal{M}, \sigma \vDash (t_1^{\mathcal{M},\sigma}, \ldots, t_n^{\mathcal{M},\sigma}) \in \sigma(f^{n,\mathcal{M}})$$

and,

$$\mathcal{M}, \sigma \vDash F^n(t_1, \ldots, t_n) \text{ iff } \mathcal{M} \vDash (t_1^{\mathcal{M},\sigma}, \ldots, t_n^{\mathcal{M},\sigma}) \in \sigma(F^{n,\mathcal{M}}),$$

for, $\mathscr{L}$-terms $t_1, \ldots, t_n$.

$$\mathcal{M}, \sigma \vDash \neg\varphi \text{ iff } \mathcal{M}, \sigma \nvDash \varphi$$

$$\mathcal{M}, \sigma \vDash (\varphi \wedge \psi) \text{ iff} : \mathcal{M}, \sigma \vDash \varphi \text{ and } \mathcal{M}, \sigma \vDash \psi$$

$$\mathcal{M}, \sigma \vDash (\varphi \vee \psi) \text{ iff} : \mathcal{M}, \sigma \vDash \varphi \text{ or } \mathcal{M}, \sigma \vDash \psi$$

$$\mathcal{M}, \sigma \vDash (\varphi \to \psi) \text{ iff} : \text{if } \mathcal{M}, \sigma \vDash \varphi, \text{ then } \mathcal{M}, \sigma \vDash \psi$$

We add the clause for universal quantification:

$$\mathcal{M}, \sigma \vDash \forall F^n \varphi(F^n) \text{ iff } \mathcal{M}, \tau \vDash \varphi(F^n),$$

And

$$\mathcal{M}, \sigma \vDash \forall f^n \varphi(f^n) \text{ iff } \mathcal{M}, \tau \vDash \varphi(f^n),$$

for every assignment $\tau$ which agrees on everything with $\sigma$, except possibly on $F^n$ (or on $f^n$).
Finally, we are able to say that for any $\mathscr{L}$-sentence $\varphi$:

$$\mathcal{M} \vDash \varphi \text{ iff } \mathcal{M}, \sigma \vDash \varphi,$$

for any variable-assignment $\sigma$.

**Remark.** Consider $\mathcal{M}$ and suppose $R$ is any one-place relation variable such that $R \subseteq Dom$. We want $\mathcal{M}$ to satisfy the following principle: $\exists F \forall v (R(v) \longleftrightarrow F(v))$. By inspection, we see that $R$ is itself a witness to the second-order quantifier. But this holds just in case $R \in Con$, and this is not guaranteed in any $\mathscr{L}$-structure. Thus, it is natural to add the following principle:

**Axiom (Comprehension principle for relations or properties).**

$$\exists F^n \, \forall v_1, \ldots, v_n \left[ F^n(v_1, \ldots, v_n) \longleftrightarrow \varphi(v_1, \ldots, v_n) \right],$$

for every formula $\varphi(v_1, \ldots, v_n)$ which does not contain the relation variable $F^n$ itself[25].

Furthermore, the notion of a Henkin model has to be restricted still further: we *accept* only those Henkin models that satisfy the comprehension principle, as well as the remaining second-order axioms. Hence, we assume that the Henkin models considered are restricted to those that satisfy the comprehension schema, and these are to be considered *faithful* models.

[25] If we do not add the following restriction, it could be always derived a Russell-style paradox. As usual, in what follows, we allow the $n$-tuples of variables, such as $x_1, \ldots, x_n$, to be abbreviated as $\overline{x}$.

Return considering abstraction principles. We want the RHS determine the truth value of the LHS in a well-founded manner: thanks to the truth value of the elements in the formula encapsulating the equivalence relation on *Dom*, we will be able to determine the truth-value of the identity statement between the two abstracted concepts. Our formal machinery is supposed to work in the following manner:

$$\underbrace{\Phi[F, G]}_{\text{equivalence relation}}$$
$$\downarrow$$
second-order formula
whose range is restricted to *Con*
$$\downarrow$$
$$\Phi[{}^{\varphi}/_{F}, {}^{\psi}/_{G}]$$
$$\downarrow$$
where $R$ determines the value of $\Phi$,
with respect to the variable-assignment $\sigma$
$$\downarrow$$
$$Val_R(\Phi[{}^{\varphi}/_{F}, {}^{\psi}/_{G}])$$
$$\downarrow$$
$R$ settles also the interpretation of $=$,
and hence the truth value of the abstracta
$$\downarrow$$
$$Val_R\Big[\ \underbrace{\underbrace{\S({}^{\varphi}/_{F})}_{\mathcal{N}(F)=\mathsf{x}} = \underbrace{\S({}^{\psi}/_{G})}_{\mathcal{N}(G)=\mathsf{y}}}_{\text{Proper abstract objects}}\ \Big]$$

As we can see from their formal general version, it becomes clear that the value of the RHS, containing an equivalence relation between two concepts, yields the value of the LHS, therefore of the identity between the two “abstracted concepts”, i.e. the *new* objects.
Now, recall that we allowed our model to think that $=$ is the identity sign. It is natural to want that Leibniz’s Law remains valid[26] and, in order to do this, we want that our equivalence relations $R$ satisfy the following closure condition:

- ∘ Let $t_1, t_2$ be two terms denoting (pre-)objects, such that for some $F, G \in Con$, $\mathcal{N}(F) = \mathsf{t}_1$ and $\mathcal{N}(G) = \mathsf{t}_2$. Let $T', T''[{}^{t_2}/_{t_1}]$ be two complex terms, such that $T''$ is the result of substituting each instance of $t_1$ in $T'$ with $t_2$. Then, if $\langle t_1, t_2\rangle \in R$, then $\langle T', T''\rangle \in R$.

This closure condition allows us to prove the following proposition:

**Proposition 13.** Let $Con \subseteq \wp(Dom)$ be the set of all first-order definable formulas with one (or more) free variable(s). Let *Dom* be the set of (pre-)objects, and $R$ any equivalence relation on *Dom* that respects ∘. Then, $\mathcal{M}[R]$ is a Henkin model of second-order logic.

[26] See Horsten and Linnebo 2016, p. 5.

*Proof.* The only part of the proof that is not immediate from the definitions of satisfaction for $\mathcal{M}$, it is the one that concerns Leibniz's Law. Let $t$ and $t'$ be terms in *Dom* and $\varphi$ any formula. We want to show that:

$$\mathcal{M}[R], \sigma \vDash \varphi \longleftrightarrow \varphi({}^{t'}/_{t}).$$

The proof goes by induction on the complexity of $\varphi$. Suppose $\varphi$ is atomic. If $\varphi =_{\text{def}} t = t$, then, by the fact that $R$ is an equivalence relation on *Dom*, it follows:

$$\mathcal{M}[R], \sigma \vDash t = t \longleftrightarrow {}^{t'}/_{t} = {}^{t'}/_{t}.$$

For the remaining atomic cases, we simply need to invoke the closure condition $\circ$. Suppose now, $\varphi =_{\text{def}} Fx_1, \ldots, x_n$, where $F$ is a second-order variable. We have to show that Leibniz's Law holds for any first-order definable formula $\varphi$. Let's prove it via induction on the complexity of $\varphi$. The base case is the one considered above. The induction steps are:

$$\text{if } \mathcal{M}[R], \sigma \vDash \neg\varphi(t = t) \longleftrightarrow \neg\varphi({}^{t'}/_{t} = {}^{t'}/_{t}), \text{ then :}$$
$$\mathcal{M}[R], \sigma \nvDash \varphi(t = t) \text{ iff } \mathcal{M}[R], \sigma \nvDash \varphi({}^{t'}/_{t} = {}^{t'}/_{t})$$

$$\text{if } \mathcal{M}[R], \sigma \vDash [\varphi \wedge \psi](t = t) \longleftrightarrow [\varphi \wedge \psi]({}^{t'}/_{t} = {}^{t'}/_{t}), \text{ then :}$$
$$\mathcal{M}[R], \sigma \vDash \varphi(t = t) \wedge \psi(t = t) \text{ iff}$$
$$\mathcal{M}[R], \sigma \vDash \varphi({}^{t'}/_{t} = {}^{t'}/_{t}) \wedge \psi({}^{t'}/_{t} = {}^{t'}/_{t}), \text{ iff}$$
$$\mathcal{M}[R], \sigma \vDash \varphi(t = t) \text{ and } \mathcal{M}[R], \sigma \vDash \psi(t = t) \text{ iff}$$
$$\mathcal{M}[R], \sigma \vDash \varphi({}^{t'}/_{t} = {}^{t'}/_{t}) \text{ and } \mathcal{M}[R], \sigma \vDash \psi({}^{t'}/_{t} = {}^{t'}/_{t})$$

$$\text{if } \mathcal{M}[R], \sigma \vDash [\varphi \vee \psi](t = t) \longleftrightarrow [\varphi \vee \psi]({}^{t'}/_{t} = {}^{t'}/_{t}), \text{ then :}$$
$$\mathcal{M}[R], \sigma \vDash \varphi(t = t) \vee \psi(t = t) \text{ iff}$$
$$\mathcal{M}[R], \sigma \vDash \varphi({}^{t'}/_{t} = {}^{t'}/_{t}) \vee \psi({}^{t'}/_{t} = {}^{t'}/_{t}) \text{ iff}$$
$$\mathcal{M}[R], \sigma \vDash \varphi(t = t) \text{ or } \mathcal{M}[R], \sigma \vDash \psi(t = t) \text{ iff}$$
$$\mathcal{M}[R], \sigma \vDash \varphi({}^{t'}/_{t} = {}^{t'}/_{t}) \text{ or } \mathcal{M}[R], \sigma \vDash \psi({}^{t'}/_{t} = {}^{t'}/_{t})$$

$$\mathcal{M}[R], \sigma \vDash [\varphi \to \psi](t = t) \longleftrightarrow [\varphi \to \psi]({}^{t'}/_{t} = {}^{t'}/_{t}), \text{ then :}$$
$$\mathcal{M}[R], \sigma \vDash \varphi(t = t) \to \psi(t = t) \text{ iff}$$
$$\mathcal{M}[R], \sigma \vDash \varphi({}^{t'}/_{t} = {}^{t'}/_{t}) \to \psi({}^{t'}/_{t} = {}^{t'}/_{t}) \text{ iff}$$
$$\text{if } \mathcal{M}[R], \sigma \vDash \varphi(t = t) \text{ then, } \mathcal{M}[R], \sigma \vDash \psi(t = t) \text{ iff}$$
$$\text{if } \mathcal{M}[R], \sigma \vDash \varphi({}^{t'}/_{t} = {}^{t'}/_{t}) \text{ then, } \mathcal{M}[R], \sigma \vDash \psi({}^{t'}/_{t} = {}^{t'}/_{t})$$

$$\begin{aligned}\mathcal{M}[R], \sigma \vDash [\forall\varphi](t = t) \longleftrightarrow [\forall\varphi]({}^{t'}/_{t} = {}^{t'}/_{t}),\ \text{then :}\\ \mathcal{M}[R], \sigma \vDash \forall\varphi(t = t) \text{ iff } \mathcal{M}[R], \tau \vDash \varphi(t = t),\\ \text{for every } \tau \text{ which agrees with } \sigma \text{ on everything, iff}\\ \mathcal{M}[R], \sigma \vDash \forall\varphi({}^{t'}/_{t} = {}^{t'}/_{t}) \text{ iff } \mathcal{M}[R], \tau \vDash \varphi({}^{t'}/_{t} = {}^{t'}/_{t}),\\ \text{for every } \tau \text{ which agrees with } \sigma \text{ on everything.}\end{aligned}$$

This completes our inner induction on the complexity of $\varphi$.
The outer induction cases are identical to the previous and, thus, our proof of Leibniz's Law is complete.
Consider now full comprehension for relations. We want our model to satisfy:

$$\mathcal{M}[R], \sigma \vDash \exists F^n \forall \overline{v} \left[ F^n(\overline{v}) \longleftrightarrow \varphi(\overline{v}) \right]$$

Let $\varphi$ be a formula with second-order variables. Thanks to our clause for quantification, every assignment $\sigma$ maps any second order variable to a first-order definable formula of $Con$. Thus, any second-order variable $F^n$ in $\varphi$ may be replaced by a first-order formula, which establishes that comprehension on $\varphi$, relatively to $\sigma$, is allowed.
This last point terminates our proof and we can easily establish that $\mathcal{M}[R], \sigma$ is a $\mathscr{L}$-Henkin model for second-order logic. ■

But, how do we determine the value of a formula in the setting we've just sketched? Leitgeb offers an answer. The notion of "dependence" we are seeking for can be informally rendered as follows: the identity between two abstracta depends upon the identities between other abstracta – the latter, having been previously settled in a well-founded manner. Our work will be firstly concerned with the notion of "dependency" and, secondly, with respect to that notion, we will provide a hierarchical construction of sets of grounded identity/difference facts, which allow us to prove some nice theorems concerning the *usage* of abstraction principles. As stated, Leitgeb's approach remains our main inspiration source.

#### 4.3.2.2 Dependence, Identity and Difference

Consider that, even if in a different context,

> "The" notion of dependence is one of the most frequently employed (and perhaps also abused) notions in philosophy. Different kinds of dependency have been considered in fields such as metaphysics, philosophy of mind and philosophy of science, but few general, systematic theories of dependence have been developed that are also applicable in a semantic context[27].

[27]Leitgeb 2005, p. 159. As remarked several times, our considerations are semantic oriented, in the sense that we are studying a faithful way to formalize, with a model-theoretic approach, notions such as "dependence", "supervenience" and "groundedness". For more about ontological and epistemological concerns about the usage and the interpretation of this model, see our Chapter 5.

The notion of dependence we are seeking for may be posed as follows: "the identities (or differences) of two (pre-)objects *depend* on the identities (or differences) between other (pre-)objects", with respect to $\mathcal{N}$ and $\Phi$ (defined above):

**Definition 36.** For all $\langle x, y\rangle \in Dom \times Dom$, for all $Z \subseteq Dom \times Dom$:

$$\langle x, y\rangle \textbf{ depends}_{\mathcal{N},\Phi} \textbf{ on } Z \text{ iff}$$

for some $A, B \in Con$, such that $x = \mathcal{N}(A),\ y = \mathcal{N}(B)$, and for any equivalence relation $R_1,\ R_2 \subseteq Dom \times Dom$, it holds that:

if, for all $A', B' \in Con$, such that $\langle \mathcal{N}(A'), \mathcal{N}(B')\rangle \in Z$,

$$Val_{R_1,\sigma}(\S(\overline{A'}) = \S(\overline{B'})) = Val_{R_2,\sigma}(\S(\overline{A'}) = \S(\overline{B'}))$$

then,

$$Val_{R_1,\sigma}(\Phi[\overline{A}, \overline{B}]) = Val_{R_2,\sigma}(\Phi[\overline{A}, \overline{B}])$$

Recall that, since $\mathcal{N}$ is injective, then $A$ and $B$ in our definition are uniquely determined. Hence, "$\langle x, y\rangle$ depends$_{\mathcal{N},\Phi}$ on $Z$" indicates that the identity between $x$ and $y$ depends univocally on which identities are already settled in $Z$. If we state the contrapositive of our if-then clause, then dependence means: if there is difference with respect to $\Phi[\overline{A}, \overline{B}]$ as determined by an equivalence relation $R$, then there is a corresponding difference with respect to the members of $Z$. Hence, here, the semantic sense of dependence means exactly that the (truth-value of the) identity (or difference) between $\S xFx$ and $\S xGx$, "supervenes" the (truth-value of the) equivalence relation $R$ as expressed by $\Phi[F, G]$. Therefore, in outline, two concepts, $A$ and $B$ of our codomain, thanks to an abstraction mapping $\mathcal{N}$, and with respect to $\Phi$, are mapped to some pre-objects, $x = \mathcal{N}(A),\ y = \mathcal{N}(B)$; where, these resulting items are what we might call proper abstract objects, i.e., $\mathsf{x} = \mathcal{N}(A)$ and $\mathsf{y} = \mathcal{N}(B)$ . Hence, with respect to the choice of $\mathcal{N}$ and $\Phi$, different sets, of proper abstract objects – whose identity or difference conditions have already been settled – will emerge. Remember, additionally, that sets generated by two different abstraction principles are not identified by the start, because, perhaps, at further stages and thanks to another well-defined equivalence relation, such a cross-identification may be successful.[28]

Furthermore, by analysing our definition of dependence, if we fix our **first relatum**[29] of the dependence relation, then the following properties may reveal useful:

**Lemma 14** (Properties of Dependence with respect to the Fixed First Relatum)**.** For all $\langle x, y\rangle \in Dom \times Dom$ and for all $\Delta, \Gamma \subseteq Dom \times Dom$:

(1) if $\langle x, y\rangle$ depends$_{\mathcal{N},\Phi}$ on $\Delta$ and $\Delta \subseteq \Gamma$, then $\langle x, y\rangle$ depends$_{\mathcal{N},\Phi}$ on $\Gamma$.

[28] Leitgeb 2017, p. 274.

[29] For our main source, see Leitgeb 2005, p. 161.

(2) if $\langle x, y\rangle$ depends$_{\mathcal{N},\Phi}$ on $\Delta$ and $\langle x, y\rangle$ depends$_{\mathcal{N},\Phi}$ on $\Gamma$, then $\langle x, y\rangle$ depends$_{\mathcal{N},\Phi}$ on $\Delta \cap \Gamma$.

(3) $\langle x, y\rangle$ depends$_{\mathcal{N},\Phi}$ on *Dom*.

Additionally, by taking some notions from set theory, let's say that $X \subseteq \wp(Dom)$ is called a **filter** iff for all $\Gamma, \Delta \subseteq Dom$:

(i) $X \in Dom$ and $\varnothing \notin X$

(ii) if $\Gamma \in X$, $\Gamma \subseteq \Delta \subseteq X$, then $\Delta \in X$;

(iii) if $\Gamma, \Delta \in X$, then $\Gamma \cap \Delta \in X$.

**Remark.** For any $\Delta \subseteq X$, with $\Delta \neq \varnothing$, the set $\{\Gamma \in Dom \mid \Delta \subseteq \Gamma\}$ is called the **principal filter generated by** $\Delta$.

Moreover, $X \subseteq \wp(Dom)$ is an **ultrafilter** iff:

(i) $X$ is a filter on *Dom*;

(ii) for all $\Gamma \subseteq Dom$, either $\Gamma \in X$ or $Dom \backslash \Gamma \in X$.

Now, for any $Z \subseteq Dom \times Dom$, let

$$\mathfrak{D}_{\mathcal{N},\Phi}(Z) = \{\langle x, y\rangle \in Dom \times Dom \mid \langle x, y\rangle \text{ depends}_{\mathcal{N},\Phi} \text{ on } Z\}$$

the **dependence chain of Z**, that is the set of all pairs $\langle x, y\rangle$ in *Dom* that depend (in the sense of our previous definition) on $Z$. In other words, we've encountered the dependence chain of a set $Z$, i.e., all the sets of pairs in $Dom \times Dom$ that depend on $Z$. Additionally, the notion of dependence chain of $\langle x, y\rangle$ may reveal useful. For $\langle x, y\rangle \in Dom \times Dom$:

$$\mathfrak{F}_{\mathcal{N},\Phi}(\langle x, y\rangle) = \{Z \subseteq Dom \times Dom \mid \langle x, y\rangle \text{ depends}_{\mathcal{N},\Phi} \text{ on } Z\}$$

In other terms, $\mathfrak{F}_{\mathcal{N},\Phi}(\langle x, y\rangle)$, – the **dependence chain of pairs of individual** –, is the collection of all subsets of $Dom \times Dom$ on which pairs of individuals, such as $\langle x, y\rangle$, depend on (as previously defined). In this context, notice that, the notion of **least** element in a set can be introduced: let $X \subseteq \wp(Dom)$, we call $\Gamma$ least in $X$ iff $\Gamma \in X$ and for all $\Delta \in X$, it holds that $\Gamma \subseteq \Delta$. In other words, $\Gamma$ is the "generator" set of $X$. As in the case of truth, least members of $\mathfrak{F}_{\mathcal{N},\Phi}(Z)$ correspond to sets of pairs that $\langle x, y\rangle$ depends on **essentially**. We may, indeed, define for $\langle x, y\rangle \in Dom \times Dom$, $Z \subseteq Dom \times Dom$:

**Definition 37.** $\langle x, y\rangle$ **depends**$_{\mathcal{N},\Phi}$ **essentially on** $Z$ iff:

$$Z \text{ is } \textbf{least} \text{ in } \mathfrak{F}_{\mathcal{N},\Phi}(\langle x, y\rangle)$$

Clearly, $\mathfrak{F}_{\mathcal{N},\Phi}$ maps pairs of individuals to the set of pairs individuals on which they depend; while $\mathfrak{D}_{\mathcal{N},\Phi}$ maps any set $Z$ on which a pair of individuals depend on, to

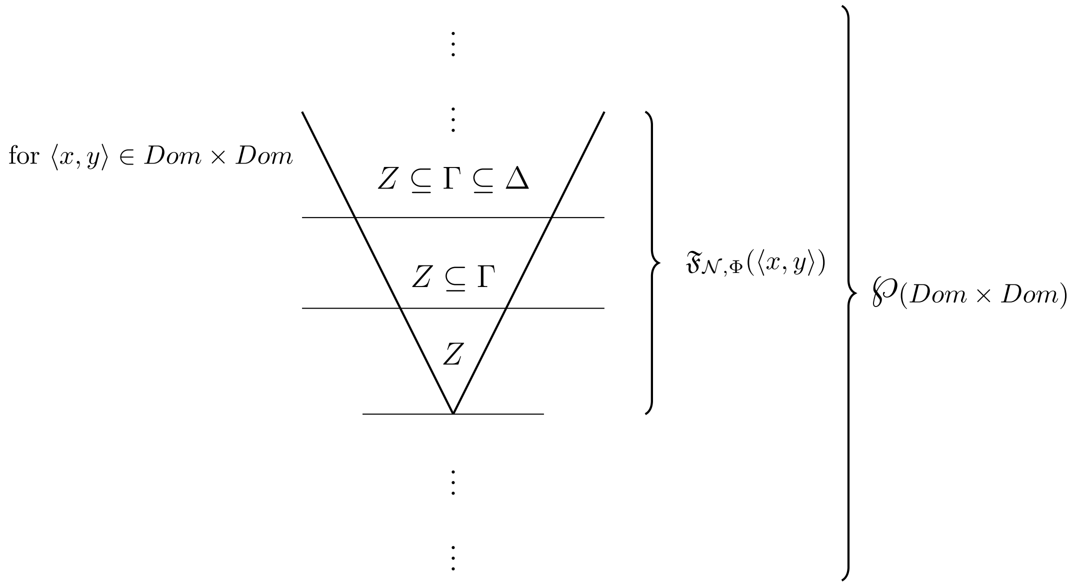



Figure 4.1: $Z$ is least in $\mathfrak{F}_{\mathcal{N},\Phi}(\langle x, y\rangle)$

the other sets characterised by the same dependence relation with respect to the pair of individuals. Hence, $\mathfrak{D}_{\mathcal{N},\Phi}$ maps sets of pairs of individuals to other sets of pairs of individuals. The relation between the two dependence chains considered is: $Z \in \mathfrak{F}_{\mathcal{N},\Phi}(\langle x, y\rangle)$ iff $\langle x, y\rangle \in \mathfrak{D}_{\mathcal{N},\Phi}(Z)$.

*Proof.* $(\rightarrow)$
Suppose $Z \subseteq Dom \times Dom$ and assume $Z \in \mathfrak{F}_{\mathcal{N},\Phi}(\langle x, y\rangle)$. By definition:

$$Z \in \{Z \subseteq Dom \times Dom \mid \langle x, y\rangle \text{ depends}_{\mathcal{N},\Phi} \text{ on } Z\}$$

Indeed, $\langle x, y\rangle$ depends$_{\mathcal{N},\Phi}$ on $Z$ means that, actually, $Z$ is one of the sets on which our pair of individuals depends on. This means that $Z$ belongs to the collection of sets on which $\langle x, y\rangle$ depends on, namely $\mathfrak{D}_{\mathcal{N},\Phi}(Z)$. Likewise, the pair of individuals depending on $Z$ is in the collection of all of the pairs of individuals that depend on $Z$, so

$$\langle x, y\rangle \in \mathfrak{D}_{\mathcal{N},\Phi}(Z)$$

$(\leftarrow)$
Suppose $\langle x, y\rangle \in Dom \times Dom$ and assume $\langle x, y\rangle \in \mathfrak{D}_{\mathcal{N},\Phi}(Z)$. By definition:

$$\langle x, y\rangle \in \{\langle x, y\rangle \in Dom \times Dom \mid \langle x, y\rangle \text{ depends}_{\mathcal{N},\Phi} \text{ on } Z\}$$

Since, $\langle x, y\rangle$ is in the dependence chain of $Z$ and, thus, $Z$ is one of the set of pairs of individuals on which $\langle x, y\rangle$ depends on, we may conclude

$$Z \in \mathfrak{F}_{\mathcal{N},\Phi}(\langle x, y\rangle)$$

■

With all this background we may interested in the following lemma:

**Lemma 15** (Properties of dependence with respect to the abstraction operator)**.** For all $\langle x, y\rangle \in Dom \times Dom$, $Z \subseteq Dom \times Dom$:

$$\langle x, y\rangle \in Z \textit{ iff},$$

for some $A, B \in Con$, such that $\mathcal{N}(A) = \mathsf{x}$, $\mathcal{N}(B) = \mathsf{y}$, and for all $R \subseteq Dom \times Dom$ it holds that:

$$\langle \mathcal{N}(A), \mathcal{N}(B)\rangle \text{ depends}_{\mathcal{N},\Phi} \text{ on } Z$$

Equivalently:

$$\langle x, y\rangle \in Z \textit{ iff } \langle \mathcal{N}(A), \mathcal{N}(B)\rangle \in \mathfrak{D}_{\mathcal{N},\Phi}(Z)$$

What we've encapsulated in the previous lemma is fundamental, since it tells us that a dependence chain of $Z$ is closed under the application of the abstraction operator to the couples of individuals members of $Z$. In other words, abstracta are within the of dependence chain of $Z$, just in case the pair of individuals that may become proper abstract objects (what we've called pre-objects) are in $Z$.

#### 4.3.2.3 The Set of Grounded Identity and Difference Facts

Well, the next step is to proceed towards the construction of the set of grounded identity and difference facts. First of all, let's begin by considering some characters of $\mathfrak{D}_{\mathcal{N},\Phi}$ itself. Consider that in this context, the operator $\mathfrak{D}_{\mathcal{N},\Phi}$ is a map $\wp(Dom) \to \wp(Dom)$, such that, $Z \mapsto \mathfrak{D}_{\mathcal{N},\Phi}(Z)$. As in the case of truth, its fundamental property is that of being **monotonic**, that is, for all $\Gamma, \Delta \subseteq Dom \times Dom$: if $\Gamma \subseteq \Delta$, then $\mathfrak{D}_{\mathcal{N},\Phi}(\Gamma) \subseteq \mathfrak{D}_{\mathcal{N},\Phi}(\Delta)$[30]. The monotonicity we've just stated implies the following points:

(i) it exists a least fixed point $E_{lf} \subseteq Dom \times Dom$ of $\mathfrak{D}_{\mathcal{N},\Phi}$.

(ii) for all, $\langle x, y\rangle \in Dom \times Dom$: $\langle x, y\rangle \in E_{lf}$ iff $\langle x, y\rangle$ depends$_{\mathcal{N},\Phi}$ on $E_{lf}$.

(iii) $E_{lf}$ can be reached from below by iterated application of $\mathfrak{D}_{\mathcal{N},\Phi}$, as follows:

$$E_0 =_{\text{def}} \varnothing$$

$$E_{\alpha+1} =_{\text{def}} \mathfrak{D}_{\mathcal{N},\Phi}(E_\alpha)$$

$$E_\lambda =_{\text{def}} \bigcup_{\alpha<\lambda} E_\alpha \text{ (for } \lambda \text{ limit ordinal).}$$

[30] *Proof Sketch*: Suppose that $\langle x, y\rangle$ depends$_{\mathcal{N},\Phi}$ on $\Gamma$, that is $\langle x, y\rangle \in \mathfrak{D}_{\mathcal{N},\Phi}(\Gamma)$. Assume that, for all $\Gamma, \Delta \subseteq Dom$ with $\Gamma \subseteq \Delta$. Then, by (1) of Lemma 15 of the present chapter, we know that $\langle x, y\rangle$ depends$_{\mathcal{N},\Phi}$ on $\Delta$ too. Hence, $\langle x, y\rangle \in \mathfrak{D}_{\mathcal{N},\Phi}(\Delta)$. Since $\Gamma \subseteq \Delta$, we may conclude that $\mathfrak{D}_{\mathcal{N},\Phi}(\Gamma) \subseteq \mathfrak{D}_{\mathcal{N},\Phi}(\Delta)$. ■

Furthermore, this sequence can be shown to be monotonically increasing and $E_{lf}$ is what we were searching for: the set of grounded identity and difference facts.
Now, recall that in order to restrict abstraction principles to their grounded instances we were searching all and only those identities on which the identities of abstracta depends on. In other words, we were interested in restricting abstraction principle in the following manner, by using an accurate terminology: $x$ is identical to $y$, according to a given abstraction principle, just in case their identity depends (only) on the identities that hold in a subset of the domain. Reformulating once again, the identity between two "proper" abstract objects hold if the identity settled previously between the $\mathcal{N}-$preimage of the labels in $Dom$ hold. With this formal framework we have already reached the set of identity and difference facts, namely $E_{lf}$. This means that the set $E_{lf}$, with the iterated application of $\mathfrak{D}_{\mathcal{N},\Phi}$, collects all the sets on which the pairs of individuals depend on and viceversa. That is, within $E_{lf}$ there are not just the identity, but also the difference facts and, indeed, $E_{lf}$ is the set that collects the dependence and not-dependence relations between pairs of individuals. Recall, indeed, that our dependency definition implies that, if in some $Z \subseteq Dom \times Dom$ the identity of $x, y$ does not hold, then the identity between the abstracted $x, y$ does not hold. It follows that, for $\alpha \in Ord$, $E_\alpha$ is increasing and there is a least ordinal $\alpha^*$, such that $E_{\alpha^*} = E_{lf}$. Anyway, $E_{lf}$ is not the only fixed point of $\mathfrak{D}_{\mathcal{N}\Phi}$ and, indeed, for instance, $Dom = \mathfrak{D}_{\mathcal{N},\Phi}(Dom)$ is its largest point. The abstraction process $E_\alpha$ has the following property:

**Lemma 16** (Properties of $(E_\alpha)_{\alpha \in Ord}$ with respect to the abstraction operator)**.** Let $E_\alpha \subseteq Dom \times Dom$. For all ordinals $\alpha$:

$$\langle x, y \rangle \in E_\alpha \text{ iff}$$

for some $A, B \in Con$, such that $x = \mathcal{N}(A), y = \mathcal{N}(B)$, and for all $R \subseteq Dom \times Dom$, it is the case that:

$$\langle \mathcal{N}(A'), \mathcal{N}(B') \rangle \in E_{\alpha+1}$$

Hence, by the previous lemma, members of $E_{lf}$ may be assigned with an ordinal rank, such that the first occurrence of the identity (or difference) fact in the progression $E_\alpha$ defines its *dependence rank*. Indeed,

**Definition 38** (Dependence rank)**.** For $\langle x, y \rangle \in E_{lf}$, $\langle x, y \rangle$ has a **dependence rank** $\alpha$ iff:

$$\langle x, y \rangle \in E_\alpha \text{ and for all } \beta < \alpha : \langle x, y \rangle \notin E_\beta$$

What we're told in the previous definition is that, every $\langle x, y \rangle \in E_{lf}$ is also a member of a set $E_\alpha$ and that there is a least ordinal $\alpha^*$ such that $E_{lf} =_{\text{def}} E_{\alpha^*}$. This least ordinal is the dependence rank of $\langle x, y \rangle$. Moreover, recall that the class $\alpha \in Ord$ is well-ordered and, thus, the dependence rank of any $\langle x, y\rangle \in E_{lf}$ is uniquely determined.

Now, in order to determine only which instances of an abstraction principles hold, let's isolate the identity facts, rather that the difference ones, from $E_{lf}$. Let's do it in the following way:

$$F_0 =_{\text{def}} \varnothing$$

$$\begin{aligned} F_{\alpha+1} =_{\text{def}} \{\langle x, y\rangle \in E_{\alpha+1} \mid \exists A, B \in Con : x = \mathcal{N}(A), y = \mathcal{N}(B), \\ Val_{RST(F_\alpha)}(\S(\overline{A}) = \S(\overline{B})) = 1\} \end{aligned}$$

$$F_\lambda =_{\text{def}} \bigcup_{\alpha<\lambda} F_\alpha \text{ (for } \lambda \text{ limit ordinal).}$$

This definition implies:

**Lemma 17** (Properties of $(F_\alpha)_{\alpha\in Ord}$ with respect to $(E_\alpha)_{\alpha\in Ord}$)**.** For $\alpha \in Ord$:

$$E_\alpha = F_\alpha \cup F_\alpha^-,$$

where, for $E \subseteq Dom \times Dom$, we set:

$$F^- = \{\langle x, y\rangle \in Dom \times Dom \mid \neg\langle x, y\rangle \in E\}.$$

As it should already be clear, $F_{lf}$ is (just) the set of grounded identities facts, while, as defined before, $F_{lf}^-$ contains just the negative identity facts, namely the differences. The union of both $F_{lf}$s gives as result the entire $E-$hierarchy. Recall, that at the beginning of the following section, we've said that we allowed our second-order language to contain the identity symbol "=", but that our $\mathcal{M}$ interprets it as the equivalence relation on $Dom$. If we have a look to the recursive definition of $F_{\alpha+1}$ we notice that the value of the elements in sets $F_{\alpha+1}$ is restricted to the reflexive-symmetric-transitive ($RST$) closure of the elements in $F_\alpha$. This means that the truth value of the identity between members of $F_{\alpha+1}$ (as given by an abstraction principle), is determined by the value of the $RST$ closure of the elements in $F_\alpha$ and from which we abstract.
By having isolated just the identity facts, let's establish the following definition:

**Definition 39** (Identity rank)**.** For $\langle x, y\rangle \in \bigcup_{\alpha\in Ord} F_\alpha$, $\langle x, y\rangle$ has an **identity rank** $\alpha$, iff:

$$\langle x, y\rangle \in F_\alpha \text{ and for all } \beta < \alpha : \ \langle x, y\rangle \notin F_\beta$$

In order to say when an identity instance is true in our *Dom*, we need the following result[31]:

**Theorem 18** (Convergence of $(F_\alpha)_{\alpha\in Ord}$)**.** For all $\alpha, \beta \in Ord$, with $\beta < \alpha$:

$$F_\alpha \cap E_\beta = F_\beta$$

The theorem implies that $(F_\alpha)_{\alpha\in Ord}$ is increasing and, moreover, that it must converge to a limit. This means that, for some $\alpha^+ \in Ord$, $\alpha^+$ is the least ordinal $\alpha$ such that, for all $\beta \in Ord$, with $\beta > \alpha$, $F_\beta = F_\alpha$. We define, finally, $F_{lf} =_{\text{def}} F_{\alpha^+}$. This implies that $\bigcup\limits_{\alpha\in Ord} F_\alpha$ can now be simply renamed $F_{lf}$.

Now, consider our dependence definition. We have that for $\langle x, y\rangle \in E_\beta$, such that $\alpha > \beta$:

$$Val_{F_\alpha}(\langle x, y\rangle) = Val_{F_\beta}(\langle x, y\rangle)$$

This fact is important since it tells us that for any $\langle x, y\rangle \in E_{lf}$, its truth value, as determined by $Val_{(F_\alpha)_{\alpha\in Ord},\sigma}$ stabilise after its dependence rank. By a previous result, we know that an identity statement in the set of grounded identity and difference facts has a dependence rank (the ordinal indicating the first time of its occurrence). Now we've just stated that the truth value, as evaluated by $Val_{R,\sigma,F_\alpha}$, stabilises after its dependence rank. Let's consider an example: if $\langle x, y\rangle$ is in some $E_\beta$, then – by isolating the identity facts, by means of the $F-$hierarchy – the evaluation function of $F_\beta$ gives 1 if the identity between $\langle x, y\rangle$ holds (0 if it is not the case). Moreover, for all $\alpha > \beta$, the evaluation function of any $F_\alpha$ will output 1 (or 0, dependently) and, hence, the truth value of an identity (or difference) statement gets settled after its dependence rank has been established. Therefore, we may conclude that for any $\langle x, y\rangle \in F_{lf}$ the identity rank of $\langle x, y\rangle$ and its dependence rank coincide:

**Lemma 19** (Rank-coincidence)**.** For $\langle x, y\rangle \in E_{lf}$:

$\langle x, y\rangle$ has a **rank** $\alpha$ iff:

$\langle x, y\rangle$ has a **dependence rank** $\alpha$, iff,

for $\langle x, y\rangle \in F_{lf}$:

$\langle x, y\rangle$ has an **identity rank** $\alpha$.

Of course, this lemma entails that $F_\alpha$ eventually reaches a fixed point $F_{lf}$ at the same ordinal as the corresponding $E-$progression. That is least fixed points of $F_\alpha$ and $E_\alpha$ coincide as well, $\alpha^* = \alpha^+$ and, indeed, let's denote $F_{lf} = F_{\alpha^*}$ and $E_{lf} = E_{\alpha^*}$. The whole theory we've just setup, finally, concedes us to introduce the following properties of the truth value of a pairs of individuals in $F_{lf}$, with respect to $E_{lf}$:

[31] *Proof sketch*: Standard transfinite mathematical induction.

**Theorem 20** (Truth in $F_{lf}$, with respect to $E_{lf}$)**.** For all $\langle x, y\rangle \in E_{lf}$,

1. $\langle x, y\rangle \in F_{lf}$ iff $Val_{F_{lf}}(\langle x, y\rangle) = 1$.

For all $\langle x, y\rangle \in E_{lf}$, for all $A, B \in Con$, s.t. $x = \mathcal{N}(A), y = \mathcal{N}(B)$ and for $R \subseteq Dom \times Dom$, it holds that:

for all $A', B' \in Con$:

2. $\langle \mathcal{N}(\overline{A'}), \mathcal{N}(\overline{B'})\rangle \in F_{lf}$ iff $Val_{RST(F_{lf}),\sigma}(\S(\overline{A'}) = \S(\overline{B'})) = 1$.

Finally, we are able to define for $\langle x, y\rangle \in Dom$:

**Definition 40.** $\langle x, y\rangle$ is **true** in $Dom$ iff $\langle x, y\rangle \in F_{lf}$.

By point (1.) of the previous theorem, this definition is adequate with respect to the members of $F_{lf}$, namely with respect to the pairs of individuals that depend on the set of grounded identity facts.
One important result, concerning in particular abstraction principles, is the following.

**Theorem 4.3.1.** Let $F_{lf} \cup F_{lf}^{-} = E_{lf}$, where $E_{lf} \subseteq Dom \times Dom$. Then, $\mathcal{M}[F_{lf}], \sigma$ is a model of any instance of any abstraction principle of the form:

$$\S x.\varphi(x) = \S x.\psi(x) \longleftrightarrow \Phi[{}^{\varphi(x)}/_{F}, {}^{\psi(x)}/_{G}] \qquad (\Sigma\mathsf{A})$$

*Proof.* Suppose first $R$ is an equivalence relation on $Dom$, encapsulated by the second-order formula $\Phi[{}^{\varphi(x)}/_{F}, {}^{\psi(x)}/_{G}]$. According to our model, $R$ settles the interpretation of "=": we will indicate this fact with the subscript $RST$ (reflexive-symmetric-transitive closure). Recall that the variable-assignment $\sigma$ assigns each second-order variable $F, G \in Con$ a first-order definable formula $\varphi(x), \psi(x)$ of $\mathscr{L}$. (1) Assume:

$$\mathcal{M}[F_{lf}], \sigma \vDash \S x.\varphi(x) = \S x.\psi(x).$$

This means that:

$$Val_{R,RST(F_{\alpha}),\sigma}(\S x.\varphi(x) = \S x.\psi(x)) = 1,$$

with $\alpha < \lambda$ (for $\lambda$ limit ordinal).
By a previous definition, this implies that:

$$\langle \mathcal{N}({}^{\varphi(x)}/_{F}), \mathcal{N}({}^{\psi(x)}/_{G})\rangle \in F_{\alpha}.$$

By the notion of dependence, we may imply:

$$Val_{R,F_{\beta}}(\Phi[{}^{\varphi(x)}/_{F}, {}^{\psi(x)}/_{G}]) = 1, \text{ for some } \beta < \alpha.$$

Hence:

$$\mathcal{M}[F_{lf}], \sigma \vDash \Phi[{}^{\varphi(x)}/_{F}, {}^{\psi(x)}/_{G}]$$

(2) By the same reasoning, assume:

$$\mathcal{M}[F_{lf}], \sigma \nvDash \S x.\varphi(x) = \S x.\psi(x).$$

This means that:

$$Val_{R,RST(F_\alpha),\sigma}(\S x.\varphi(x) = \S x.\psi(x)) = 0,$$

with $\alpha < \lambda$ (for λ limit ordinal).
By a previous definition, this implies that:

$$\langle \mathcal{N}(^{\varphi(x)}/_F), \mathcal{N}(^{\psi(x)}/_G)\rangle \notin F_\alpha$$

By the notion of dependence, we may imply:

$$Val_{R,F_\beta,\sigma}(\Phi[^{\varphi(x)}/_F, {}^{\psi(x)}/_G]) = 0, \text{ for some } \beta < \alpha.$$

Hence:

$$\mathcal{M}[F_{lf}], \sigma \nvDash \Phi[^{\varphi(x)}/_F, {}^{\psi(x)}/_G].$$

Terminating, thus, the proof. ■

### 4.3.3 Hume's Principle, Basic Law V and Grounded Concepts

Now, in order to see the formal machinery we've just proposed at work, consider the following example concerning HP.[32]

**Example** (Hume's Principle)**.**

$$\#F = \#G \longleftrightarrow F \sim G$$

$Dom = \{0, 1, 2, 3, \ldots, \aleph_0\}$, $Con = \{\varnothing, \{0\}, \{0,1\}, \{0,1,2\}, \ldots, \mathbb{N}\}$
$\mathcal{N}: \ \varnothing \mapsto 0, \{0\} \mapsto 1, \{0,1\} \mapsto 2, \{0,1,2\} \mapsto 3, \ldots, \mathbb{N} \mapsto \aleph_0$
$\Phi[F, G]$ : $F$ is equinumerous to $G$

We collect the pairs into our $E-$hierarchy of grounded identity/difference facts:

$E_1$ : $\langle 0,0\rangle$, $\langle 1,1\rangle$, $\langle 2,2\rangle$, ..., $\langle 0,1\rangle$, $\langle 0,2\rangle$, $\langle 0,3\rangle$, ... $\langle 1,0\rangle$, $\langle 2,0\rangle$, $\langle 3,0\rangle$, ..., $\langle \aleph_0, \aleph_0\rangle$, $\langle 0, \aleph_0\rangle$, $\langle \aleph_0, 0\rangle$

$E_2$ $\langle 1,2\rangle$, $\langle 1,3\rangle$, $\langle 1,4\rangle$, ..., $\langle 2,1\rangle$, $\langle 3,1\rangle$, $\langle 4,1\rangle$, ..., $\langle 1, \aleph_0\rangle$, $\langle \aleph_0, 1\rangle$

$E_3$ $\langle 2,3\rangle$, $\langle 2,4\rangle$, $\langle 2,5\rangle$, ..., $\langle 3,2\rangle$, $\langle 4,2\rangle$, $\langle 5,2\rangle$, ..., $\langle 2, \aleph_0\rangle$, $\langle \aleph_0, 2\rangle$

⋮

[32] We take Leitgeb's example, see Leitgeb 2017, pp. 277–279.

Here, our abstraction principle settles all the identity/difference facts for abstracta taken from the concepts in $Con$. For instance, $\langle 0,0\rangle \in E_1$ since $Val_{R,\sigma}(\Phi[\overline{\varnothing},\overline{\varnothing}])=1$, for any $R$. $\langle 0,1\rangle$ is in $E_1$, and $Val_{R,\sigma}(\Phi[\overline{\varnothing},\{\overline{0}\}])=0$, since a non-empty set $\{\overline{0}\}$ cannot be equinumerous to the empty set. Once the identities for pairs in $E_1$ have been settled, for instance, $\langle 1,2\rangle$ will appear at the next stage, in $E_2$, since – indeed – $Val_{R,\sigma}(\Phi[\overline{\{0\}},\overline{\{0,1\}}])$ depends only on the differences and identities that have been already established at the previous stage, in $E_1$. That is the identity/difference facts that have been determined at $E_1$ determine further identities or differences between pairs at next stages of the $E-$hierarchy.
Now, by isolating the corresponding identity facts (namely di $F-$sets), we indicate them by double-underlining their instances:

$E_1$ : $\underline{\underline{\langle 0,0\rangle}}$, $\underline{\underline{\langle 1,1\rangle}}$, $\underline{\underline{\langle 2,2\rangle}}$, $\ldots$, $\langle 0,1\rangle$, $\langle 0,2\rangle$, $\langle 0,3\rangle$, $\ldots$ $\langle 1,0\rangle$, $\langle 2,0\rangle$, $\langle 3,0\rangle$, $\ldots$, $\underline{\underline{\langle \aleph_0,\aleph_0\rangle}}$, $\langle 0,\aleph_0\rangle$, $\langle \aleph_0,0\rangle$

$E_2$ $\langle 1,2\rangle$, $\langle 1,3\rangle$, $\langle 1,4\rangle$, $\ldots$, $\langle 2,1\rangle$, $\langle 3,1\rangle$, $\langle 4,1\rangle$, $\ldots$, $\langle 1,\aleph_0\rangle$, $\langle \aleph_0,1\rangle$

$E_3$ $\langle 2,3\rangle$, $\langle 2,4\rangle$, $\langle 2,5\rangle$, $\ldots$, $\langle 3,2\rangle$, $\langle 4,2\rangle$, $\langle 5,2\rangle$, $\ldots$, $\langle 2,\aleph_0\rangle$, $\langle \aleph_0,2\rangle$

$\vdots$

Here, $E-$sets determine dependences among their members, while the identity relation is determined by proceeding through the $F-$sets. We have already seen that once in $E_1$ identities or differences between pairs of individuals, namely 0 and 1, have been settled, $\langle 1,2\rangle$ will be determined at the stage further. Since, $Val_{R,\sigma}(\Phi[\overline{\{0\}},\overline{\{0,1\}}])=0$ and we know, from the previous stage, that the pre-object 0 is different from 1 with respect to $F_1$, it follows that $\langle 1,2\rangle$ is a pair of distinct object. This, simply, means that $\langle 1,2\rangle$ is not in $F_1$. Likewise, e.g., at $E_2$ we have established the difference between 1 and 2 and, since $Val_{R,\sigma}(\Phi[\overline{\{0,1\}},\overline{\{0,1,2\}}])=0$, the pairs of individual $\langle 2,3\rangle$, appearing in $E_3$, are distinct objects and, therefore, are not members of $F_2$. Notice that our initial assumption, that our second-order quantifiers respect the identity relation that each model establishes, is here verified. Identity relations, indeed, are to be understood as equivalence relations, compatible with second-order quantification: there is no possibility for our models to establish the equinumerosity, for instance, between $\overline{\{0\}}$ and $\overline{\{0,1\}}$ when the model treats 0 and 1 as distinct. Viceversa, we cannot count $\overline{\{0\}}$ and $\overline{\{0,1\}}$ as not equinumerous, if the model treats 0 and 1 as identical. Trivially, 0 is identical to 0 and, therefore, the equinumerosity relation between $\overline{\varnothing}$ and $\overline{\varnothing}$ is determined. Differently, 0 is different from 1 and, hence, $\overline{\varnothing}$ cannot be established as equinumeorus to $\overline{\{0\}}$.

In our example, all the pairs of the form $\langle x,x\rangle$ are to be considered as identity facts. Now, at the last stage $\omega$ all the pre-objects have "become" proper abstract objects, with respect to their identity or difference as given by $F_{lf}$. In this sense, every subset of $Dom$ turns out to be grounded in a sense that, now, we may state explicitly:

**Definition 41** (Groundedness)**.** For $y \in Dom$, $A \in Con$, $R \subseteq Dom \times Dom$, such that $\mathsf{y} = \mathcal{N}(A)$, we say that

A set $X \in Con$ is **grounded** iff:

$$Val_{R,F_{lf},\sigma}(\Phi[\overline{X}, \S(\overline{A})]) = 1$$

In other words, a subset of *Dom* is grounded just in case it stands in a $\Phi-$equivalence relation with an $\mathcal{N}-$pre-object of *Dom*, that has become a proper abstract object at the stage $\alpha$ where $\Phi$ is evaluated by $Val_{F_{lf},\sigma}$. In our example, all concepts are grounded, since any of them stands in an equivalence relation with some (pre-)object: $\overline{\varnothing} \sim 0, \overline{\{0\}} \sim 1, \overline{\{0,1\}} \sim 2, \ldots, \overline{\mathbb{N}} \sim \aleph_0$. In this example, likewise, all subsets of *Dom* are grounded, since anyone of them is equinumerous with a concept: $0 \sim \overline{\varnothing}, 1 \sim \overline{\{0\}}, 2 \sim \overline{\{0,1\}}, \ldots, \aleph_0 \sim \overline{\mathbb{N}}$. Thus, in this framework, every pre-object $x \in Dom$, determined by a member $A$ of *Con*, such that $\mathsf{x} = \mathcal{N}(A)$, has, finally, "become" a proper abstract object. In other words, at the last $\omega-$stage every pre-object of *Dom* will turn into a proper abstract object and every concept will be grounded with respect to $Val_{\sim,F_{lf},\sigma}$.
In this sense,

1. We determined identity and difference facts between pre-objects (or labels) along well-ordered stages;
2. The abstraction map $\mathcal{N}$ and $\Phi$ determined the set of proper abstract objects, for which all identity and difference facts have already been been settled at the fixed point stage;
3. The concepts from which we abstract the proper abstract objects are grounded in the sense that they stand in a $\Phi-$equivalence relation to any of the former concepts at the fixed point stage;
4. Impredicative abstraction principles are understood, thus, in a well-founded fashion and, indeed, what we are calling "abstraction process", looks more alike an "individuation process". Let's be clearer: if the process, which starts always with the emptyset, reaches some pair of individuals $\langle x, y\rangle$ of labels, then they're reached at some well-determined ordinal stage, which, in turn, is completely determined by identity and difference facts settled at stages before.
5. Our pair of individuals $\langle x, y\rangle$ of labels can be reached by iterated application of an abstraction principle from the emptyset just by involving identities and differences between $x$ and $y$.
6. If an abstraction principle leaves indeterminate the question of whether $x$ and $y$ are identical, always starting from $\varnothing$, is a point we will discuss soon[33].

[33] There are different options either invoking classical logic or employing three-valued models for second-order languages. See Leitgeb 2005, p. 171 and Leitgeb 2017, p. 281 – this latter regarding exactly the case of abstraction principles.

It is interesting to notice that HP, when added to second-order logic, gives raise to what we nowadays call "Frege's Arithmetic". Many scholars, since C. Parson's article (1965) appeared and conjectured that the banning of BLV in favour of HP, could restore – even if in an altered fashion – Frege's *Grundgesetze*, added (i) a non-logical symbol, #, to their vocabulary, that applies to second-order variables and (ii) HP as an axiom. The main result has settled positively Parson's conjecture[34] and the derivation of the Dedekind-Peano axioms for arithmetic from second-order logic (+ HP) has shown successful. Moreover, unlike BLV, HP is not an inflationary principle and, hence, second-order full comprehension is admitted. In the following decades, other Fregean studies emerged and, in particular, by modifying the setting in which BLV is evaluated or by modifying BLV itself (in the sense of considering either its axiomatic or its schematic version), some nice conclusions have been drawn.
In this spirit, our intent is limited in showing how our second-order model – which is based upon a non-standard semantics – yields a model for BLV. The foregoing example concerning HP served as "informal" description of how our framework is supposed to work. In turn, consider the second-order language to which the non-logical symbol, $\epsilon F$, is added and which has to be read as "the extension of $F$". Recall that our $F, G \in Con$ were defined as ranging over first-order definable formula of our first-order $\mathscr{L}$. In order to be clear, if our second-order domain would be the entire powerset of the domain, in what follows, we would not need to take care about the distinction between the axiomatic and the schematic version of BLV:

$$\epsilon F = \epsilon G \longleftrightarrow [F \equiv G] \tag{BLV$_A$}$$

$$\epsilon x.\varphi(x) = \epsilon x.\psi(x) \longleftrightarrow \forall x(\varphi(x) \equiv \psi(x)) \tag{BLV$_S$}$$

Instead, since we have posed some restrictions, namely by allowing comprehension just with respect to the members of $Con$, then it is important to establish which formulation to analyse. Indeed, in order to avoid paradoxical conclusions, we restrict the following results to the axiomatic version of BLV. In order to continue remember that a variable-assignment of a second-order variable outputs a first-order formula of our domain. In this sense, if $\sigma(G) = \varphi(u)$, and $\epsilon\overline{G}$, then, according to $\sigma$, it may be concluded $\epsilon u.\varphi(u)$. First,

**Theorem 4.3.2.** $\text{BLV}_A$ has grounded instances.

In order to be clear, we want to show that there are some grounded instances for which $\text{BLV}_A$ comes out true.

*Proof.* First of all, consider that $\Phi[\overline{F}, \overline{G}]$ encapsulates the equivalence relation $R$ of material equivalence between the variable-assignment of $F$ and $G$.
Pick a formula $\varphi(u) =_{\text{def}} (u = u)$ and let $\sigma(F) = \varphi(u)$. By logic alone, we establish that $\varphi(u)$ is materially equivalent to itself. Hence, for some ordinal $\alpha$:

[34] See, among others, Zalta 2018, p. 29.

$$Val_{R,\sigma,F_\alpha}(\Phi[\overline{F},\overline{F}]) = 1.$$

In other words, the material equivalence between a concept $F$ and itself, in this context, is positively evaluated and hence, the identity between the extension of $F$ with itself may be easily checked.
Aside the determination of $\Phi$, recall that $R$ settles also the interpretation of =, and, given $\text{BLV}_A$ and our variable-assignment $\sigma(F)$, we may infer:

$$\epsilon u.(u = u) = \epsilon u.(u = u)$$

Thus, for the pairs $\langle \epsilon\overline{F}, \epsilon\overline{F}\rangle$ being in some set $F_\alpha$ of the grounded hierarchy $F_{lf}$ (as settled by the equivalence relation), this implies that the $RST$ value of $\epsilon(\overline{F}) = \epsilon(\overline{F})$ is 1. ■

Thus, the label $[\epsilon u.(u = u)] = x$ – identified with the extension of the objects identical with themselves – turns as proper abstract object, $\mathsf{x} = [\epsilon u.(u = u)]$, at the very ordinal stage when evaluated by the $\Phi-$equivalence relation, in this case, with itself. Therefore, the property $F$ defined by the first-order formula $u = u$ is a grounded instance of $\text{BLV}_A$. Henceforth, a more general result, concerning our study of $\text{BLV}_A$ may be established.

**Theorem 4.3.3.** $\mathcal{M}[F_{lf}], \sigma$ is a model for $\text{BLV}_A$.

*Proof.* We want to show that for some variable-assignment, $\sigma$, such that $\sigma(F) = \varphi(x)$ and $\sigma(G) = \psi(x)$:

$$\mathcal{M}[F_{lf}], \sigma \vDash \epsilon x.\varphi(x) = \epsilon x.\psi(x) \longleftrightarrow \Phi[{}^{\varphi(x)}/_F{}^{\psi(x)}/_G]$$

Suppose that $\Phi[F, G]$ is a second-order formula determined by the equivalence relation $R$. In this case, we have to consider material equivalence. Assume:

$$\mathcal{M}[F_{lf}], \sigma \vDash \epsilon x.\varphi(x) = \epsilon x.\psi(x)$$

This means that, for some ordinal $\alpha$:

$$Val_{\sigma,RST(F_\alpha)}(\epsilon x.\varphi(x) = \epsilon x.\psi(x)) = 1.$$

By dependence, then we may imply for some ordinal $\beta < \alpha$:

$$Val_{R,\sigma,F_\beta}(\Phi[{}^{\varphi(x)}/_F{}^{\psi(x)}/_G]) = 1.$$

Hence,

$$\mathcal{M}[F_{lf}], \sigma \vDash \Phi[{}^{\varphi(x)}/_F{}^{\psi(x)}/_G]$$

For the other direction, simply invert the starting point. ■

**Further Researches.** Abstractionism and the semantic, mathematical and philosophical studies of abstraction principles are one of the most debated and fruitful perspectives within the philosophy of mathematics. Indeed, the literature surrounding abstractionism is very vast and, indeed, we claim that what we've produced within this chapter should be deepened – also with respect to other positions and contributions, different from the one quoted in this chapter. Indeed, what we claim is that, our sketchy considerations can be even consolidated by explicitly studying the strength of our model. In this sense, – by having a second-order Henkin semantics, plus impredicative abstraction principles, – we would like to study how much of the so-called **Frege's Arithmetic** can be secured within our model and, consequently, which are the divergences between ours and other approaches. Predicative second-order systems, to which Basic Law V is added as an axiom, have been studied from different perspectives and, interestingly, many of them have been shown as strong as Q, i.e., as Robinson's arithmetic[35]. So, for what concerns the proof-theoretical strength and the possibility of encapsulating Frege's Arithmetic, we reserve these inquiries to a second moment of the present analysis.

**Back to Platonism!** With respect to the background logic, our attempt has consisted in limiting, in a Henkin way, the Comprehension Schema, by allowing second-order quantification just within a specific subset of the powerset of our domain. With respect to abstraction principles, we've allowed them being impredicative and studied their consequent "logical" behaviour. Finally, we've sought a model-theoretical construction for which the impredicativity of any abstraction principle gets disentangled in a well-founded and well-individuated manner – thanks to which, the axiomatic version of Frege's unfortunate **Basic Law V** gets a model. So, as remarked several times our study has focused upon truth-theoretical considerations, i.e., we have mainly worked within semantic abstractionism, and, in this sense, we have no clear "picture" of the main features of the "objects" that our abstraction principles introduce in a well-founded manner. In this sense, we think it's useful – for the moment – to pause our formal and semantic study of abstraction principles and trying to justify how our considerations, according to us, can render justice to a Gödel-like philosopher – allowing us, anyway, to prefer an ontologically and epistemologically moderate Platonistic interpretation of mathematics[36].

[35]For an introductory overview on such a formalization of arithmetic see also our Chapter 1. For similar remarks see, Linnebo 2004, Heck 2011 and Linnebo 2017.

[36]For the main differences we think subsist between a moderate and an extreme Platonistic philosopher of mathematics see our Chapter 1. In that context we've also sketched our first reflections on why we think "moderatism" is a better way to traverse. For a deepening of the comparison see also Chapter 2, last section "Lightened or Heavy Platonism?". In what follows (Chapter 5), we are going to sum up our philosophical observations, by trying to give a plausible (moderate Platonistic) interpretation of mathematics.

# Chapter 5

# Philosophy of Abstraction and Platonism

**Overview.** In our previous discussion we have seen how the model sketched in Leitgeb 2017 can be adapted in order to give a model for $\mathrm{BLV}_A$. Our reasoning has involved many features, such as the notions of identity, of grounding, and the usage of (Henkin) second-order models, which should be discussed more deeply from a philosophical point of view.
First of all, let's consider that we have not proposed a theory which derives the Dedekind-Peano axioms from logic alone, but we have simply proposed a formal model in which precise abstraction principles may be evaluated. In this spirit, it is useful to point out that the abstract (pre-)objects have no characterization yet and that our logical machinery does not tell us anything about their main features. In this sense, by a very simple inspection, it is clear that our model cannot directly clarify ontological and epistemological issues concerning abstracta in general, but, we claim, this can be done secondly and by considering our model as a starting point for philosophical debates. Indeed, since our main purpose was that of evaluating the tenability of ontologies as based on abstract objects (introduced thanks to abstraction principles), we will restrict our focus on this point.
Recall that Gödel-like philosophers held that abstract object may fit well within an ontology of mathematics and that we apprehend all of them thanks to a incomplete intellectual perception. We have taken seriously Gödel's platonistic ontological position, especially trying to circumvent Benacerraf's main criticism (1965)[1]. As we have seen, Benacerraf misunderstood the difference between identity statements and faithful representations. In the foregoing chapter, indeed, it was argued that while $\{\{\varnothing\}\} \neq \{\varnothing, \{\varnothing\}\}$, both can, anyway, viewed as reducible to the same natural number, 2, in this case, just by considering a precise equivalence relation.
Here's the fundamental role of our formal model: things are not so general as Benacerraf had desired and, indeed, we have established a way in which equivalences relations

[1]See chapters 1-3.

determine identity statements without violating Leibniz's Law of Indiscernibles. We reason in the following way:

a. Pick an equivalence relation between the concepts from which you want to abstract.

b. Construct the hierarchy of grounded identity/difference facts as established by the equivalence relation.

c. Isolate the identity facts.

Hence, of course, Benacerraf, while understanding the identity relation as the starting point of his considerations was brought to affirm that no set-theoretical reduction can accomplish the work of representing univocally a natural number, we have introduced a step earlier. Start with equivalence relations and let them determine the interpretation of the identity symbol.
As it already should be clear, in order to do so we have employed a lot of notions, which might well be objected. Indeed, consider the following questions:

1. What is an abstract object?
2. How does our formal machinery address the so-called *bad-company* objection?
3. Are we closer or even farer from giving a plausible answer to epistemological concerns (such as to the much debated *Julius Caesar* problem)?

We think that trying to restrict our focus on these three questions, our reasoning might reveal other (positive or negative) details of an abstractionist program. Moreover, we believe that some of ours Gödel-like doubts, especially epistemological, may be addressed in a way which get rid of any sort of neural or psychologistic process.

## 5.1 Towards a Metaphysical Characterization of Abstract Objects

### 5.1.1 The "Parallelism" Thesis

In developing our formal model we have fixed a domain of abstract (pre-)objects by defining these items as the entities which, once our evaluation process has terminated, may become proper abstract objects. An equivalence relation that operates on concepts, additionally, settles the interpretation of the identity sign. So, for instance, in more fregean terms, if some objects fall under a concept $F$ and some other objects fall under concept $G$, and we are able to establish a biunivocal correspondence between the two concepts involved, then the identity between the number of the $F$s and the $G$s immediately follows. Henceforth, by following our reasoning, the abstract pre-object representing the number of the $F$s (or the number of the $G$s), namely $\#F$ (or $\#G$), "becomes" a proper abstract object. In other words, the main difference occurring between pre- and proper abstract objects is that the latter has been given precise and definite identity conditions. Aside the epistemic problem

(how do we apprehend abstracta?), we have to care one moment on some ontological concerns. Let's recall that some paragraphs ago we've sketched one of the main differences lying between Gödel-like philosophers and some contemporary forms of neo-fregeanism: shortly, while the former invoke a "strong" conception of existence, the latter, instead, argue in favour of a "light" (or "thin") conception of being. In this spirit, since our final objective in this chapter is both, ontological/metaphysical and epistemic, we believe the an investigation within the conception of existence, involved by mathematical abstract objects, may reveal the best route in order to provide tenable epistemological considerations.
Consider again, for the moment, our Gödel-like philosopher of mathematics and the following claim:

**(Par)** Mathematical abstract objects exist and describe a non-sensory reality, just as physical bodies and scientific theories do with respect to our empirical reality.

Let's call **(Par)** the *parallelism thesis*, which we've already introduced in the first chapter. Recall, additionally, that what we've called Extreme Platonism had **(Par)** as one of his leading premisses. Indeed, an inquiry within the *admissibility* of such a parallelism should be conducted, as we think that one of the main characters of mathematical abstract objects is that their existence is «less demanding»[2] on the world, then the existence of physical bodies. In other words – we argue – the tenability of **(Par)** is a misleading route to understand ontological and metaphysical questions concerning abstract objects.
But – let's think for a moment – **1.** if we get rid of **(Par)**, from where should we start understanding the main ontological characters of our mathematical abstract objects? Additionally – begin asking yourself – **2.** if the parallelism thesis is false and we believe in the existence of abstract objects (at least, for an ontology of mathematics), then any argument which tries to compare scientific/empiric perception to mathematical/intellectual perception, has to be refused?
**1.** In order to shed light on our point of view, re-consider the following abstraction principle:

$$\#F = \#G \longleftrightarrow F \sim G \tag{HP}$$

Now, suppose that $F =_{\text{def}}$ "the pens on the table" and $G =_{\text{def}}$ "the bottles in the fridge". If we are able to verify the equinumerosity ($\sim$) between the $F, G$, that is we are able to establish a one-to-one correspondence between the two concepts considered, then the identity between the numbers of the elements falling under $F, G$ is identical. Hence, the number of the pens equals the number of the bottles, in our case. The objects standing on the LHS of the biconditional – the numbers – are proper abstract objects, whose identity conditions – as settled by our informal example – grant their existence and, moreover, we claim, that those identity conditions are

[2]Expression taken from Linnebo 2018.

their unique "existence conditions". This latter point explains precisely the meaning of the concept of *thin existence* that we've already simply cited. At this point, then, it should already be clear why we are refusing **(Par)** from our ontological considerations: we believe that something as numbers, sets and lines exist, but that their existence has to be characterized very differently and in a less demanding way than that of physical bodies. Recall, now, the general working of our formal second-order model. Thanks to the grounded hierarchy we have been able to isolate the concepts rightly standing into an equivalence relation each to the other and, hence, to "transform" (metaphorically speaking) into proper abstract objects, with definite identity conditions. The strategy has been that of letting the equivalence relation determine the identity symbol. Reconsider our example: once the elements of the table can be put into a biunivocal correspondence with the elements in the fridge, then the identity conditions for the (abstract) objects representing the number of those elements follow in a grounded way.

This is, in addition with some gödelian features, our archimedean starting point: abstract mathematical objects are not concrete, not spatio-temporally located and, moreover, their existence is grant by their grounded identity conditions been given. This latter point will shed more light on the conception of light existence we have in mind and, secondly, will conduct us directly to epistemology[3].

First of all, consider the following reflection:

> The vast ontology of mathematics may well be problematic when understood in a thick sense. If mathematical objects are understood on the model of, say, elementary particles, there would indeed be good reason to worry about epistemic access and ontological extravagances. But this understanding of mathematical objects is not obligatory. If there are such things as *thin* objects, then the existence of mathematical objects need not make much of a demand on the world. [...] More generally, the less of a demand the existence of mathematical objects makes on the world, the easier it will be to know that the demand is satisfied[4].

But, what does it actually mean that "thin objects are less demanding on our worlds than concrete ones"? Are there just abstract and concrete objects or – as suggested by C. Parsons – a third mid-way may be conceived?

Recall, for example, the fridge containing some determined objects. Now, if we pick some object $b$ in the fridge, say, the bottle of water, then, by inspection, we see that it occupies a specific portion of space and a specific region of time. Let's say that our bottle $b$ is divided into spatiotemporal parts $b =_{\text{def}} b_1, b_2, \ldots, b_n$. Moreover consider that any $b_i$ must be necessarily part of the entire material that we call bottle and that we identified with $b$. Hence, a concrete object, such as $b$, makes a substantial demand on our world, since – to exist – it has to occupy a determined space-time region. In other words, objects that "demand" to occupy a determined region of

[3]Notice that by asserting that **(Par)** is false and by assessing a conception of thin existence, we are somehow aiming not to invoke any sort of mathematical and/or intellectual perception, whatsoever meant. We will discuss this point soon.

[4]Linnebo 2018, p. 9. Our emphasis.

space-time, exist in a way to be conceived as "thick" or "robust". Differently, if we say that numbers, figures and sets are "less demanding" on the world, we mean that their existential status is characterized by ontological "thinness". Consider, for the sake of the argument, HP that allows us to define "numbers as collections of objects". Now, let the collection of bottles in the fridge, say $B$, be the collection of (concrete) objects that share the "bottle-being" property. Now, consider the collection of eggs (i.e., of concrete objects that share the "egg-being") and let's call it $E$. Suppose that we are able to establish a biunivocal correspondence between $B$ and $E$, letting, thus, the two sets be equipotent. According to HP, then, it follows that the number of $B$ is identical to the number of $E$. By existential generalizations, finally, it is possible to conclude that something as the number of a collection (represented by a concept) of things, exists.

As it should already be clear – we claim –, our strategy settles the existence of abstract objects, such as the natural numbers (as introduced by HP), by not "occupying" any region of space-time. Indeed, the abstract object, identified with the natural number representing the bottles (or eggs) in the fridge, «is far less demanding than what we found in the case of physical bodies», and, in «general, abstract objects are thinner than concrete objects because they do not make demands on any particular region of spacetime»[5]. By inspection, our second example, actually, may be generalized by saying that any well defined collection of objects provides a referent for the (abstract object) representing a natural number. In different terms, independently from our sets $B$ or $E$, numbers, meant as thin abstract objects, exist. Therefore, as I have said up to now, we may conclude that the first step in trying to furnish the ontological definition of thinness is obtained by excluding that abstract objects may occupy a portion of space and time.

### 5.1.2 The "Indefinite Extensibility" Thesis

The second step is linked to the previous, but, instead of focusing on the possibility, for abstracta, to occupy a region of space and time, it concerns the different notions of *determinacy* invoked by concrete and abstract objects. Another time, examples will shed light on what we mean. Suppose we possess an almost perfect definition of the concept of "liquid" and we assume that scientists won't betray us. The definition, for the details it contains, yields (or, should yield) the extension of the concept of "liquid". In other words, thanks to the definition of liquid we are able to settle all the concrete things which fit the definition of liquid. For instance, by a simple inspection, if I have in front of me a laptop, I won't never assert that it bears the "liquid-being" property, while, if I consider the material within my bottle of water, I will see that it perfectly falls under the concept "liquid". Hence, the *determination* of the concept of liquid is perfectly given by its extension, namely, once we have a clear idea of which concrete features belong to a liquid, then the world immediately furnishes an answer to the question concerning the liquid-existence. In other words – we claim –, by

[5] Linnebo 2018, p. 45.

comparing what scientist mean with liquid, to some concrete objects in the reality, then we can have an almost complete lists of possible referents for the concept liquid. Mixing both components together, we may say that thick existence (i) grants that concrete objects are spatiotemporal located and, (ii) that, this location grants their determinacy:

> For most, if not all, concepts of physical objects, the determinacy of the concept ensures the determinacy of its extension, given the usual input from reality[6].

Now, let's turn to mathematical abstract objects. We've said that, for instance, sets or numbers lack spatio-temporal connotations and that, this particular feature, enables us to introduce the first characterization of their thinness. The previous example – concerning thick objects – has shown that, by comparing determinate definitions to reality, a more or less precise list of concrete objects falling under the definition considered will arise. But, if an objects is not inserted in our spatio-temporal reality, how should we understand their determinacy? We claim, in accordance to Linnebo's line of argumentation[7], that mathematical objects are characterized by their *indefinite extensibility*. Consider, now, numbers as introduced by HP and recall that, by abstraction from collections of things, bearing a precise relation to one another, it is possible to form the number belonging to the collections of elements considered. Let's say that $a, b, c$ are the only objects falling under a concept $P$ and that $\#P$ is the number of the concept considered. Let $a', b', c'$ other three objects falling under another concept, say $Q$ and let $\#Q$ be as before. By HP, since, $P \sim Q$ holds, then $\#P = \#Q$ holds too. Using our common number-sign we may therefore conclude $\#P = \#Q = 3$. By inspection of the definition of liquid, for instance, and thanks to the input coming from the world, it was noticed that it is (or should be) possible to settle down all the objects that fall under the "being-a-liquid" property. Following this line of thought, it is natural to ask whether mathematical concepts possess the same extensional determinacy as physical objects do. We claim that, any sort of collection can form the number of its elements and, hence, even if we have an almost clear way to form number statements, we cannot have their complete extension: «[f]or any plurality of instances, we can use this plurality to define a new instance of the concept, that is, an instance that is not member of the given plurality»[8]. That is, by having many instances, for example, of sets characterized by the presence of 3 elements, we can use the totality including all of these 3-sets as a new number object. In other words, mathematical abstract objects are indefinitely extensible: given some acceptable instances of some abstract object, then, by collecting the totality of those instances, we do not obtain – as in the case of liquids – the almost complete list of all objects sharing their "being-a-liquid", but what we get is another mathematical abstract object. This is exactly what we mean by the terms "having no definite extension". In other words, for example, the totality of the "3-sets" is not the entire

[6] Linnebo 2018, p. 191.
[7] Linnebo 2018, pp. 189–192.
[8] Linnebo 2018, p. 191.

list containing the collections sharing 3 elements, but it is the new formed abstract object, identified with the number representing the totality of “3-sets”. Let’s call the totality of 3-sets $S$. If $S$ would have a determined extension, then we would able to produce an almost complete list of 3 elements collections, but – as we have claimed – this totality does not determine the extension of 3, in our case, but determines another number that can be defined, just by starting from the totality $S$, namely $\#S$. The situation we are sketching can be put graphically as follows:

$$S \underbrace{\begin{cases} \#A & \#B & \#C & \#D & \dots \\ \{a, a', a''\} & \{b, b', b''\} & \{c, c', c''\} & \{d, d', d''\} & \dots \end{cases}}_{\#S}$$

Thus, we have a clear way in which collections of objects form numbers (meant as abstracta) and, moreover, we’ve understood that, despite this clearness in forming number-like abstract objects, they do not possess a determinate extension. Indeed, as showed by our example, from any totality (seen as the *indefinite* extension of any number-concept), such as $S$, it is always possible to build a new abstract object, namely the number of the totality of abstracta falling under the specific number-concept considered, that is $\#S$. In conclusion, concrete concepts do have, more or less, precise extensions, while abstract objects have indefinite ones.

### 5.1.3 Our Metaphysical Picture

This second point terminates our discussion of the tenability of **(Par)** – the *parallelism* thesis. At this point, indeed, some fundamental feature of our abstracta finally arise:

**AO1.** Abstracta lack spatio-temporal location.

**AO2.** Abstracta are indefinite extensible.

The combination of **AO1** with **AO2** corresponds to the first characterization we have provided for the sense of “thin existence” we are trying to defend. The first purpose, here, was, indeed, that of seeing whether the parallelism between physical bodies and abstracta hold. Our answer has been negative and we’ve proposed, in connection to **AO1**, to understand the abstracta existence as “poorly” demanding on the world. Secondly, we have seen that, by analysing the notion of “determinacy”, with respect to abstract and concrete objects, respectively, **AO2** follows.
Last point of this section. Our previous reasoning was involving two categories for objects (abstract and concrete) and we have said that something as numbers and sets do not share the same ontological features of tables and bottles. Moreover, we argued, for example, that a number (meant as abstracta) is less demanding on the world than the table (meant a concrete), and what we had in mind with this formulation is exactly embodied by **AO1** and **AO2**. The argumentation seems to fit

for all these objects which are *purely thin*: «an object is *pure abstract* if it lacks both spatiotemporal location and any kind of intrinsic relation to space and time. The natural numbers and pure sets are examples»[9].

But, if we consider letters or figures, then, by looking back at our everyday experience, we may observe that these entities – letters or figures –, generally conceived, lack spatiotemporal location, while some of their realizations are inserted in some space and in a specific time. For instance, while letters lack spatial and temporal relations, their concrete realizations – namely tokens – are in some space and time and, likewise, for geometrical figures. So, in the case of purely thin abstracta we can never encounter one of their concrete realizations (e.g., sets and numbers), resulting therefore as characterized by a very poor existential demand on the world. In the case of thick concrete objects, instead, we are in front of a very strong demand on the world. And, finally, as it should already be clear, objects such as figures and letters do not belong to none of the two preceding categories and, indeed, following some recent literature, let's introduce a midway between abstracta and concrete objects, i.e., *quasi-concrete* entities:

> The surprising discovery is that quasi-concrete objects are somewhat thicker than pure abstract objects. The existence of a quasi-concrete object makes a non-trivial demand on spacetime, however weak and indirect: there must be, or at least possibly be, concrete realizations of the object somewhere or other in space and time.[10]

*So far so good.* Recall our Quinean definition of ontology and metaphysics: roughly, the former inquires "what there is", the latter investigates "what is what there is". By starting from this subdivision, let's precisely summarize, in the next table, the way we are conceiving as a good route to construct an ontology – and consequently, a metaphysics – for mathematics.

| | **Ontology** | **Metaphysics** | **Specification** |
|---|---|---|---|
| Concrete objects | ✓ | Thick existence | Non-trivial, strong and direct demand on spacetime. |
| Quasi-concrete objects | ✓ | Thin/Thick existence | Non-trivial, weak and indirect demand on spacetime. |
| Abstract objects | ✓ | Thin existence | No demand on spacetime. |

Finally, the three-partition of objects – as sketched above – summarizes what we've

[9]Linnebo 2018, p. 45.
[10]Linnebo 2018, p. 45.

specified up to now. From an ontological point of view, abstract, quasi-concrete and concrete objects exist. From a metaphysical point of view, we've argued that different conceptions of existence could be traced. Indeed, thanks to the deny of the *parallelism thesis* **(Par)**, we have seen that, by considering the different possible demands that an objects has with respect to reality, three different sense of the word "existence" can be defined. In this, sense, even if mathematical abstract (or, quasi-concrete) objects exist, this does not mean that the way in which they exist is the same. In other words, we believe that there is no extravagance in asserting that physical bodies exist differently – with respect to their demanding on reality – from, for example, sets.

## 5.2 Platonism & Ontological Remarks II

The discussion concerning the existence of abstract objects has an ancient route and, indeed, many scholars track back the origin of this debate to Plato and to the school he founded. The influence of such a view – that postulated the existence of a realm of abstracta – has exercised much strength on the philosophy immediately succeeding Plato and, indeed, Platonism has remained as one of the prominent line of thoughts, not only in philosophy of mathematics, but in philosophy generally conceived. As R. Heck writes, «much of our ordinary discourse seems to involve reference to abstract objects»[11] and, indeed, our entire work has seriously taken Heck's affirmation. Anyway, it is useful to point out that – as always in philosophy – there is no general agreement for what concerns *the* right way to understand the ontology of abstract objects and, consequently, the debate can be regarded as overloaded. In any case, since the position we are trying to outline, has used some particular "techniques" or notions let's deepen our philosophical considerations. Recall that in the first section devoted to Gödel's metaphysics of mathematics, we've introduced Platonism as the conjunction the following two theses:

**P1.** Mathematical objects exist.

**P2.** Mathematical objects are abstract.

Now, as it should already be clear – even if with some changes in the general setting – our considerations are compatible with P1 and P2. Moreover, we've tried to accomplish a step further, by not simply assuming the existence of abstract objects, but, by trying to understand what we could have in mind while using terms such as "abstract" and "existence", always with respect to mathematics. Indeed, by considering our three-partition of objects, a more detailed specification, for what concerns P2, can be added:

**P2**$^{\star}$. Mathematical objects are either abstract or quasi-concrete.

Thus, the realm of mathematical entities is twofold covered. We claim, indeed, that both, abstract entities (such as numbers or sets) which lack any sort of spatio-temporal

[11] Heck 2017, p. 50.

relation, and quasi-concrete ones (objects which can have concrete, spatio-temporal located realizations) populate Plato's heaven.
In any case, some precisions are needed. The first part of this work ("Hunting abstract objects") has been devoted to the analysis of Benacerraf's "reductive" argument (1965). The second part – of which this chapter is the closing point – differently, has analysed a possible way of introducing abstract objects. Our intention was that of seeing whether the concept of "abstraction principle" (as developed by neo-fregeans) could have helped us in forming a clear picture of we mean by mathematical abstract object. The first of our concerns has been another time Quinean and, in this spirit, we have tried not to admit any entity, in our ontology, which lack identity conditions. This route has brought us in building a formal model capable of suggesting in which sense abstracta were introduced. Philosophically, we claim, the main advantage of the model we've built, has been that of giving us a general method to construct hierarchies of identity/differences facts, as settled by an equivalence relation. This has been done thanks to the definition of dependence we have assumed and stated already from the beginning of chapter 3. For example, suppose that the RHS of an abstraction principle holds for some concepts, that is, there is an equivalence relation positively occurring between these concepts. From the point of view of our model, these latter equivalent concepts determine the identity between the newly introduced objects. Hence, differently put, the (identical) objects standing in the LHS of an abstraction principle depend on the equivalence relation occurring between the concepts (in the RHS) and from which our new abstract objects have been abstracted. As it has been argued, the hierarchy proceeds in a well-founded manner and this proceeding grants that, whatsoever identity claim is made at some $n$-stage, then there is an equivalence relation, at some stage $n-1$, determining the further identities themselves. Indeed, the impredicativity of any abstraction principle gets disentangled exactly once a dependence-hierarchy may be established. Furthermore, this hierarchical strategy, when implemented with thin conceptions of existence, reveals an interesting point: before it was argued that some abstraction principles are untenable since the quantifiers they involve presupposes the totality on which it quantifies and, hence, some philosophers have argued in favour of predicative (or symmetric) abstraction principles. The main difference is that a predicative principle allows us to introduce objects which where not already included in the RHS and, thus, by not including the objects in the LHS within the range of the quantifier. Simply put, predicative principles are not presupposing. We see that the adoption of the distinction between thin and thick existence may furnish an answer to this problem and may reveal helpful in revaluating the value of impredicative principles: «The left-hand side of an abstraction principle makes demands on the world that go beyond those of the right-hand side. Thin objects are nevertheless secured because the former demands do not *substantially* exceed the latter. For the truths on the left are grounded in the truths on the right»[12]. In this sense, we have not provided an

[12] Linnebo 2018, p. 5.

answer to the tenability of asymmetric principles just from a model-theoretic point of view, but – and this is a fundamental point – we've argued in their favour, also and mostly, philosophically[13].

## 5.3 Groundedness, Impredicativity and "Bad Companions"

In the previous chapter, we have began our discussion by enumerating three of the main concerns that abstractionist philosophers usually face and tried to circumvent them by establishing some model-theoretical considerations. Recall that the first two problems we have exposed, are the impredicativity of abstraction principles and the so-called "bad company" problem. Roughly, their respective main claim can be rendered as follows:

**(Imp)** Impredicative abstraction principles are untenable, since the LHS of any of them presupposes the RHS.

**(BC)** Consistent abstraction principles are surrounded by "bad companions", i.e., unacceptable ones.

In what follows, we wish to discuss how our previous semantic work, accompanied by our metaphysical characterization of abstracta, can lighten both problems. First of all, the main idea of our semantic work has been that of developing a model in which the ontology of mathematics gets constructed through a well-determined process. This latter, more closely, has been developed as a well-ordered individuation of entities, where the notions of "dependence" and "presupposition" have played crucial roles. But, how does this process determine a lightening with respect to **(Imp)** and **(BC)**?
First of all, consider that – according to **(Imp)** – there is presupposition between the two sides of a biconditional expressing an abstraction principle and, in this sense, what is on their left side depends on what belongs to their right side. Our model, we think, by previously modelling the notion of presupposition, provides a way thanks to which impredicativity gets disentangled. Indeed, the entities that an abstraction principle introduces, are given just with respect to what comes before, or, in other terms, the *new* objects presuppose just what is available at the stages coming before their introduction. More specific, an abstraction principle allows us to introduce entities by specifying their identity conditions within the LHS, but, as based upon

[13] A significant distinction has to be made: in the whole of our work we were not aiming in proposing specific abstraction principles capable of deriving portions of mathematics (remarkable examples can be found in Zalta 1999, Zalta 2001 and Anderson and Zalta 2004; consider also Linnebo 2018) but, instead, our main objective was that of seeing whether the adoption of some formal framework, in which principles may be evaluated, could be useful in discussing some philosophical problems surrounding abstractionism.

the equivalence relation settled within the RHS. For clarity, consider always our $\Sigma$:

$$\S\alpha = \S\beta \longleftrightarrow R(\alpha, \beta). \qquad (\Sigma)$$

So, for example, let $P, Q \in Con$, pick two relations $R_1$, $R_2 \subseteq Dom$ and let them determine two second-order formulae with exactly two free variables, i.e., $\Phi_1[P, Q]$ and $\Phi_2[P, Q]$. If we see that the value of $\Phi_1[P, Q]$, as determined by $R_1$, is different from the truth-value of $\Phi_2[P, Q]$, determined by $R_2$, this means that the values of their corresponding abstract objects are different. In other terms, if the values of the two formulae $\Phi_1[P, Q]$ and $\Phi_2[P, Q]$ are different, then the two pre-objects, to which the abstraction map links the concepts involved, are likewise different. In this sense, different objects – with respect to $\Phi_1[P, Q]$ and $\Phi_2[P, Q]$ as determined $R_1$ and $R_2$ respectively – will turn out as proper abstract objects from the abstraction from $P$ and $Q$.

Importantly, if we recall our construction $E$, it is easy to check that we have imposed a **groundedness** condition, i.e., a restriction for which, the identity conditions between two entities of every stage of the grounded hierarchy of identity/difference facts $E$, presuppose exactly those, and just those, identity conditions between entities that were available at some stage earlier. The same applies to the set of grounded identity facts $F$ – as extracted from the $E$-hierarchy by considering just the identity statements between two pre-objects, this latter influenced by the Reflexive-Symmetric-Transitive closure between the concepts from which the (proper) abstract objects gets abstracted. In this sense, as it should already be clear, we have addressed the impredicativity problem for abstraction principles, not by banishing them from our philosophical attempts, but rather by trying to disentangle the charge of presupposition which affects them. Moreover, from the beginning of this work, we have tried to argue in favour of the presence of abstracta within an ontology of mathematics – now, we know, as introduced by abstraction principles and thanks to a groundedness condition – by trying to indicate and defend a (Light or Thin) Platonistic conception of their existence. Hence, in conclusion, we claim that, by seriously considering **(Imp)**, one might give raise to a construction in which the presupposition problem has not necessarily to be excluded, but gets modelled and encapsulated as a fundamental notion of our semantic study concerning abstraction principles.

The same elements can now be considered and adapted to the **(BC)** problem. Roughly, some abstraction principles are consistent and acceptable, while some others – which are, at least, intuitive – turn out as inconsistent and as unacceptable. This argument has led many philosopher to abandon and declare abstractionist approaches within the philosophy of mathematics as inherently mistaken and corrupt. Differently, we claim that our model-theoretic approach can also indicate an answer for what concerns **(BC)**, by, consequently, arguing in favour of the applicability of abstraction principles within a foundational abstractionist program. In this respect, a central role has been played by the groundedness condition we have imposed on our construction. Our main idea has been that of, this time, banishing the **(BC)** problem by indicating that *any* abstraction principle should be restricted to its grounded instances. What we actually had in mind was to furnish a clear device whose objective was that of withdrawing, in order to exclude, the distinction good *vs* bad abstraction principles. As clear, for instance, BLV[14] belongs to the so-called "bad companions", but – as previously showed – if restricted to its grounded instances, as suggested by us, it gets a model. As Leitgeb spells this situation out, «no good guys left, no bad company either»[15].

## 5.4 Epistemological Suggestions

In Chapter 1 we have analysed Benacerraf's problem from two separated, but connected perspectives. In his 1965 paper Benacerraf challenged any Platonistic position which introduces abstract objects to explain mathematical or scientific ontologies. We've argued that his premises – as majorly based upon set theory – can be weakened by carefully analysing the logic of his argumentation (Chapters 2-3). Moreover, starting from our attempt to undercut Benacerraf's ontological "elimination", we have tried to develop a method that allows us to introduce whatsoever sort of abstract object, especially thanks to a semantic study of the notions of "dependence" and "groundedness" (Chapter 4). Finally, the resulting construction we have proposed, as remarked several times, does not tell us anything about the abstract objects that abstraction principles allow us to introduce. This has led us in re-considering the tenability (of some form of) Platonism. With respect to this situation, then, we have tried to provide a first main clarification – whose objective was to give a first metaphysical and ontological characterization of the "existential" sense of abstract objects. Up to now, anyway, no explicit answer, for what concerns Benacerraf's[16] initial epistemic worries, has been given. In what follows, indeed, we wish to give some epistemological suggestions, which, we believe, fit in a coherent manner with our previous (semantic and metaphysical/ontological) considerations.

### 5.4.1 Frege, Logicism and Epistemology

In what follows, we wish to briefly discuss Frege's original logicism from an epistemological point of view. Recall that the main purpose of Frege's logicist attempt was that to give a clear and undoubted foundation of arithmetic, by deriving the most fundamental laws governing natural numbers from just logical definitions and notions. Additionally, ontologically, to secure reference to entities – such as concepts, extensions and so on – Frege posited a "third realm" of logical objects. In the whole of this context, if we were asked the motivation for which Frege developed such a program, then we were immediately led to his epistemic worries concerning logic and

[14] For limits and further researches with respect to the model of Chapter 4, see section "Hume's Principle and Grounded Concepts".

[15] Leitgeb 2017, p. 270.

[16] See Chapter 1, section "Against Platonism II: Epistemology". See also Benacerraf 1973.

mathematics. By describing his logicism, in his *Grundlagen*, Frege clearly wrote:

> The problem becomes, in fact, that of finding the proof of the proposition and of following it up right back to the primitive truths. If in carrying out this process, we come only on general logical laws and on definitions, then the truth is an analytic one, bearing in mind that we must take account also of all propositions upon which the admissibility of any of the definitions depends. If, however, it is impossible to give the proof without use of truths which are not of a general logical nature, but belong to the sphere of some special science, then the proposition is a synthetic one. For the truth to be a posteriori, it must be impossible to construct a proof of it without including an appeal to facts, i.e., to truths which cannot be proved and are not general, since they contain assertions about particular objects. But if, on the contrary, its proof can be derived exclusively from general laws, which themselves neither need or admit of proof, then the truth is a priori[17].

In this part of his *Grundlagen*, Frege was trying to distinguish what does it mean for a statement being or synthetic either analytic, and, consequently,what does it mean for a truth to be or a priori either a posteriori. First of all and from a general point of view, we have to consider that Frege – in a Kantian spirit[18] – divided scientific propositions in the following way:

- **Analytic Statements**. A proposition $p$ is said to be analytic if its truth depends just on a small package of undoubted primitive logical propositions.
- **Synthetic Statements**. A proposition $p$ is said to be synthetic if its truth does not follow from logical laws alone, but, depends on the laws derived from the objects of inquiry of some special science.

In the same vein as Kant, then Frege proceeded in defining the notions of truth a priori and a posteriori. Let's summarize the division as follows:

- **A Priori Truth**. A proposition $p$ is said to be a priori, if it is proved to be true just by involving general logical laws and any appeal to empirical evidences.
- **A Posteriori Truth**. A proposition $p$ is said to be a posteriori, if it is proved to be true thanks to the appeal to particular laws, established starting from the objects of inquiry of a particular science.

For instance, generally conceived, biological statements are synthetic and a posteriori, for their truth not depending just on the pure laws of thought, but following from the particular biological laws, i.e., those laws established thanks to observations and experiments. Likewise, geometry is synthetic, i.e., we do possess – according to Kant

[17]Frege 1884, p. 4.

[18]It is interesting to notice that I. Kant has reserved much space of his philosophical work to mathematics, see, for introductory remarks, Shabel 2013.

and Frege – something as "spatial intuition", at least, to see the truth of geometrical statements. But its truth has an a priori character: since the more sophisticated and complex areas of geometry treat spaces which are very hardly perceivable or empirically testable, we do possess – in addition to spacial intuition – the a priori concept of "all that is spatial" – as not limited to the perceivable or imaginable. For what concerns mathematics, the thoughts of Kant and Frege diverged. For the first one, mathematical statements have a synthetic, even if a priori, character, i.e, in Kant's view, in mathematics, we start by concrete representations of some objects (synthetic component), that we encapsulate in propositions which will be proved just thanks to the appeal to general logical laws (a priori component). Differently, for Frege, mathematical cognition does not involve any appeal to empiric representations, and, indeed, any mathematical statement $p$ gets its truth only on the basis of the meaning of the primitive logical terms occurring in it. Importantly, for Frege, while proving the truth of a mathematical proposition $p$, we do not consider anything except the logical terms and the general logical laws that govern them, and, therefore, – in a pretty non Kantian way – no synthetic component is required for mathematical statements to be true. In conclusion, so, in Frege's philosophical view, mathematics is analytic and a priori, and, exactly, in pursuing the objective of deriving the natural numbers laws from logic alone, − apart his foundational aim −, we believe, Frege was trying to prove that, at least, arithmetic, among all mathematical disciplines, is analytic and a priori:

> The gaplessness of the chains of inferences contrives to bring to light each axiom, each presupposition, hypothesis, or whatever one may want to call that on which a proof rests; and thus we gain a basis for an *assessment of the epistemological nature* of the proven law.[19]

So, in order to assess the "epistemological nature" of mathematical, or better arithmetical, propositions, Frege thought as necessary to show, in a clear and precise way, that all arithmetical laws are derived from logical laws. In other words, according to Frege, our knowledge and understanding of any mathematical proposition $p$, flows analytically from some basic and universally valid principles, these latter, considered as true a priori. Hence, general logical laws – in Frege's original project – have a twofold objective, namely that of providing an ontological and epistemological foundation of mathematics:

> He aimed to identify a few select general logical laws, or basic laws, that were needed to provide an epistemic foundation: namely, mathematical knowledge was meant to "flow" from those basic principle and (what is now called) second-order logic. In addition, Frege also thought of the basic principles as providing an ontological foundation. Basic Law V was meant to identify the logical objects (extensions) by means of which numbers could then be defined. For Frege, logic was the most general of all sciences

[19] Frege 1893/1903, p. XXVI. Our emphasis.

and concerned with the laws of thought[20].

As remarked several times, Frege's logicism failed and, consequently, also the project of giving a secure, at least, epistemic foundation of mathematics failed.
Several attempts – in the context of contemporary abstractionism – have tried to argue in favour of a renewed Fregean epistemology of mathematics. For instance, if we accept that HP plus second-order logic suffice to carry out the derivation of the Dedekind-Peano axioms for arithmetic, then we should also contemplate whether we are working with an analytic or a synthetic proposition, with an a priori or an a posteriori truth. Being very far from proposing an epistemology that argues in favour or against the epistemic tenability of some particular abstraction principles – such as BLV or HP –, we would like just to indicate some epistemological suggestions, which we think could fit with our Platonistic – ontological and metaphysical – conceptions of abstract objects as introduced by abstraction principles.

### 5.4.2 Abstract Objects: Benacerraf, Doxology and Epistemology

**Summary.** Some paragraphs ago we have argued that the existence of abstract objects has not, necessarily, to be conceived in a "hard" way – on the model of physical objects –, but, rather, we have suggested that this parallelism could be misguiding. What we meant, was exactly that, unlike physical ones, abstract objects, to exist, simply request their identity conditions being well-defined and specified. The demand on our reality, that an abstract object imposes, with respect to the demand of a concrete body, is – as argued – very poor. Additionally, we have reasoned that some kind of abstract objects – different from the "pure" ones – can have concrete realizations that we may encounter (e.g., letter and tokens). We have called the kind of existence that characterizes pure abstracta as light or thin. Our reasoning concerning the demand that some abstract (or quasi-concrete) objects have with respect to our reality, has led us in revaluating the function of our abstraction principles: what an abstract object demands for its existence to be ensured is the settling of its identity conditions. Consider again our Σ:

$$\S\alpha = \S\beta \longleftrightarrow R(\alpha, \beta) \qquad (\Sigma)$$

From a semantic point of view (Chapter 4),we have studied abstraction principles by trying to restrict their usage and application, to their "grounded instances". In other words, we have proposed a model in which the truth-value of the LHS of any Σ depends on the truth-value of its RHS. Now – by having explained how metaphysically the existence of abstracta may be characterized – we additionally know that the identity conditions that any Σ expresses in its LHS, are the unique existential condition requested for abstract objects to exist. In other words, an abstract object to exist requires simply that its identity conditions gets introduced

[20]Ebert and Rossberg 2007, p. 54.

in a grounded manner thanks to a well-defined Σ. This is the main reason for which we are characterizing thin, or light, existence as poorly demanding on the reality.

Recall, now, that, in Chapter 1, we have explained that Benacerraf-Field's Dilemma is concerned with the problem of finding a plausible explanation of the connections between abstract objects and our ability to refer to them, given that abstracta do not participate in the causal order. In other terms, following Field's elaboration of the problem, we aim to search for a motivation that justifies our way of forming true (mathematical) beliefs, given our metaphysical characterization of mathematical entities as abstract (or quasi-concrete) objects. Finally, as emerged from the end our previous analysis[21], Benacerraf-Field's problem is concerned not only with the possibility of "apprehending" abstract objects (epistemology), but also, and maybe primarily, with the possibility of "referring" to them at all (doxology). In this respect, given our metaphysical and ontological characterization of abstract and quasi-concrete objects (as introduced thanks to laws such as Σ) and concrete bodies, we will try to give plausible explanations to the following two doubts:

> **Doxology**. How do we refer to an abstract or quasi-concrete objects, given that they lack spatio-temporal location, causal powers and are indefinite extensible? What does it grant to us that, words or symbols ,occurring in statements, really refer to abstract or quasi-concrete entities?
>
> **Epistemology**. If there is way thanks to which humans refer to abstract and quasi-concrete objects, then, how should we explain our consequent knowledge of their existence and their force of giving meaning to mathematical or scientific theories – given that we are rejecting any appeal to mysterious and psychologistic epistemic faculties?

#### 5.4.2.1 Suggestion 1

**Mathematical Doxology.** In this little paragraph we wish to discuss how we think that our semantic and metaphysical works can lighten the problem of referring to abstract objects. If we recall how our model works, we can see that we have tried to unpack a mechanism for which, some kind of entities, may be introduced by well-defined identity conditions. Furthermore, we have claimed that what characterizes their thin existence is exactly and simply given by their identity conditions – which are the unique demand that they impose on our reality to exist. We have, hence, a **minimal condition** that an entity has to respect in order to be classified as an existent abstract object. Furthermore, apart characterizing their existential status, grounded identity conditions – we believe – furnish also the minimal condition that grants human reference to abstracta. Consider the case of (unfortunate) Basic Law

[21]See Chapter 1, Section "Against Platonism II: Epistemology", Subsection "Another Challenge: What is Mathematical Doxology?".

V:

$$\epsilon F = \epsilon G \longleftrightarrow \forall x(F(x) \equiv G(x)). \tag{BLV}$$

Now, the abstract objects that BLV introduces are the so-called predicate or concept extensions. As it emerges, any concept extension, $\epsilon F$, has determined identity conditions – as explicitly stated within the LHS of the biconditional –, that are dependent on the truth conditions of the equivalence relation expressed within its RHS. We additionally know that any extension "exists" in a thin manner, i.e., they're not spatio-temporal located and their existential demand requires just the settling of the identity conditions that BLV expresses. We can reason as follows: by considering a concept, such as $F$, we see that its application is defined on the basis of the objects that satisfy $F$, i.e., all the $F(x)$. These latter elements all together form the extension of $F$, i.e., $\epsilon F$. Furthermore, if we consider another concept, let's say $G$, and we inspect again its applicability, i.e., all the entities $x$, such that $G(x)$, then $\epsilon G$ denotes its extension. By inspection, so, if for any $x$, $F(x)$ is materially equivalent to $G(x)$, then the same objects fall under the concept $F$ and under $G$. Provided this, and by previously having defined a concept-extension, $\epsilon F$ (or $\epsilon G$), as the collection of objects that satisfy the specific concept $F$ (or $G$), it might be concluded that the two extensions are identical, $\epsilon F = \epsilon G$. This latter identity fact, finally, – we claim – grants that the extension exist (in our sense) and confirms our ability to pick out abstract objects.

What we actually wanted to render evident with this explanation of BLV, is that, by introducing abstract objects in a thin way, i.e., by simply stipulating their identity conditions, you not merely characterize them ontologically, but, importantly, you refer to them. We know that abstracta are introduced in a grounded manner and that their identity conditions are the minimal requirement for them to exist. In this sense, identity conditions – aside furnishing the minimal requirement for abstracta to exist – furnish also a grant for our referring to them. The example of BLV is illuminating: the collection of entities which fall under $F$ constitutes its extension, which is – as obvious – an abstract object. By simple inspection, indeed, we never encounter something as *the* collection of the $F$s itself, but rather we meet its constituents – i.e., any of the individuals $x$ which falls under $F$ and which its "union" gives me the entire extension of $F$. We recognise, furthermore, that, thanks to the relation of material equivalence between two concepts (stated in the RHS of BLV), denoted by the predicate letters $F$ and $G$, determines also when to concept extensions are identical. So, while observing that $F(x)$ is materially equivalent to $G(x)$, we pick out the same extensions, i.e., the fact that $\epsilon F$ and $\epsilon G$ share the same objects. Thus, our first suggestion is to extend what we are calling the minimal existential condition for abstract objects – which defined their main metaphysical characters – to the study of the possibility of referring to them. In other words, an abstraction principle, by introducing abstract objects through simply stipulating their identity conditions, allows us to consider them as (thin) existent and, moreover, to pick them out.

We conclude, so, that – in according to Quine's celebrated *dictum* – whatsoever entity may be conceived of, it must be ontologically introduced, metaphysically characterized and doxologically explained thanks to its explicit and specific identity conditions.

#### 5.4.2.2 Suggestion 2

**Towards Platonized Naturalism.** In this conclusive paragraph, we wish to give some epistemological suggestions concerning the central question of how are we supposed to apprehend abstract objects – given their thin existence and referentiality. Moreover, we wish to draw some suggestions that are, somehow, supposed to explain our knowledge of abstract entities and their fruitfulness for scientific and mathematical theories. First of all we have to take as background the synthetic/analytic – a priori/a posteriori distinctions. In what follows, our main source of inspiration, are the considerations of the two American philosophers, B. Linsky and E. Zalta. They expressed, with respect to the possibility of giving a plausible epistemology for abstract objects, some interesting considerations, which we think could fit with what we've called **minimal condition** for abstracta to exist and that grants us our referring to them[22].
As remarked several times, Platonism is one of the most prominent schools in the philosophy of mathematics and, for many Platonistic philosophers, abstract entities were considered as part of the "natural realm" – i.e., as something to be "caught" and considered as necessary for our best scientific and mathematical theories. Let's call this view – in accordance with Linsky and Zalta – **Naturalized Platonism**:

> Naturalism is the realist ontology that recognizes only those objects required by the explanations of the natural sciences. But some abstract objects, such as mathematical objects and properties, are required for the proper philosophical account of scientific theories and scientific laws. This has led some naturalists to locate properties or sets (or both) in the causal order, and to suggest that philosophical claims about properties and sets are empirical, discovered a posteriori, and subject to revision.[23]

Now, in order to be clear, we have to take into account another form of epistemology, principally devoted to W. V. Quine[24] and bettered, somehow, by H. Putnam. They, respectively, claimed:

- **Quine**. Some abstract objects (sets and all the mathematical entities thought to be reducible to sets themselves) are required for obtaining the

[22]See Linsky and Zalta 1995, pp. 525–555. Consider that both authors defend their epistemological considerations by taking as basis what they call Axiomatic Metaphysics. This latter approach, characterized by a formal way of doing metaphysics, has been proposed by Zalta 1983. We do not discuss the features of Axiomatic Metaphysics, that both authors link to the epistemology developed in their 1995 paper, but, rather, we will focus exclusively on their epistemological proposal and try to connect it to our previous – metaphysical and doxological – dialogue.
[23]Linsky and Zalta 1995, p. 525.
[24]See again his Quine 1948.

best explanations of the physical world. In other words, for some abstract mathematical objects to exist means simply to be in the range of the quantifier of some scientific theory.

- **Putnam**. Some abstract objects are not only required, but, moreover, **indispensable** for scientific theories to be formulated. Sets are not the only mathematical entities considered as necessary, but, also and importantly for science, properties figure as objectively existing.

Summing up all together we may characterize Naturalized Platonism with the following two statements:

- **Parsimony**. Not all of the abstract mathematical objects are required for our best scientific theories and, indeed, any form of **ontological commitment** is reduced at the minimum indispensable for working scientists.
- **Reduction**. Any entity – apart sets or properties – that serve as a proper explanation of some natural phenomena, have to be reduced to sets (Quine) or to sets and properties (Putnam).

Now, consider that «Quine's formulation of a limited Platonism was seen by many as incomplete, however, for it did not provide an account of our access to abstract objects. How do we obtain *knowledge* of individual abstract objects?»[25] We have already analysed Gödel's idea of a form of cognitive perception guided by the axioms and, with Putnam[26] and Benacerraf-Field[27], agreed that intellectual and rather inexplicable perception-like faculties are not the best way for developing an epistemology for a Platonistic philosophy of mathematics. By having in mind our objective and by starting from Quine-Putnam's epistemological suggestion, we wish to take the distances from their position in different respects: (i) the failure of the parallelism between abstract objects and physical ones[28] and (ii) the position for which *all* abstract object may be rightly introduced in an ontology of mathematics. Let's begin.
First, as clear, for our account of "thin" or "light" existence, it is not plausible that abstract objects participate in the causal order thanks to our "magic" perceptions of them and, indeed, we've argued that – by simple stipulation of their identity conditions within an abstraction principle – we have the warranty that they exist and that we may non vacuously refer to them. From a traditional point of view, for Platonistic philosophers, anyway, abstract objects are – as in our metaphysical account – non spatio-temporal and, hence, outside of the causal order. Moreover, we have different principles that assert the existence of different abstract objects, such HP, in the case of numbers, or BLV, in the case of extensions. Moreover, previously, we have said, that, in second-order logic, a special principle – that asserts

[25]Linsky and Zalta 1995, p. 527.
[26]See Putnam 1980, pp. 466, 471.
[27]Benacerraf 1973.
[28]See the beginning of our Chapter 5, where we discuss the "Parallelism Thesis".

the existence of properties or relations – has to be added. Consider it as follows:

$$\exists F\,\forall x\,(F(x) \longleftrightarrow \varphi(x)). \tag{Comp}$$

It simply assert that for any schematic condition $\varphi$, it exists the corresponding second-order property $F$. In this sense, for instance, we are granted that the property represented by $F$ is the abstract property which exists and is satisfied by all the objects which satisfy $\varphi$. If we stipulate that – in order to instantiate RHSs of abstraction principles – it is somehow needed that the properties to be substituted for the second-order variables exist, then we may think that this "existential grant" comes from our Comprehension Principle. In other words, we may easily believe that principles, specially such as (Comp), deliver abstracta:

> Comprehension principles are very general existence claims stating which conditions specify an object of a certain sort. Some of these principles are distinguished by the fact that they assert that there are as many abstract objects of a certain sort as there could possibly be (without logical inconsistency); [...] Any theory of abstract objects based on such comprehension principles constitutes a *principled Platonism*. Some principled Platonisms are built around comprehension principles for properties, relations and propositions[29].

It is actually true that we have not proposed a formal ontology, in which (Comp) absolves exactly the job of delivering abstract objects, indeed, – for the moment –, we simply believe ,that the «claim that the comprehension principle is required for our understanding of any possible scientific theory is stronger than the claim that it is part of the best scientific theories», and we wish to discuss this view. So, differently from Quine, we do not think that abstract objects serve just as a "proper" explanation of the theories we possess and that they exist in this arbitrary way; indeed, we claim that their existence is to be considered as required for our understanding of those scientific theories themselves and non arbitrary. In this sense, providing a framework with a principles, such as (Comp), may be useful to characterize a *plenum* of abstracta and to introduce them in a well-founded manner. If we desire to apply, for example, HP to some concepts $P$ and $Q$, we must be sure the our second-order quantifier range over a collection of predicates such that it includes $P$ and $Q$. We may be sure that $P$, $Q$ or whatever concept we may desire, exist, exactly by (Comp). In this sense, continuing our example, we have the grant that $P$ and $Q$ exist – thanks to (Comp) – and we may apply HP to see whether the objects – which we abstract from $P$ and $Q$ – exist, i.e., if the identity between the number of $P$ and of $Q$ holds, allowing us then to introduce them into our grounded hierarchical construction of identity/difference facts, $E_\alpha$. In this sense, we agree that Platonistic claims – concerning an ontological introduction of abstracta – may be defended, if based upon principles which allow philosopher to introduce whatsoever abstract mathematical object or concept they might have in mind. The epistemology

[29] Linsky and Zalta 1995, p. 533.

based upon an ontology developed by starting form principles, such as (Comp), has been renamed – by Linsky and Zalta – **Platonized Naturalism**:

> [...] know the comprehension principle if we can rationally conclude that it is part of the best analysis and offers the most uniform understanding of scientific theories.[30]

Furthermore, a crucial role is played – as happens in most philosophical cases – by the settlement of identity conditions between abstracta. From this point of view, our work, has been different from Linsky's and Zalta's and, indeed, we have not developed, as remarked, a formal ontology for abstract objects, but, rather, we have have furnished some semantic considerations regarding a special way of conceiving abstraction processes. In this sense, the work of our comprehension principle for properties or relations does not deliver individuals, meant as objects, but rather properties or relations which serve as basis for our future application of some abstraction principle – this one, meant as delivering abstract individuals. Consider the following example. By (Comp) – as said – properties, such as $P$ and $Q$, exist; instantiate HP RHS with those properties:

$$P \sim Q.$$

If this equivalence, i.e., equipotency, holds between the two concepts involved, then we may easily conclude:

$$\#P = \#Q.$$

That is, the number of the $P$s is equal to the number of $Q$s. We may reason as follows:

1. A Comprehension Principle (with the needed restrictions) grants us the existence of a complex of abstract concepts, properties or relations.
2. Starting from some of these concepts, properties or relations, it is possible to individuate, between them, some equivalence relations, as imposed by the RHS of some well-defined abstraction principle.
3. By starting from the truth-value of the equivalence relation that occurs in the RHS of the abstraction principle under consideration, it possible to abstract from those items, in order to get abstract individuals and the truth-value of their identity or difference relation.
4. Abstract individuals, as abstracted from concepts, properties or relations, are meant to be introduced, in stepwise manner and, exclusively, by considering that, if there is difference, with respect to two or more equivalence relations, then there is a corresponding difference between the abstracted items.

[30] Linsky and Zalta 1995, p. 550.

So, roughly, if we believe that existence of, and the reference to, abstract objects may be introduced this way, then our perspective reconcile in an epistemology, for which the existence of, the reference to, and the knowledge of abstract objects, satisfies the definition of Naturalism we have quoted at the beginning of this section:

> We defined naturalism at the outset, somewhat ambiguously, as the realist ontology that recognizes only those objects required by the explanations of the natural sciences. Platonized naturalism satisfies this definition (in two senses) because it only recognizes the objects falling under the quantifiers of scientific theories and the objects required for a proper philosophical account of those theories. [...] Moreover, Platonized naturalism postulates nothing outside space-time which could be subject to enquiry by one of the natural sciences.[31]

Aside the different metaphysical framework, we think that Linsky's and Zalta's suggestion to develop a principled Platonism, could actually be a good move to restore a sort of naturalistic conception of abstract objects. We've actually tried to clarify the main differences between two different ways of understanding Platonism and Naturalistic epistemology, in order to draw our desired conclusions. As remarked several times, this little paragraph – or, more in general, the entire section devoted to epistemology – has just a general character and, indeed, all the positions here defended, or attacked, should be deepened and developed within a separate and specific chapter. We wanted just to light up two suggestions, which – we think – could be addressed by our much deeper, semantic and metaphysical, work. Nonetheless, we think that by expanding and developing – as said at the end of Chapter 4 – our model-theoretical framework, we may reach some answers with respect to doxological and epistemic question concerning the philosophy of mathematics.

[31]Linsky and Zalta 1995, p. 551.

# Bibliography


Anderson, D.J. and E.N. Zalta (2004). "Frege, Boolos, and Logical Objects." In: *Journal of Philosophical Logic* 33.1, pp. 1–26.

Armour-Garb, B.P. and J.C. Beall, eds. (2005). *Deflationary Truth.* Vol. 1. Open Court Publishing: Chicago.

Benacerraf, P. (1965). "What numbers could not be." In: *Philosophy of Mathematics: Selected readings.* Ed. by H. Benacerraf P. Putnam.

– (1973). "Mathematical Truth." In: *Philosophy of Mathematics: Selected readings.* Ed. by H. Benacerraf P. Putnam.

Benacerraf, P. and H. Putnam, eds. (1983). *Philosophy of Mathematics: Selected readings.* Cambridge University Press: Cambridge.

Berto, F. (2008). *Tutti pazzi per Gödel! La guida completa al Teorema d'Incompletezza.* Laterza: Roma-Bari.

Biggs, N. L. (2002). *Discrete Mathematics.* 2nd ed. Oxford University Press: New York.

Boolos, G. (1990). "The Standard Equality of Numbers." In: *Meaning and Method. Essays in Honor of Hilary Putnam.* Ed. by G. Boolos. Cambridge University Press: Cambridge, pp. 261–277.

Button, T. and S. Walsh (2011). *Philosophy and Model Theory.* Oxford University Press: Oxford.

Cantor, G. (1883). "Foundations of a General Theory of Manifolds: A Mathematical Philosophical Study in the Theory of the Infinite." In: *From Kant to Hilbert: A Source Book in the Foundations of Mathematics.* Ed. by W. Ewald. Vol. II. Clarendon Press: Oxford, pp. 879–920. URL: http://cogito.lagado.org/sites/default/files/Cantor%20-%201883%20-%20Grundlagen.pdf.

– (1899). "Letter to Dedekind." In: *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931.* Ed. by J. van Heijenoort. Harvard University Press: Cambridge (Mass.), pp. 113–117.

– (1915). *Contributions to the Founding of the Theory of Transfinite Numbers.* Open Court Publishing: Chicago.

Carlucci Aiello, L. and F. Pirri (2005). *Strutture, logica, linguaggi.* Pearson: Milano.

Costantini, F. (2016). *Pensare l'infinito. Filosofia e Matematica dell'Infinito in Bernard Bolzano e Georg Cantor.* Mimesis: Milano-Udine.

Dedekind, R. (1890a). "Letter to Keferstein." In: *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931*. Ed. by J. van Heijenoort. Harvard University Press: Cambridge (Mass.), pp. 98–103.

– (1888a). "Letter to Weber." In: *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*. Ed. by W. Ewald. Vol. II. Clarendon Press: Oxford, pp. 834–835.

– (1888b). "Was sind und was sollen die Zahlen?" In: *Gesammelte mathematische Werke*. Ed. by R. Fricke et al. Vol. 3. (1932). F. Vieweg Verlag: Braunschweig, pp. 787–833.

Dummett, M. (1995). *Frege: Philosophy of Mathematics*. Harvard University Press: Cambridge (Mass.)

Ebert, P. A. and M. Rossberg (2007). "What is the purpose of Neo-Logicism?" In: *Travaux de logique* 18, pp. 33–61.

– (2017). "Introduction to Abstractionism." In: *Abstractionism. Essays in Philosophy of Mathematics*. Ed. by P. A. Ebert and M. Rossberg. Oxford University Press: Oxford, pp. 3–33.

Field, H. (1989). *Realism, Mathematics and Modality*. Blackwell: Oxford.

Fraenkel, A. A. (1922). "The notion of 'definite' and the indipendence of the axiom of choice." In: *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931*. Ed. by J. van Heijenoort. Harvard University Press: Cambridge (Mass.), pp. 284–289.

Frege, G. (1893/1903). *Grundgesetze der Arithmetik [Basic Laws of Arithmetic]*. Ed. by P. A. Ebert and M. Rossberg. Vol. I and II. (2013). Oxford University Press: Oxford.

– (1879). "*Begriffsschrift*, a formula language, modeled upon that of arithmetic, for pure thought." In: *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931*. Ed. by J. van Heijenoort. Harvard University Press: Cambridge (Mass.), pp. 5–82.

– (1884). *The Foundation of Arithmetic. A logico-mathematical enquiry into the concept of number*. Ed. by J. Austin. (1953). Blackwell: Oxford.

– (1895). "Kritische Beleuchtung einiger Punkte." In: *Vorlesungen über die Algebra der Logik. Archiv für systematische Philosophie I*. Ed. by E. Schröders. Teubner: Leipzig, pp. 433–456.

– (1902). "Letter to Russell." In: *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931*. Ed. by J. van Heijenoort. Harvard University Press: Cambridge (Mass.), pp. 126–128.

Gödel, K. (1931). "Über Formal Unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I." In: *Monatshefte für Mathematik und Physik* 38.1, pp. 173–198.

– (1944). "Russell's Mathematical Logic." In: *Philosophy of Mathematics: Selected readings*. Ed. by P. Benacerraf and H. Putnam.

– (1947). “What is Cantor’s continuum problem?” In: *Philosophy of Mathematics: Selected readings*. Ed. by P. Benacerraf and H. Putnam.

– (1951). “Some basic theorems on the foundations of mathematics and their implication.” In: *Collected Works. Unpublished Essays and Lectures*. Ed. by Feferman S. et al.

– (1961). “The modern development of the foundations of mathematics in the light of philosophy.” In: *Collected Works. Unpublished Essays and Lectures*. Ed. by Feferman S. et al.

– (1995). *Collected Works. Unpublished Essays and Lectures*. Ed. by S. Feferman et al. Vol. 3. Oxford University Press: Oxford.

Harzheim, E. (2006). *Ordered Sets*. Springer: New York.

Heck, R. G. (1995). “Definition by induction in Frege’s *Grundgesetze der Arithmetik*.” In: *Frege’s Philosophy of Mathematics*. Ed. by W. Demopoulos. Harvard University Press: Cambridge (Mass.), pp. 295–333.

– (2011). “Ramified Frege Arithmetic.” In: *Journal of Philosophical Logic* 40.6, pp. 715–735.

– (2017). “The Existence (and Non-Existence) of Abstract Objects.” In: *Abstractionism. Essays in Philosophy of Mathematics*. Ed. by P. A. Ebert and M. Rossberg. Oxford University Press: Oxford, pp. 50–78.

Hersh, R. (1997). *What is mathematics, really?* Oxford University Press: Oxford.

Horsten, L. (2007). “Philosophy of Mathematics.” In: *Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta.

Horsten, L. and H. Leitgeb (2009). “How abstraction works.” In: *Reduction - Abstraction - Analysis*. Ed. by H. Leitgeb and A. Heike. Proceedings of the 31th International Ludwig Wittgenstein-Symposium in Kirchberg. de Gruyter: Berlin, pp. 217–226.

Horsten, L. and Ø. Linnebo (2016). “Term Models for Abstraction Principles.” In: *Journal of Philosophical Logic* 45.1, pp. 1–23.

Kanamori, A. (2003). “The Empty Set, the Singleton, and the Ordered Pair.” In: *Bulletin of Symbolic Logic* 9.3, pp. 273–298.

– (2009). “Set theory from Cantor to Cohen.” In: *Philosophy of Mathematics*. Elsevier/North Holland: Amsterdam, pp. 395–459. URL: `http://math.bu.edu/people/aki/16.pdf`.

Kennedy, J. (2007). “Kurt Gödel.” In: *Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta.

Kripke, S.A. (1975). “Outline of a theory of truth.” In: *Journal of Philosophy* 72, pp. 690–716.

Leitgeb, H. (2005). “What truth depends on.” In: *Journal of Philosophical Logic* 34, pp. 155–192.

– (2007). “What Theories of Truth Should be Like (but Cannot be).” In: *Philosophy Compass* 2.2, pp. 276–290.

Leitgeb, H. (2017). "Abstraction Grounded: A Note on Abstraction and Truth." In: *Abstractionism. Essays in Philosophy of Mathematics*. Ed. by P. A. Ebert and M. Rossberg. Oxford University Press: Oxford, pp. 269–282.

Lewis, D. (1983). "New Work for a Theory of Universals." In: *Australasian Journal of Philosophy* 61.4, pp. 343–377.

Linnebo, Ø. (2004). "Predicative Fragments of Frege Arithmetic." In: *Bulletin of Symbolic Logic* 10.2, pp. 153–174.

– (2006). "Epistemological Challenges to Mathematical Platonism." In: *Philosophical Studies* 129.3, pp. 545–574.

– (2009a). "Bad company tamed." In: *Synthese* 170.3, pp. 371–391.

– (2009b). "Introduction [to special issue regarding the 'Bad Company Problem']." In: *Synthese* 170.3, pp. 321–329.

– (2009c). "Platonism in the Philosophy of Mathematics." In: *The Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta.

– (2011). *Philosophy of Mathematics*. Princeton University Press: Princeton.

– (2017). "Impredicativity in the Neo-Fregean Program." In: *Abstractionism. Essays in Philosophy of Mathematics*. Ed. by P. A. Ebert and M. Rossberg. Oxford University Press: Oxford, pp. 247–268.

– (2018). *Thin Objects: An Abstractionist Account*. Oxford University Press: Oxford.

Linnebo, Ø. and R. Pettigrew (2014). "Two types of Abstraction for Structuralism." In: *The Philosophical Quarterly* 64.255, pp. 267–283.

Linsky, B. and E.N. Zalta (1995). "Naturalized Platonism versus Platonized Naturalism." In: *The Journal of Philosophy* 92.10, pp. 525–555.

McGee, V. (1993). "A semantic conception of truth?" In: *Deflationary Truth*. Ed. by B.P. Armour-Garb and J.C. Beall, pp. 111–142.

Moschovakis, Y. (2006). *Notes on Set Theory*. Springer: New York.

Oliver, A. and T. Smiley (2006). "What are sets and what are they for?" In: *Philosophical perspectives* 20.1, pp. 123–155.

Parsons, C. (1995). "Platonism and Mathematical Intuition in Kurt Gödel's Thought." In: *The Bulletin of Symbolic Logic* 1.1, pp. 44–74.

Plebani, M. (2011). *Introduzione alla filosofia della matematica*. Carocci editore: Roma.

Potter, M. (2000). *Reason's Nearest Kin. Philosophies of Arithmetic from Kant to Carnap*. Oxford University Press: Oxford.

Putnam, H. (1980). "Models and Reality." In: *The Journal of Symbolic Logic* 45.3, pp. 464–482.

Quine, W.V. (1948). "On what there is." In: *The Review of Metaphysics* 2.5, pp. 21–38.

Reck, E. (2013a). "Frege, Dedekind, and the Origins of Logicism." In: *History and Philosophy of Logic* 34, pp. 242–265.

– (2013b). “Frege or Dedekind? Towards a Reevaluation of their Legacies.” In: *The Historical Turn in analytic Philosophy*. Ed. by E. Reck. Palgrave Macmillan: London, pp. 546–571.
– (2003). “Dedekind’s Structuralism: An Interpretation and Partial Defense.” In: *Synthese* 137, pp. 369–419.
– (2009). “Dedekind, Structural Reasoning, and Mathematical Understanding.” In: *New Perspectives on Mathematical Practices*. Ed. by B. van Kerkhove. WSPC Press: Singapore, pp. 150–173.
– (2016). “Dedekind’s Contributions to the Foundations of Mathematics.” In: *Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta.
– (2017). “Dedekind as a Philosopher of Mathematics.” In: *In Memoriam Richard Dedekind (1831-1916)*. VTM-Verlag: Munster, pp. 36–49.
– (2018). “On Reconstructing Dedekind Abstraction Logically.” In: *Logic, Philosophy of Mathematics, and their History: Essays in Honor of W.W. Tait*. Ed. by E. Reck. College Publications: London, pp. 113–138.
Russell, B. (1902). “Letter to Frege.” In: *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931*. Ed. by J. van Heijenoort. Harvard University Press: Cambridge (Mass.), pp. 124–125.
– (1903). *The Principles of Mathematics*. (1996). W.W. Norton & Company: New-York.
– (1907). “The Regressive Method of Discovering the Premises of Mathematics.” In: *Essays in Analysis*. Ed. by D. Lackey. (1973). George Braziller: New York, pp. 272–283.
– (1919). *Introduction to Mathematical Philosophy*. (2008). Spokesman Books: Nottingham (UK).
Shabel, L. (2013). “Kant’s Philosophy of Mathematics.” In: *Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta.
Skolem, T. (1922). “Some remarks on axiomatized set theory.” In: *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931*. Ed. by J. van Heijenoort. Harvard University Press: Cambridge (Mass.), pp. 290–301.
Smith, P. (2013). *An Introduction to Gödel’s Theorems*. Cambridge University Press: Cambridge.
Tait, W.W. (1996). “Frege versus Cantor and Dedekind: On the Concept of Number.” In: *Frege: Importance and Legacy*. Ed. by M. Schirn. de Gruyter: Berlin, pp. 70–113.
van Heijenoort, J. (1967). *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931*. Harvard University Press: Cambridge (Mass.)
von Neumann, J. (1925). “An axiomatization of set theory.” In: *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931*. Ed. by J. van Heijenoort. Harvard University Press: Cambridge (Mass.), pp. 393–413.
– (1929). “Über eine Widerspruchfreiheitsfrage in der axiomatischen Mengenlehre.” In: *Journal für die reine und angewandte Mathematik* (160), pp. 227–241.

Wang, H. (1996). *A Logical Journey: From Gödel to Philosophy.* MIT Press: Cambridge (Mass.)

Weir, A. (2003). "Neo-Fregeanism: An Embarrasment of Riches." In: *Notre Dame Journal of Formal Logic* 44.1, pp. 13–48.

Yablo, S. (1982). "Grounding, Dependence and Paradox." In: *Journal of Philosophical Logic* 11, pp. 117–137.

Yap, A. (2009). "Logical Structuralism and Benacerraf's Problem." In: *Synthese* 171.1, pp. 157–173.

Zalta, E.N. (1983). *Abstract objects: An introduction to Axiomatic Metaphysics.* D. Reidel Publishing Company: Dordrecht.

– (1999). "Natural Numbers and Natural Cardinals as Abstract Objects: A Partial Reconstruction of Frege's *Grundgesetze* in Object Theory." In: *Journal of Philosophical Logic* 28.6, pp. 617–658.

– (2001). "Fregean Senses, Modes of Presentation, and Concepts." In: *Philosophical Perspectives* 15, pp. 335–359.

– (2018). "Frege's Theorem and Foundations for Arithmetic." In: *Stanford Encyclopedia of Philosophy.* Ed. by E.N. Zalta.

Zermelo, E. (1904). "Proof that every set can be well-ordered." In: *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931.* Ed. by J. van Heijenoort. Harvard University Press: Cambridge (Mass.), pp. 139–141.

– (1908a). "A new proof of the possibility of a well-ordering." In: *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931.* Ed. by J. van Heijenoort. Harvard University Press: Cambridge (Mass.), pp. 183–198.

– (1908b). "Investigations in the foundations of set theory I." In: *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931.* Ed. by J. van Heijenoort. Harvard University Press: Cambridge (Mass.), pp. 199–215.

– (1930). "On boundary numbers and domains of sets: New investigations in the foundations of set theory." In: *From Kant to Hilbert: A Source Book in the Foundations of Mathematics.* Ed. by W. Ewald. Vol. II. Clarendon Press: Oxford, pp. 1208–1233.