



Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Kant, Cantor, and the Unconditioned

Thesis submitted in order to obtain the degree of
PhD in Philosophy.

Damián Bravo Zamora.



Thesis 9681

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Damián Bravo Z

106 / 2012

Summary

In this thesis I inquire into the possible connections between the philosophical problem that Immanuel Kant called the First Antinomy of Pure Reason and some of the paradoxes that were discovered in set theory in the second half of the nineteenth century and at the turn of the twentieth century. I provide arguments in support for the thesis that there are motivational, structural and conceptual similarities between these two problems. The main claim of the investigation is that Kant's discussion of the Antinomy can be interpreted as an early episode in the attainment of a philosophical idea which after the discovery of the set-theoretic paradoxes seems to suggest itself, namely that the concept of an absolutely all-embracing totality (Everything) is problematic in the sense that either paradoxes or puzzlements arise out of it.

In the first chapter I examine Kant's claim that there is a first antinomy of pure reason, namely that we must claim that the world, conceived as the mathematical whole of spatiotemporal objects and events, is both finite and infinite, at least under the presupposition of a doctrine that Kant calls "transcendental realism". I argue that Kant fails to derive an antinomy from the concept of the world under the presupposition of transcendental realism, and hence I disagree with some contemporary interpreters of Kant, who in an attempt to defend Kant's "indirect argument for idealism" (a *reductio ad absurdum* of realism), have tried to restate the Kantian first antinomy.

In the second chapter I present the paradoxes of set theory in a way that does not presuppose any knowledge of set theory by the reader. I then move on to a discussion of a philosophically significant way to understand the outcome of some of the paradoxes: if every collection must be thought of a set, given the paradoxes generated by the concept of a set of all sets, it seems that we must reject our existential commitment to an all-embracing collection (Everything). On the other hand, we can keep the concept of an all-embracing totality but we must give up the idea that every collection must be a definite one, i.e. a set. I present these two options as the two available ones as long as we adhere to the standard system of set theory, the axiomatic system known as Zermelo-Fraenkel.

In the third chapter I examine two non-standard ways of avoiding the paradoxes generated by the concept of an all-embracing collection. These are the system of set theory called "New Foundations", which stems from a 1937 article by Willard Van Orman Quine, and a system of non-set collections called "Plenum Theory", proposed by Nicholas Rescher and Patrick Grim in their recent work (2007 and 2011). I argue that these two systems do manage to avoid the classical paradoxes regarding an all-embracing class, but that this does not mean that no obscurities or puzzlements remain with regard to the collection of Everything. In the case of Quine's system, a

problem with the self-membership seems to be arise in the presence of the axiom of extensionality. In the case of Plenum Theory, to the obscurities of self-membership one must add some difficulties which, I claim, arise out of Rescher's and Grim's claim that plena are indeterminate collections.

In the fourth chapter, having looked at the paradoxes and puzzlements that are generated by the idea of Everything, I try to connect these issues with the Kantian discussion of the so-called First Antinomy of Pure Reason. Although in the first chapter I had argued that there is no such thing as a genuine Kantian antinomy, in this fourth chapter I focus on passages of Kant's text where he describes the Antinomy as being generated by the idea of the unconditioned. I then characterize the concept of an absolutely all-embracing totality as a radicalization of Kant's concept of the unconditioned. I then defend the main claim of this thesis, namely that one can interpret the paradoxes and puzzlements generated by the concept of Everything as more convincing and more radical versions of the Kantian antinomy. The main claim is that the ambitious idea of the unconditioned has little to offer to our theorizing, apart from either contradictions or obscurities or both things.

Acknowledgements

This text is a PhD thesis that I have elaborated under the supervision of Lilian Alweiss. I wish to express my gratitude to her for giving me as much freedom and time as I needed in order to find the exact problem I deal with. I also wish to thank her for her friendship and her support during all these years, and for her willingness to discuss my work in detail, especially those parts of it where her approach to these topics differs from mine, or where she thought that my argumentation in previous versions of this text was weak.

I am also greatly indebted to James Levine. It was through him that I learned that the Kantian antinomy could be looked at rigorously –hence critically- without necessarily taking it to be a picturesque episode in the history of philosophy. Although I am not sure that he would agree with much of what I say or suggest in this thesis, I tend to think that if there is anything correct in it, it is something I learned with his help.

During four of the five years in which the investigation that led to this text was carried out (from 2006 to 2009 and from 2010 to 2011), I received a grant from Trinity College Dublin which paid my student fees and allowed me to live with commodity. Without that financial support, the present investigation would have been impossible. I am extremely grateful for that. During the remaining year (from 2009 to 2010), I was awarded a scholarship by the German Academic Exchange Service (DAAD: Deutscher Akademischer Austausch Dienst), which in the context of an academic exchange enabled me to carry out my research in the universities of Konstanz (in the Wintersemester) and Humboldt in Berlin (in the Sommersemester). I wish to thank that funding German institution, as well as Professor Tobias Rosefeldt for having made this academic exchange possible.

This corrected version of my thesis is submitted after the *viva voce* examination which took place on 15 March 2012. I wish to express my gratitude to Professors Peter Simons and Øystein Linnebo, who were the internal and examiners respectively, for their very helpful comments, criticisms and suggestions.

to Rosario, Emilia and Sofia

to Nieves

Table of Contents

Introduction.....	15
Chapter I: Kant's Alleged Antinomy.....	21
1.1 The Systematic Role of the First Antinomy in Kant's Theoretical Philosophy.....	23
1.2 The Alleged Antinomy.....	27
1.2.1 The Thesis.....	27
1.2.2 The Antithesis.....	32
1.3 Does Transcendental Idealism Strictly Follow from the Alleged First Antinomy?.....	37
1.4 Why I am not a transcendental idealist.....	43
Chapter II: A Collection of All Collections: A First Approach to its Problems.....	51
2.1 The Paradox of the Largest Cardinal.....	53
2.2 Russell's Paradox.....	61
2.3 How to React in Case of Paradox.....	66
2.4 Potential Universe, Indefinite Being.....	71
Chapter III: Two Ways of Saving Everything from Paradox.....	77
3.1 Avoiding the Diagonalised Collection.....	78
3.1.1 Quine's New Foundations: An alternative set theory.....	78
3.1.2 Plenum Theory: An alternative to set theory.....	86
3.2 "... and perhaps the mind ought to boggle..."	92

Chapter IV: Kant, Cantor, and the Unconditioned.....	103
4.1 A totality of What?.....	104
4.2 Totalisation and Beyond.....	113
4.3 A Farewell to the Unconditioned?.....	119
 Conclusions.....	 127
 Bibliography.....	 133

... Juzgan que la metafísica es una rama de la literatura fantástica.

Jorge Luis Borges, on the philosophers of Tlön

Introduction

Reason falls into this perplexity through no fault of its own. It begins from principles whose use is unavoidable in the course of experience and at the same time sufficiently warranted by it. With these principles it rises (as its nature also requires) even higher, to more remote conditions. ... But it thereby falls into obscurity and contradictions

Kant

Philosophy seems to be unable to dispense with totalisation. It seems to rest on the presupposition that even if we cannot know all the elements and aspects of a particular multiplicity of objects, we can have a conceptual access to at least some (essential) features of that multiplicity taken as a whole or totality, and hence as an individual object in its own right, not to be identified with any of its constitutive elements or with a mere aggregate of them. Thus understood, totalisation seems to have been fundamental in the formation of the early Pre-Socratics' concept of *physis*, and it seems to be at the core of our contemporary philosophical and scientific concept of *the spatiotemporal world*. In its most ambitious variety, totalisation has led some philosophers to form the concept of the most comprehensive totality: the multiplicity –taken as an individual object- of absolutely all things. The first philosopher to enunciate this radical form of totalisation was perhaps also a Pre-Socratic, namely Heraclitus, who thought that it was a matter of listening 'not to him but to the logos' in order wisely to agree that 'all things are one'.

The suspicion, however, that the most ambitious sort of totalisation might be unaccomplishable since it is plagued by puzzlement, if not by paradoxes, is at least as old as Plato. In this investigation we will be concerned with two episodes in the history of this suspicion. One of them is the 'antinomies' that Kant says arise with regard to the concept of the spatiotemporal world –at least as long as one does not adopt his doctrine of transcendental idealism-. In particular, we will be concerned with Kant's claim that under the quite natural assumption of transcendental realism (and

Kant thought that all his predecessors shared this assumption) reason is bound to affirm that the spatiotemporal world is *both* finite (Thesis) and infinite (Antithesis). This is what Kant calls the First Antinomy of Pure Reason. The second subject matter of this investigation are the paradoxes which were discovered, at the dawn of Cantorian set theory, with regards to oversize totalities. In particular, we will be concerned with the paradoxes that might arise if we assume that there is an all-embracing totality which is an individual object of the same (definite and iterable) kind as the objects it embraces. That is to say, we will look at the problems surrounding the concept of a set of all sets and, more generally, of a collection of all collections.

There is one important reason why one might think that the only connection that can be drawn between these two episodes of thought is a superficial and in any case a negative one. As we will see, an essential element in Kant's "proof" of the thesis of his first antinomy is the assumption that an infinite series can never be completed, which is just the traditional Aristotelian view that a series can only be said to be potentially but not actually infinite. Now, apart from the fact that this assumption, in the context in which it appears, will betray Kant's "proof" as a fallacy of circularity, it was precisely the assumption that a series can also be actually and not only potentially infinite what paved the way for the development of Cantorian set theory. True, the theory soon encountered paradoxes. But, as we will also see, the paradoxes can be avoided by giving up the commitment to the existence of some problematic oversize totalities while at the same time keeping the commitment to the existence of all other actually infinite totalities which are instrumental in the accomplishment of set theory's scientific task: to provide the conceptual foundations of classical mathematics. This is the way in which the paradoxes are avoided in the standard system of set theory, i.e. the axiomatic system known as Zermelo-Fraenkel (or ZF for short). In sum, not only was Kant not allowed to presuppose that an actually infinite series is impossible; we, after the development of set theory, know that the assumption that infinite series can be actual is a fruitful one in our attempt to understand mathematical practice. If set theory provides a correct account of infinity, as it seems to do, pointing at its development is perhaps the best way to show that there is no such thing as a Kantian "first antinomy" at all.

During the course of this investigation, I will show that this way of looking at the Kantian first antinomy from the vintage point of set theory is, although no doubt correct, extremely partial. Not only are there philosophically significant *structural* similarities in the way both Kantian and Cantorian antinomies are *generated*; a deep connection between the two episodes of thought becomes evident when one takes Kant by his word when he says, in the Preface to the second edition of the *Critique of Pure Reason*, that the most essential task of metaphysics is to *think of the unconditioned*,

and that the Antinomies chapter shows that his philosophical predecessors (and, we can now say, many of his successors, since some of them are transcendental realists too) cannot accomplish this task *without contradiction*. If by “the unconditioned” Kant meant, as he did, the spatiotemporal world, he most certainly failed to show that transcendental realists cannot think of it without contradiction, as many critics have pointed out. This is a clear conclusion of our investigation –and it is also the least interesting of our conclusions-. It is much more interesting to ask what Kant could have meant by “the unconditioned” or at least what we can interpret him as meaning by that phrase. As I will propose, we can characterize in the same way the oversize totalities which the paradoxes of set theory show to be problematic, at least under certain assumptions. In particular, given that the relation that a totality has to its members is called “comprehension”, can we not say that an absolutely all-embracing totality is a totality which, as regards comprehension, is *absolutely unconditioned*? And do the paradoxes of set theory not show that this object cannot be thought of without contradiction, at least as long as we keep some quite natural presuppositions (like the axiom of comprehension)? In this investigation I will answer to these two questions positively, and I hope that this will show that the Kantian first antinomy is not (only) a picturesque historical curiosity. It can also be interpreted as an antecedent of the paradoxes of set theory and, unless our totalising ambitions are fully satisfied by one or other of the contemporary ways to avoid the paradoxes, it might be worth the while to inquire into the possible connections between our past and our present puzzlements.

Amongst the authors who have been working on the connections between Kantian and Cantorian antinomies, I should like to mention three of them from whose work this investigation takes its cue: Gottfried Martin, A.W. Moore, and Graham Priest. I will not always agree with their views or share their intentions, but the main suggestion, namely that there are motivational, structural and conceptual similarities between Kantian and Cantorian antinomies, is explicitly stated in their writings.¹

In chapter one I look at Kant’s first antinomy and to the role it plays in Kant’s theoretical philosophy. No reference is made to set theory in that chapter, since my intention is to show that the antinomy is not genuine anyway. By way of preventing future temptations, I also argue that the “solution” Kant proposes to the alleged antinomy, namely transcendental idealism, was neither the only

¹ Personally, I owe my awareness of this fascinating topic to a series of lectures given by Jim Levine in Trinity College Dublin.

possible nor even a plausible way out of the antinomy –all this under the quite generous assumption that the antinomy is genuine.

In chapter two I look for the first time at the paradoxes of set theory. I first present the paradox of the largest cardinal (also called “Cantor’s paradox”) and Russell’s paradox in a way which, I hope, is fully accessible to a reader with no previous knowledge of set theory. I then move on to a discussion of the philosophical significance of these paradoxes. I defend the claim that one of the most appealing suggestions that can be drawn from the paradoxes is that, if we are going to commit to the existence of an all-embracing collection at all, it cannot be thought of as an entity of the same kind as the objects that make it up. Thus, we must either affirm with (the late) Russell that there is no such thing as Everything, or we must make room, as Cantor himself did, for oversize collections which are not of the same kind as the objects that make them up. If every collection is a Cantorian set, then there is no set of all sets, and hence no absolutely all-embracing collection. If on the other hand we want to keep our commitment to the existence of such a thing as Everything, it seems that we must give up the thought that Everything is a definite and iterable thing –a set-. These are, anyway, the two options that we have on the table as long as we work with a series of assumptions equivalent to the ones that are made in the standard system of set theory, ZF.

In chapter four I look at two non-standard ways of theorizing about collections. One of them (namely, Quine’s system *New Foundations*) is an alternative set theory, while the other one (namely, Nicholas Rescher’s and Patrick Grim’s “*Plenum Theory*”) is an alternative *to* set theory. For all their important differences, what these two systems have in common is a commitment to the existence of an all-embracing collection. There is, according to these systems, such a thing as Everything. I will claim that the fact that the traditional paradoxes are avoided by these systems does not mean that no obscurities or dissatisfactions are left by them with respect to the concept of an all-embracing totality. In the case of Quine’s NF, the self-membership of some sets (like the set of all sets), together with the fact that those sets must conform to the axiom of extensionality, imply that not even an idealized Omniscient Mathematician could non-circularly identify them. In the case of *Plenum Theory*, to the mind-boggling idea of self-membership one must add the complications introduced by the characterization of plena as indeterminate and ever-growing. Granted, these problems are not paradoxes. They are problems nonetheless. At least for some of us.

In the fourth and final chapter we go back to Königsberg. Having seen in the two previous chapters that the idea of an absolutely all-embracing collection (Everything) generates either paradoxes or puzzlements, I characterize that idea as a radicalization of Kant’s concept of *the world*.

The characterization is allowed, I propose, by Kant's own statement that the antinomies of pure reason are generated by the idea of the unconditioned, and by the quite evident fact that the concept of an absolutely all-embracing collection (a collection whose comprehension need not be limited to spatiotemporal objects and events) is the concept of something unconditioned if anything else is. The main claim of this chapter, as well as that of the whole investigation, is perhaps not new. It is perhaps an old suspicion that suggests itself once one understands the paradoxes of set theory. But the suspicion is couched, in this fourth chapter, in Kantian language. The idea of the unconditioned brings paradoxes or at least puzzlements with it. It looks as if as long as we are the kind of cognitive beings we contingently are, our most ambitious totalising urges must be qualified in some way or other, at some point or other.

Chapter I. Kant's Alleged First Antinomy

The discussion of what Kant calls the Antinomies of Pure Reason is interesting for at least two reasons. One of them is that these problems are supposed to be genuine, independent problems in their own right, at least under the presupposition of a doctrine that Kant sees as being presupposed in one sense or other by all his philosophical predecessors (namely transcendental realism). The second reason, which is related to the first, is that these problems are supposed to be avoided by adopting the doctrine of transcendental idealism, which is arguably the core of Kant's theoretical philosophy. A discussion of the Kantian antinomies is thus at the same time a discussion of traditional metaphysical problems and of a key element in the justification of Kant's doctrine of transcendental idealism.

An antinomy is supposed to be a contradiction or, better said, a paradoxical situation in which reason gets involved whenever it allows itself to be led by an ambitious theorizing principle or law (*nomos*). This principle is Kant's elaborate version of Leibniz's *principle of sufficient reason*: that there must be a sufficient reason for everything to be the way it is and not otherwise. Kant's version can be called the *principle of the unconditioned*: whenever something conditioned is given, the entire series of its conditions, and hence the absolutely unconditioned, must be given as well.¹ For the moment, what is important is that Kant thinks that this principle can be understood in two different ways, giving birth to the two antagonist claims that constitute an antinomy: a thesis and an antithesis. The principle can be understood as the claim that whenever there is an ascending series of conditioned objects, one of these objects, a privileged member of the series, is the unconditioned demanded by the principle. This way of understanding the principle accounts for the thesis of a Kantian antinomy. On the other hand, the principle can be understood as the claim that absolutely every member of an ascending series of conditions must have a determining condition, and hence only the series itself can be regarded as unconditioned. This second way of understanding the principle gives birth to the antithesis of what Kant calls an antinomy.

¹ I will elaborate on the connection between Kant's and Leibniz's principle below, namely in section 1.2.2 of this chapter.

So much for the etymology of the word 'antinomy'. As the above shows, there is nothing particularly esoteric about Kant's claim that his philosophical opponents must face antinomies. And this, as we said, generates two interesting questions: 1) whether those problems are really unavoidable for his opponents –i.e. whether transcendental realists must regard them as genuine antinomies-, and 2) whether Kant succeeds in showing that only by adopting transcendental idealism can the antinomies be avoided.

In this chapter I will address these questions with regard to one of the four Kantian antinomies. According to Kant, the first antinomy of pure reason is that, under the assumption of transcendental realism, we must claim that the spatiotemporal world is both finite and infinite.¹ In section 1.1 I will explain the role that this antinomy plays in Kant's theoretical philosophy; in particular, I will explain Kant's claim that the first antinomy² provides an indirect proof of transcendental idealism. In section 1.2 I will examine the antinomy itself, i.e. I will address question 1) above, whether transcendental realists must hold both that the world is finite (subsection 1.2.1) and that it is infinite (subsection 1.2.2). In section 1.3 I will address the second question, i.e. whether transcendental idealism is the *only possible* way out of the antinomy, granting that it is a genuine antinomy. Finally, in section 1.4 I will consider the question whether, even if transcendental idealism is not the only way out of the alleged antinomy, the doctrine can nonetheless be regarded at least as a *plausible* way out of the antinomy.

My answer to all these questions is an unequivocal 'No'. There is no genuine Kantian first antinomy (the transcendental realist is neither forced to say that the world is finite nor to say that it is infinite). It is also not the case that transcendental idealism is the only way out of the antinomy, even under the generous presupposition that the antinomy is genuine (if the self-contradictory nature of the world had been proved, all that would have followed is that we must be as committed to the world's existence as we are committed to the existence of a round square). Finally, transcendental idealism cannot be considered as a plausible way out of the alleged antinomy (for the doctrine, even under sympathetic interpretations, creates as many problems as it solves).

¹ Although Kant's antinomy consists of proofs against the finitude and infinitude of the world with respect to both space and time, I will focus on the *temporal* side of the Antinomy. This is not only for the sake of brevity, but because Kant's argument for the spatial finitude of the world takes for granted the essential premise of the argument against the temporal infinitude of the world. So there is some interpretative clarity to be gained by focusing on the temporal side of the Antinomy. Besides, I believe that none of the claims I advance in this chapter would be rendered incorrect, but rather confirmed, by a detailed analysis of the spatial side of the first antinomy.

² Together with the second one, with which I will not deal.

A legitimate question is likely to have arisen already in the reader's mind: What are the motivations for dealing with Kant's first antinomy in any detail, if it is considered to be a false problem to which Kant gave an implausible solution? My answer is to ask the reader to have some patience. I hope that in chapters II, III and IV it will become increasingly clearer that the Antinomy section of the first *Critique* can be read as something much more interesting than a picturesque episode in the history of philosophy. But this cannot be understood unless we rigorously begrudge Kant every pinch of salt. If there are any lessons to be learned from Kant's first antinomy, they can only be learned if we are as critical readers as the *Critique* deserves them.

1.1 The Systematic Role of the First Antinomy in Kant's Theoretical Philosophy

Seen from within Kant's theoretical philosophy, the first two antinomies are assigned the hardest of all tasks. They are supposed to provide a *reductio ad absurdum* of transcendental realism and thereby justify the adoption of transcendental idealism.

Transcendental idealism is the doctrine according to which space, time, and all spatiotemporal objects and events are mere representations and not things in themselves. To this doctrine Kant opposes transcendental realism, i.e. the doctrine according to which space, time and all spatiotemporal objects and events are things in themselves. Here is an explicit definition of both doctrines:

We have sufficiently proved in the Transcendental Aesthetic that everything intuited in space or time, hence all objects of an experience possible for us, are nothing but appearances, i.e., mere representations, which, as they are represented, as extended beings or series of alterations, have outside our thoughts no existence grounded in itself. This doctrine I call *transcendental idealism*. The realist, in the transcendental signification, makes these modifications of our sensibility into things subsisting in themselves, and hence makes *mere representations* into things in themselves. (A491-2/B 518-9)¹

¹ References to the *Critique of Pure Reason* will be made from now on by indicating only the canonical pagination. "A" refers to those passages which appear in the first edition only, "B" to those appearing only in the second edition, and A/B for those passages which appear in both.

This passage is an unequivocal statement of the first thesis by which we might identify the doctrine of transcendental idealism. It is the *Non-spatiotemporality Thesis*. Things in themselves are not spatiotemporal. The second defining thesis of the doctrine is inseparable from the first one if one accepts two further claims that Kant defends in the Transcendental Aesthetic, namely that knowledge can only be attained with respect to objects of possible experience, and that the latter are spatiotemporal objects and events. This implies, if we accept the Non-spatiotemporality Thesis, that things in themselves are unknowable. We can refer to this second thesis as the *Unknowability Thesis*.¹ Since, as we will see, the strength of the Antinomy chapter depends on our *not* accepting Kant's claims about the objects of possible experience, we will identify from now on transcendental idealism with the first thesis: that things in themselves are not spatiotemporal.

The task of the first two antinomies is, then, to provide a justification for the claim that things in themselves are not spatiotemporal. But, as the passage above-quoted makes clear, Kant thinks that he has already proved this claim in the Transcendental Aesthetic. A discussion of this first or "direct" argument is beyond the scope of this paper.² It is crucial to bear in mind, however, that not only is the argument of the Antinomies independent of the one in the Aesthetic; the Antinomies argument is actually directed at those who "did not have enough in the direct proof in the Transcendental Aesthetic." (A506/B534). The indirect argument, that is to say, is supposed to make up for any

¹ Kant-sympathetic discussions of transcendental idealism as identified by these two theses can be found in Gardner, S. (1999) pp. 99-100 and Buroker, J.V. (2006) pp. 21-2 and 17-27.

² The direct argumentation in support of transcendental idealism consists actually not of one single argument but of a series of arguments, with varying strength and (given their presuppositions) with varying relevance for contemporary discussions. One of the most important of these arguments (given in A 46-9/B64-7) turns on the explanation of the possibility of synthetic *a priori* knowledge. Kant thinks that those judgments that constitute the most important part of our scientific knowledge (namely those that express necessity and universality) could not be accounted for under the presupposition that space and time are either things in themselves or properties or determinations of things in themselves. Given that Kant thinks that this sort of knowledge is exemplified in geometry (Cf. B 3-19), the argument has sometimes been called "the argument from geometry". As is well known, however, when Kant puts forward this sort of argument, he has in mind not only geometry but also pure physics, and he intends the argument to prove not only that space and time are nothing but pure forms of intuition, but that the *objects* in space and time are nothing but appearances, i.e. representations ("modifications of our sensibility", as the passage just quoted in the text says). It is of course debatable, and well-debated, whether we are or can be in possession of necessary and universal knowledge in the way Kant understood it, and whether he is at any point in his argumentation entitled to assume that we are in possession of it (Cf. Guyer, P. (1987) pp. 354-64). See Gordon Brittan's (1978) especially chapters 1-3, for a sympathetic reconstruction of Kant's views on *a priori* scientific knowledge. A further question is whether, even granting that we have synthetic *a priori* knowledge and that we can only account for it by considering space and time as pure forms of intuition, this implies that *things in themselves* are not spatiotemporal. The claim that the implication does not hold has been called the "neglected alternative" charge: why can't space and time be *both* pure forms of intuition *and* determinations of things in themselves? For an overview of this discussion, see Gardner (1999) pp. 107-111 and Buroker (2006), pp. 64-68.

alleged inconclusiveness of the Aesthetic's argument. The mathematical antinomies are the decisive *coup de grâce* to transcendental realism.¹

Now the *coup de grâce* is, as we said, a *reductio ad absurdum*. But this means that any minimally critical evaluation of Kant's arguments must never disregard the following two considerations. First, the indirect argument works if it shows that the realist must get involved in the relevant contradictions; *it fails otherwise*. To escape the antinomy, all that the realist has to do is to show that no contradiction is implied by one of his or her relevant beliefs, the thesis or the antithesis. The second consideration is at least as important as the first one. Given that collectively the two sides of a mathematical antinomy are supposed to constitute the *reductio ad absurdum* of transcendental realism, and thereby a proof of transcendental idealism, if there is one doctrine that cannot be presupposed by either side of an antinomy, it is precisely the doctrine of transcendental idealism. If it was found that Kant's idealism plays even the slightest role in either the thesis or the antithesis argument, the realist would be entitled to leave the battlefield denouncing unfair play. It is the *realist* who is supposed to get involved in contradictions, not someone who already endorses transcendental idealism. Besides, to presuppose idealism in a "proof" of idealism is, well, circular.

The contradiction in question is that the spatiotemporal world is finite and infinite.² But before analysing the arguments for the thetical and antithetical claims, we must identify the concept of *the world* that is relevant for the present discussion:

¹ I take the phrase from Carl Posy's "Dancing to the antinomy: a proposal for transcendental idealism" In *American Philosophical Quarterly*. Volume 20, Number 1. pp. 81-94. (1983) p. 81. Henry E. Allison, who has been a prominent defender of Kant against the neglected alternative charge referred to in the previous footnote, has recently affirmed: "We learn in the Transcendental Aesthetic and Analytic that the projection of the spatiotemporal structure of our experience onto things in general is *unwarranted*, but it is only in the Dialectic that we come to appreciate that it is also disastrous. This is the clear message of the most important part of the Dialectic: The Antinomy of Pure Reason, where Kant warns ominously that 'If we would give in to the deception of transcendental realism, then neither nature nor freedom would be left' (A543/B571). The loss of nature on the assumption of transcendental realism is the main lesson to be learned from the 'mathematical antinomies'" ("Transcendental Realism, Empirical Realism and Transcendental Idealism" *Kantian Review*. Volume 11, 2006, p. 13, my italics). To say that all that the Aesthetic (and the Analytic) i.e. all that the direct argument shows is that transcendental realism is *unwarranted* is to accept that those sections of the first *Critique* do not conclusively establish that transcendental realism is untenable, let alone that it is an absurd position. Other commentators of the *Critique* who in spite of their sympathy for Kant recognize the limitations of the argument in the Aesthetic, and the fact that the Antinomies argument is supposed to make up for those limitations, are Arthur Melnick, in *Themes in Kant's Metaphysics and Ethics* (Washington, D.C.: The Catholic University of America Press) 2004. p. 171, and Sebastian Gardner, in *Kant and the "Critique of Pure Reason"*, pp. 113.

² Although Kant's antinomy consists of proofs against the finitude and infinitude of the world with respect to both space and time, I will focus on the *temporal* side of the Antinomy. This is not only for the sake of brevity, but because Kant's argument for the spatial finitude of the world takes for granted the essential premise of the argument against the temporal infinitude of the world. So there is some interpretative clarity to be gained by

We have two expressions, *world* and *nature*, which are sometimes run together. The first signifies the mathematical whole of all appearances and the totality of their synthesis in the great as well as in the small, i.e., in their progress through composition as well as through division. But the very same world is called nature insofar as it is considered as a dynamic whole and one does not look at the aggregation in space or time so as to bring about a quantity, but looks instead at the unity in the *existence* of appearances. (A 418-9/B446-7)

The concept of the world, in this sense of a mathematical whole, is the one that is said to generate the first two antinomies –the ones which accordingly are called the “mathematical antinomies”-. The concept of nature, on the other hand, is the one that is relevant for the remaining two antinomies, the so-called “dynamical antinomies”. We can therefore concentrate in the first concept, that of the world as the totality of appearances. Kant’s task is to prove that under the assumption of transcendental realism, the totality of appearances is both finite and infinite. Note that the word “appearance” *cannot* be understood in the way a transcendental idealist understands it. “Appearance” here means simply object of experience, and “mathematical whole of all appearances”, i.e. world, means the totality of objects of experience. If this neutral sense is not attached to the word, a transcendental realist has no particular reason to get involved in the present controversy. But what Kant wants to say is that this conflict between thesis and antithesis, which for a transcendental realist is unavoidable and unsolvable, can only be solved by means of transcendental idealism:

The [indirect] proof would consist in this dilemma. If the world is a whole existing in itself, then it is either finite or infinite. Now the first as well as the second alternative is false [...]. Thus it is also false that the world (the sum total of all appearances) is a whole existing in itself. From which it follows that appearances in general are nothing outside our representations, which is just what we mean by their transcendental ideality. (A506-7/B534-5)

If the word ‘appearance’ were understood in the idealist sense, meaning representation, there would be no need to *derive* any contradiction from the claim that the world is a whole existing in itself. For this claim itself would *be* the contradiction. On the contrary, the indirect proof works by genuinely presupposing that transcendental realism is not only a coherent but a true doctrine, so that in principle there should be no problem with the assumption that the world-whole existing in

focusing on the temporal side of the Antinomy. Besides, I believe that none of the claims I advance in this paper would be rendered incorrect, but rather confirmed, by a detailed analysis of the spatial side of the first antinomy.

itself is either finite or infinite. The bad news that Kant has for the realist is that s/he must affirm that it is both, and *herein* lies the contradiction required for the indirect proof, not in transcendental realism as such.

We are now ready to delve into the dialectics of Kant's First Antinomy. Let us see whether Kant succeeds in showing that the transcendental realist cannot avoid falling into a contradiction about the spatiotemporal world.

1.2. The Alleged Antinomy

1.2.1. The Thesis.

The Thesis of the Antinomy is that the world is temporally finite since it had a beginning in time. That is to say, the claim to be defended is that the series of past events is finite. This claim is arrived at by a *reductio* argument, i.e. we start by presupposing the opposite claim and then derive a contradiction from it. Here is the crucial passage (A426/B454) of Kant's proof:

For if one assumes that the world has no beginning in time, then up to every given point in time an eternity has elapsed, and hence an infinite series of states of things in the world, each following another, has passed away. But now the infinity of a series consists precisely in the fact that it can never be completed through a successive synthesis. Therefore an infinitely elapsed world-series is impossible, so a beginning of the world is a necessary condition of its existence [...].

This argument can be recast in the following seven steps:

1. Assume that the world did not have beginning in time (i.e. that the series of past events is infinite).
2. If so, then up to any given moment (up to, say, a second ago) an eternity has elapsed.
3. That up to any given moment an eternity has elapsed means that up to that given moment an infinite series of successive events has occurred, i.e. that an infinite series has been completed.
4. But an infinite series, by definition, can never be completed by means of a successive synthesis.
5. If we assume that the world did not have a beginning in time, then, we arrive at a contradiction.
6. Therefore, it is false that the world did not have beginning in time (i.e. that the series of past events is infinite).
7. Therefore the world had a beginning in time (i.e. the series of past events is finite).

As we can see, Kant's strategy is to point out that the idea of an infinite past is the idea of an *elapsed eternity* and that this idea, in its turn, entails the idea of a *completed infinity*. But, Kant claims, the notion of a *completed infinity* is a contradictory one. Why? Because to call something (a

set, a series) infinite is “precisely” to say that it cannot be completed. Kant elaborates further in the Remark on the First Antinomy: “The true (transcendental) concept of infinity is that the successive synthesis of unity in the traversal of a quantum can never be completed.” (A432/B460). In short: anyone committed to the idea of an infinite past time must be committed to the claim that an eternity has elapsed up to any given moment; now, since an elapsed eternity is a completed infinity, and since the latter is a contradiction in terms, the idea of an infinite past entails a contradiction. This is how the *reductio* works.

Interpreted in this way, the argument relies on two very specific points. The first is the claim that an elapsed eternity is a completed infinity. The second is the claim that an infinite series is *by definition* an incompletionable one. What should we make of these two points?

Let us start with Kant’s first point. What grounds do we have to accept the claim that an elapsed eternity is a completed infinity? There seems to be nothing wrong with Kant’s characterization of *eternity* as a kind of *infinity*, namely the infinite series of past events. But what about ‘elapsed’ and ‘completed’? Is the connection between these two words as unproblematical as that of ‘eternity’ and ‘infinity’? I think this question constitutes a dilemma for Kant. On the one hand, he seems to think that “elapsed” just means *passed away*. But if this is so, there is no reason why an elapsed series, i.e. a series *passed away*, is necessarily a *completed* series. Were we, on the other hand, to try to help Kant here, and say that the completeness of the series of past events is given by the fact that it has an end (terminus) at any given present moment, then the move would be like giving to our sick friend a medicine that turns out to be poison. For then he would have no right to assume the definition of ‘infinity’ that he has given, i.e. the second point of his argument would be devoid of any force. In effect, if saying that a series is completed just means that it has an end in the way that the series of past events has an end in the present moment, then one is not entitled to say that an infinite series cannot be completed, let alone *define* infinity as incompleteness. For, quite evidently, infinite series *can* be completed if “completed” means having *one* end or terminus. Think of the series of natural numbers $\langle 0, 1, 2, \dots \rangle$; it has a terminus in 0. And it would be wrong to say at this point that any infinite series must be of the form $\langle \dots, -2, -1, 0, 1, 2, \dots \rangle$, i.e. that it must lack any terminus at all; for this move would immediately invalidate Kant’s first point, i.e. ‘completed’ would

mean *having two termini*, and thus Kant would not be allowed to assume that ‘elapsed’ means *completed*.¹

In other words, either “elapsed” just means what it means, i.e. “passed away”, but then the interpretation of it as “completed” makes Kant’s argument fallacious because he *defines* “infinite” as “incompletable”; or else one could perhaps put up with an idiosyncratic Kantian meaning for both “elapsed” and “completed” (namely, a meaning which would allow us to say that the *ex hypothesi* infinite series of past events is “completed” because it has an end in the present moment), but then the concept of a completed infinity would be as unproblematical as that of the series of natural numbers. Each of the points on which Kant rests his argument constitutes, therefore, a horn of a dilemma. Either way the argument is invalid.

Some authors have tried to resuscitate Kant’s argument.² They would usually point out that Kant is not trying to expose an inconsistency in the notion of an infinite series in general, but only in that of an infinite series of past events. This is true, of course. But then the question is what makes the notion of an infinite *temporal* (past) series incoherent whereas the notion of an infinite series *simpliciter* is unproblematical? The answer, they say, is that the series of past events is what Kant would call a “totum syntheticum”, i.e. a whole which is possible only through the composition (“product”, “conjunction”, “synthesis”) of its parts. *Tota synthetica* are opposed to *tota analytica*, i.e. to wholes which are prior to their parts. Space and time are for Kant analytical wholes and they are not only infinite but even *actually* infinite. Any part of space and time is just a limit within space or time; space and time are therefore conceived of as all-embracing, actually infinite wholes.³ Thus – so the argument goes- the target of the Thesis-argument is by no means the notion of infinity, not even actual infinity; the problem is with the notion of an infinite series which at the same time is a synthetic whole.

This defense of Kant’s argument, then, depends on two claims: A) the claim that the notions of an infinite series and of a totum syntheticum are logically incompatible, and B) the claim that the series of past events must be, for the transcendental realist, a totum syntheticum.⁴ It does not suffice to defend one of these claims; either both are true or there simply is no reason to follow Kant’s

¹ The present criticism of Kant’s argument for the Thesis of the first Antinomy is, in essentials, due to Russell [1914] p. 161. Very elucidating elaborations on Russell’s criticism are to be found in Strawson (1966) pp. 199-201 and Bennett (1974) p. 119.

² Cf. Al-Azm (1972) p. 11-4, Allison (2004) pp. 366-72 and Buroker (2006) pp. 232-6.

³ These are doctrines which Kant defends in the Transcendental Aesthetic. See A245/B39-40 for space and A31-2/B46-8 for time.

⁴ I owe the formulation of question A) to Tobias Rosefeldt.

defenders in saying that for a transcendental realist the series of past events, being a totum syntheticum, can't be infinite.

Quite evidently, none of these two claims is true. And quite tellingly, in trying to argue for them, Kant's well-intentioned readers make the whole Thesis-argument vulnerable to the bluntest of criticisms leveled against it: that it presupposes idealism.

Take Henry E. Allison's argument for claim A), that the notion of an infinite totum syntheticum is logically incoherent:

The point can be clarified by noting that the concept of a totum syntheticum is here operationally defined in terms of the intellectual procedure through which it is conceived. Consequently, the problem is that the rule or procedure for thinking a totum syntheticum clashes with the one for thinking an infinite quantity. The former demands precisely what the latter precludes, namely, completability (at least in principle).¹

A totum syntheticum, says Allison, must be *completable in principle*. But what does "completability in principle" mean? Surely, whatever it means, it cannot mean completability *by us, finite human beings with the psychological capacities that we happen to have*. If it meant something like this, the transcendental realist would quite understandably say that she is not discussing about any "totum syntheticum" at all; she is discussing about *the world*, and what defines her as a transcendental realist is that she regards the world's existence as independent of any mental capacities. So if the argument is going to be of any interest for the realist, "completability in principle" must mean something like *completability by some idealized rational subject*. But why is this idealized rational subject, which could be God Herself, assumed to be finite? This assumption begs the whole question.²

¹ Allison (2004) p. 370.

² These considerations show that Allison is mistaken when he says that the characterization of infinity that Kant provides and he (Allison) defends –as inexhaustibility in principle of the process by which the members of a set are enumerated– is compatible with Russell's conception according to which the members of a (finite or infinite) set are given all at once by the defining property of its members, "and presumably [the Kantian definition is compatible] with the Cantorian conception as well" (p. 369). Allison goes on to point out in a footnote that the Cantorian definition of infinity is the one according to which a set is infinite if there is a one-to-one correspondence between itself and a proper subset of it. Now, when one assumes, as Kant did and Allison explicitly recognizes, that the enumeration of the members of a set is a process by which a number *n* is assigned to each of those members, and that *n* is a member of the series of natural numbers, then there is all the difference in the world between the Cantor-Russell conception, on the one hand, and the Kant-Allison conception, on the other. As we know, any member *n* of the series of natural numbers, however big, is *finite*. Thus, acceptance of Kant's definition of infinity is tantamount to saying that a set is infinite when its members cannot be enumerated by an idealized but *finite* rational subject. But if this is so then the Russellian objection applies. To presuppose that the idealized rational subject who would carry out the enumeration is a finite

Much the same mistake is made by Jill Vance Buroker in trying to defend claim B), that a transcendental realist must conceive of the world as what Kant calls a totum syntheticum. She writes:

Because the world is *there to be encountered* by the subject, the parts must be thought of as *given independently of the constructive process*. Since humans do not intuit the world as a whole, Kant seems justified in claiming that the idea of the world-whole arises by synthesis of the parts. Now since all synthesis is successive for humans, whether the parts exist successively or

being is to beg the question. And if the subject is not idealized then Kant's argument incorporates, "more or less surreptitiously, that reference to mind by which all Kant's philosophy was infected". (Russell [1914] p. 161). On the Cantor-Russell conception, in contrast, one is not forced to see every enumeration as the assignment of a finite number to the elements of a set. There are infinite or transfinite numbers for them, like the number which is assigned to the series of natural numbers itself, $\aleph_0$. This means that if anyone raised the question of "enumeration" or "completion by synthesis", as Kant does, they can say that it could well be the assignment of an *infinite* number to that set. Hence Russell's famous passage: "[...] when Kant says that an infinite series can 'never' be completed by successive synthesis, all that he has even conceivably a right to say is that it cannot be completed *in a finite time*. Thus, what he really proves is, at most, that if the world had no beginning, it must have already existed for an infinite time." (Russell [1914] p. 161). Particularly vulnerable to this reply is Arthur Melnick's defence of the thesis-argument in his (2004). In defence of Kant's definition of the infinity of a series as its incompleteness, Melnick says: "Kant is denying, that is, that an infinite regress can be (represented as) now done with [...]. Indeed, if it is now done with (now proper to be beyond all stages) then what was it proper to be doing a moment ago? Producing a *k*th downbeat? That's impossible since I would then be now up to $k + 1$. Similarly, there is nothing it was proper to be doing two moments ago, three moments ago, etc." (p.177) Evidently, if Melnick's k is a finite number, like 1, 2 or 3, then neither he nor anyone would dare to say that it is impossible to be up to $k + 1$. But if k is an infinite number, it is an argument for the impossibility of being at $k + 1$ that needs to be provided rather than presupposed (compare set theory's $\aleph_0 + 1$ or $\omega + 1$). This also shows that it would be wrong to suggest, as some authors do (e.g. Melnick (1989) p. 331 and Buroker (2006) pp. 233-4), that in the Remark on the Thesis of the First Antinomy Kant puts forward the same defense of the concept of the actual infinite that he had advanced in a footnote to the first paragraph of the 1770 *Dissertation*. It is true that in both passages Kant criticizes a certain way to reject the possibility of an infinite magnitude, namely that rejection which is based on a defective concept of the infinite as a *maximum* (largest possible). The argument that Kant criticizes is the next one: An infinite magnitude would be a largest possible magnitude, but given that one or two units (of an arbitrarily chosen measure) can always be added to any magnitude, the notion of an infinite magnitude is absurd. But, as Kant goes on to say in both the *Dissertation* (2:389) and in the Remark on the Thesis (A430-2/B458-60), this is not the correct concept of infinity. The correct concept of infinity is not one by which one tries to determine, absolutely, *how big* a given magnitude is. Rather, by thinking of a magnitude as infinite, one only thinks of the *relation* of that magnitude to a multiplicity of arbitrarily chosen units, viz. the relation of being *larger than any number* of such units. Thus, both in the *Dissertation* and in the *Critique*, the correct concept of infinity is said to be relational (as opposed to "absolutist"). However, there is an important aspect in which the two texts differ. In the *Dissertation*, given that the relation in question is that of a quantity to a given enumeration of arbitrarily chosen units (or rather to its result: the natural number n assigned to the multiplicity of units), Kant understandably points out that an idealized subject could see an infinite magnitude "at a single glance". Enumeration would not be at issue in the idealized case. This is why the *actual* infinite is said to be possible. As we can see, the Kant of the *Dissertation* is not miles away from Russell: either enumeration ("synthesis") is not at all at issue or, if it is, then to presuppose that the idealized enumerating subject must be finite is to beg the question. For some reason, this consideration is completely absent from the Remark on the Thesis of the First Antinomy. Accordingly, here Kant speaks only of *mathematical* infinity (which he explains as the relational concept we mentioned above); regarding *actual* infinity Kant remains completely silent.

simultaneously, the *thought* of its composition requires a successive synthesis.¹

The first two sentences of this passage can be granted. The first one is actually an accurate description of the transcendental realist's position; the second one is not very problematical at the moment: we have not intuited the whole spatiotemporal universe, so our idea of it must have arisen through the successive synthesis or its parts. So far so good. But why must the fact that the process by which *humans* think of the world is a successive synthesis -a psychological or anthropological fact- have any bearing on how *the world is*? The transcendental realist, at any rate, denies precisely this.

In conclusion, Kant's claim according to which a transcendental realist must be committed to the claim that the spatiotemporal world is finite is incorrect. Kant fails to provide a proof of the Thesis of his alleged First Antinomy.

1.2.2. The Antithesis

Let us now turn to the Antithesis. The claim to be defended is, of course, that the series of past events is infinite. And this defense comes again in the form of a *reductio* argument. The crucial passage (A428/B457) is the next one:

For suppose that [the world] has a beginning. Since the beginning is an existence preceded by a time in which the thing is not, there must be a preceding time in which the world was not, i.e. an empty time. But now no arising of any sort of thing is possible in an empty time, because no part of such a time has, in itself, prior to another part, any distinguishing condition of its existence rather than its non-existence (whether one assumes that it come to be of itself or through another cause). Thus many series of things may begin in the world, but the world itself cannot have any beginning, and so past time is infinite.

The argument can be reconstructed in the following six steps:

1. Suppose that the world has a beginning in time.
2. That the world has a beginning means that it comes into existence preceded by a time in which it (the world) *was not*.
3. But since we are speaking about the *world* (the *totality* of appearances), a time in which the world was not is an *empty time*. Thus if the world has a beginning, this means that it came into being in an empty time.

¹ Buroker (2006) p. 235.

4. But no coming into being in an empty time is possible at all, because no part of an empty time possesses, as compared with any other, a distinguishing condition of existence rather than nonexistence.
5. Therefore the world couldn't have had a beginning in time.
6. Therefore the world is infinite with respect to past time.

This argument assumes that time is infinite, since step 2 and 3 imply that for any finite series of past events, there is an empty time preceding it. The argument also assumes that time is real independently of events in time, for otherwise the notion of an empty time preceding worldly events is meaningless. Another important assumption is that empty time is completely homogeneous, i.e. that any part of empty time is indistinguishable from any other part. Given these assumptions, the argument clearly rests on a Kantian version of Leibniz's principle of sufficient reason, "viz. that nothing happens without a reason why it should be so rather than otherwise."¹

In fact, in his *Correspondence* with Samuel Clarke, Leibniz used a similar argument against the Newtonian view of time (and space) as absolute.² Since Newtonians conceived time as *absolute and infinite*, they could not avoid the view that the world is eternal, because they could not give a reason why God did not create it "some millions of years sooner". Leibniz proceeds: "For since God does nothing without reason, and no reason can be given why he did not create the world sooner; it will follow, either that he has created nothing at all, or that he created the world before any assignable time, that is, that the world is eternal."³

Kant's argument is a secular version of Leibniz's argument. Kant is saying that a condition of existence *must be able to be distinguished* if the world actually came into existence. So Kant is presupposing two things here: first, that no existence and no coming to be of anything –no "happening"– is possible without its determining condition; second, and in line with transcendental realism, that the determining condition of existence *must be able to be distinguished* by our cognitive powers. The first of these presuppositions is actually what, according to Kant, generates the entire antinomial conflict, namely the principle that *If the conditioned is given, then the whole sum of conditions, and hence the absolutely unconditioned, is also given.* (cf. e.g. A307-8/B364-5,

¹ §1 of Leibniz's second paper in Alexander (1956).

² This historical point about the source of the Antithesis argument is forcefully defended by Al-Azm (1972), chap. 1.

³ §14 of Leibniz's forth paper in Alexander (1956). Leibniz was convinced that this argument refuted the absolutist conception of time, and that it proved his own relationist theory. We need not concern ourselves with this controversy at the moment.

A409/B436 and A497/B525). We can call this principle the principle of the unconditioned.¹ Some commentators have pointed out that this principle is Kant's elaborate version of the principle of sufficient reason.² This suggestion correct, at least in the context of the Antinomy, for the equivalence between the two principles obtains only when the principle of sufficient reason is taken to its ultimate consequences and when one assumes the truth of transcendental realism.

Kant thought that the principle of the unconditioned was present also in the Thesis of the Antinomy, for he thought that the unconditioned is differently considered by each side of the conflict (A417-8/B445-6): Both sides agree that the unconditioned is given once the *whole series of conditions* is given, i.e. when there is *absolutely nothing (relevant) left to be explained*.³ But each part of the antinomy has a distinct way of claiming that the unconditioned must be given. In the Thesis argument, this claim is *argued for* by deriving a contradiction from the presupposition that it is not given. The principle of the unconditioned is therefore a conclusion of the argument. By contrast, in the Antithesis argument the principle of the unconditioned is one of the crucial *premises* upon which the whole argument rests.

The discussions generated by this argument concern not so much whether the conclusion follows from the premises, but whether or not to accept those premises and, in that case, how to interpret them. Some critics, for example, grant the definition of "beginning" in step 2 ("a beginning is an existence preceded by a time in which the thing was not") but, since we can interpret the Leibnizian-Kantian claim that we cannot conceive a temporal limit of the world as a claim about an *epistemological* predicament, not about a *metaphysical* necessity, those critics think that the only thing that follows is that the question "Why did the world begin when it did?" is an unanswerable one, not that the world is eternal.⁴ This criticism is incorrect because for a transcendental realist who accepts the principle of sufficient reason –and nobody else is taken by Kant to endorse the Antithesis argument–, reality in general is thoroughly intelligible and we *must be able* to access this intelligibility by means of our faculty of reason.⁵ So if evidence for a beginning of the world is *inconceivable*, the realist concludes that a beginning of the world is metaphysically impossible. The

¹ This is what Michelle Grier (2001) refers to as "P₂", but her nomenclature is designed in order to clarify Kant's doctrine of reason's "inevitable" transcendental illusion. Now, for reasons that hopefully will become clear in what follows, we are not allowed to assume this Kantian doctrine. So we are well advised to reflect our circumspection in our terminology.

² Cf. e.g. Walsh. W.H. (1975) pp.204-5 and Gardner (1999) p. 218.

³ See Nieman (1994) p. 63, Wiggins (1996) pp. 120-1, and Gardner (1999) p. 217.

⁴ See Strawson, P. (1966) pp. 177-178 and Guyer (1987) p. 408.

⁵ See Wiggins (1996) pp. 120-2 and Nieman (1994) p. 13.

word “evidence” can be misleading here. It certainly doesn’t mean empirical evidence. In the present context, to say that evidence for something’s being the case is possible means that such a fact is identifiable by a *rational* subject, be it a human being, a Laplacean demon, or God.

But there is, as in the case of the Thesis, a criticism which is not so easily disposed of. The attack consists in a rejection of the definition provided for the word ‘beginning’. Bennett has a clear formulation of this line of criticism and hence we will focus on it in what follows.¹

Bennett admits that the notion of ‘time before the first event’ (empty, extra-mundane time) should be rejected. And he admits this in accordance with the argument presented in defense of the Antithesis, namely, the (Leibnizian) argument that no evidence for the existence of such a time is conceivable at all. But Bennett thinks that the rejection of the notion of a time before the first event does not oblige him to draw the conclusion that there was no first event. From the impossibility of a pre-mundane time Kant infers the impossibility of the first event. But why? We could also draw the conclusion, says Bennett, that *if there was a first event, then it happened at the first time*.²

What Bennett is challenging is the presupposition that the concept of *beginning* must include the idea that there was a time in which the thing or event that begins was not. Bennett doesn’t give an explicit formulation of his alternative conception of beginning, but it is easy to know what he must have in mind, at least in the present case. A beginning of the world is an absolutely first event, and an event is the first event *if there are no events previous to it*. This definition does not imply that there was a time when the world was not.

To be sure, if this is right, then time would no longer be conceived of as infinite, which was one of the presuppositions of the Antithesis argument. But a commitment to the infinity of time is not amongst the logical or systematical requirements for entering the present cosmological discussion. (Kant hasn’t given an argument to the effect that a transcendental realist has to be Newtonian in this respect). So if one is to reject the idea of the finitude of time one has to find an independent argument. But Bennett wouldn’t lay too much hope on that project. He thinks that the idea of a first time, and therefore that of the finitude of the series of past events, is perfectly intelligible:

People who think there is a radical incoherence in the idea of a first time never succeed in displaying one. The logic of it is as follows. Let *H* be some unique historical event, say the death of Berlioz, and let *n* be the number of years back from *H* to the first event. Then the phrase ‘*n* years before *H*’

¹ It must be noted, however, that the argument was first stated by Leibniz himself. See § 56 of his fifth paper to Clarke.

² Ibidem.

names the first time, and any phrase of the form '*k* years before *H*', for *k* > *n*, makes sense but does not refer to any time.¹

So the overall picture whose intelligibility Bennett is defending is the next one: a natural number *n* is assignable to the whole series of past events. Of course, for any natural number we can think of, we can also think of its successor. But if the world had an absolutely first event the successor of *n* is not assigned to any member of the series of past events (for there are no more such members), and it doesn't refer to an empty time either (since the first event happened at the first time).

Allison has tried to defend Kant also on this front. That is to say, Allison tries to defend the view that the transcendental realist must be committed to the impossibility of an absolutely first event. Thus, Allison points out that if by "event" one means a change of state or alteration of a thing, given that the notion of a change implies that there was a previous time in which the thing in question was in a different state, then Bennett's notion of a first event-at-a-first-time must be incoherent. Now, Allison notes that this would not be sufficient to discard the possibility envisaged by Bennett. For Bennett can deny that the first event is a change. Allison replies that if an event is not a change, then we cannot meaningfully say that it occurred at the first (or at any other) time. Now, Allison does not argue for this reply. The only thing he says is that "it is a condition of possibility of conceiving of the *change* of a thing in time that we are able to contrast the state of a thing at an earlier time with its state at a later time."² But this reply cannot be right, for Allison had granted a few lines earlier that it is conceivable that the first event is not a change. So his point about a previous time as a condition of conceivability of a *change* is irrelevant. Even if one agrees with Allison that the notion of a first *change* at a first time is incoherent, no defense has been provided for the claim that the notion of a first *event* at a first time is incoherent. It is the latter which has to be done if the Antithesis is to be defended.

In conclusion, Kant's claim, according to which the transcendental realist must be committed to the view that the series of past events is infinite, is also incorrect. Kant has failed to provide a proof of the Antithesis of the First Antinomy.

¹ Ibid. p. 161.

² Allison (2004) p. 376. Italics are mine.

1.3 Does Transcendental Idealism Strictly Follow from the Alleged First Antinomy?

If what I've been defending is correct, there are at least two legitimate ways in which a transcendental realist could escape the First Antinomy. One of them is to reject the Thesis argument on the grounds that its strongest and most fundamental premise –the claim that an elapsed eternity would entail the completion of an infinite series, which is impossible- rests on the failure to realize that “a series both infinite and completed is problematic only on the assumption that it had a beginning” (in Strawson's words). That is to say, the argument assumes precisely what it shouldn't assume, i.e. that the series of past events had a beginning. The second way in which a transcendental realist could escape the Antinomy is by challenging step 2 of the Antithesis argument, the definition according to which “a beginning is an existence preceded by a time in which the thing was not”. In the present case, this can be rejected by defining the beginning of the world as its first event, and defining the first event as that which occurs earlier than every other. The result would be the possibility that *the first event happened at the first time*.

In this third section I'll explore a third way to escape the Antinomy. Let us suppose, for the sake of the argument, that the Thesis and the Antithesis had both been validly established. After all, the defender of the Thesis might have in mind the Zenonian paradoxes and Aristotle's reply to them, and s/he therefore would reject the idea of an actual infinity, at least as regards “*tota synthetica*” like the empirical objects. As regards the Antithesis, the realist might be committed to the claim that time is not only infinite but also absolutely real, and thus s/he would be unable to avoid the conclusion that the material universe is infinite (other than by an appeal to the extraordinary nature of God's will, in the way Clarke replied to Leibniz). So let us assume that the realist finds both arguments convincing. We have, then, an Antinomy and we have to look for a way out of it. Kant thought that the only right thing to do at this moment is to reject transcendental realism and to adopt transcendental idealism. This is the sense of Kant's “indirect proof” of transcendental idealism: that this doctrine can solve a problem which under the assumption of transcendental realism is insoluble. In what follows I will argue that the only conclusion we are allowed to draw from the existence of the present Antinomy is *not* that the spatiotemporal world is a world of phenomena, as opposed to a world of things in themselves, but that the concept of the world is a self-contradictory one and that it therefore does not refer to an existing object.

Let us first look at Kant's solution. As is well known, the key to understanding Kant's diagnosis and solution of the problem is his distinction between the *analytical* and the *contradictory* kinds of opposition (A503/B53). He provides the example of the sentence: "All bodies have either a good smell or a smell that is not good". Although this sentence is a disjunction, it is not an exhaustive one, because there are bodies which do not fall under either of the opposed predicates. There are bodies which have no smell at all and, having no smell, they have neither a good smell nor a smell that is not good. Only the ungrounded presupposition that all bodies have a smell could give the disjunction an appearance of exhaustiveness. Since this kind of opposition is based on an unacceptable assumption, Kant calls it a "dialectical" one. On the other hand, an "analytical" or "contradictory" opposition, is the one in which the alternatives simply negate one another, as in "All bodies are either good-smelling or not good-smelling". The truth of this disjunction is based on the principle of non-contradiction and the disjunction is obviously exhaustive.

Kant thinks that in the case of the first Antinomy what we have is a dialectical and not an analytical opposition. Both the Thesis and the Antithesis rely on an "ungrounded presupposition", namely the presupposition that the world is a thing in itself. These are Kant's words:

If one regards the two propositions, "The world is infinite in magnitude," "The world is finite in magnitude," as contradictory opposites, then one assumes that the world (the whole series of appearances) is a thing in itself. For the world remains, even though I may rule out the infinite or finite regress in the series of its appearances. But if I take away this presupposition, or rather this transcendental illusion, and deny that it is a thing in itself, then the contradictory conflict of the two assertions is transformed into a merely dialectical conflict, and because the world does not exist at all (independently of the regressive series of my representations), it exists neither as *an in itself infinite* whole or as *an in itself finite* whole. It is only the empirical regress of the series of appearances, and by itself it is not to be met with at all. Hence if it is always conditioned, then it is never wholly given, and thus does not exist as such a whole, either with infinite or with finite magnitude. [A504-5/ B532-3]

The structure of Kant's argument is therefore the next one: If you think that the spatiotemporal world is something existing in itself, the paradox which constitutes the first Antinomy arises. Take this inadmissible presupposition away and the paradox disappears. But given that the disjunction between transcendental realism and transcendental idealism is an exhaustive one, the rejection of the latter is tantamount to an endorsement of the former. The antinomial conflict would thus provide an indirect proof of transcendental idealism.

I think that Kant's argument is incorrect. It is true that if one assumes that the world is something existing in itself, and one generously grants Kant the validity of his "proofs", the First Antinomy arises. But it is not *on account of being a realist* that one is forced to say that the world is both finite and infinite. None of the analyzed arguments turns on the presupposition that spatiotemporal objects exist independently of our cognitive faculties. This presupposition is semantically different from the crucial assumptions of the two arguments (world as totum syntheticum [Thesis], absoluteness-plus-infinitude of time coupled with the principle of the unconditioned [Antithesis]). So it is true that if one grants these assumptions one must arrive at the conclusion that the world is both finite and infinite, which means that the concept of the world is self-contradictory. But then the only thing that we are allowed to do is to treat the concept of the world as a pseudo-concept and to treat the world as a pseudo-object. Now if the transcendental realist treats the concept of the world as unproblematic, it follows that s/he is wrong in so doing. But it would be wrong to think that only transcendental idealists have the privilege of being able to treat the concept of the world as a pseudo-concept. There is no logical incompatibility between the claim that Mars and Jupiter exist outside the "modifications of our sensibility" and the claim that the concept of *the absolutely unconditioned totality of spatiotemporal objects* does not refer to any existing object. The realist could even adduce that s/he refuses to treat this concept as a referential one because otherwise the Antinomy arises. This concession does not make a realist into an idealist.

One of the few commentators to have noticed the flaw is Moltke Gram:

From [the Antinomy] it follows that the referent [of the expression 'the totality of appearances'] is nonexistent. Now, if this is the real structure of Kant's argument, then he has proved that the totality of appearances does not exist as a thing in itself; for he has established that it does not exist at all, and a fortiori it does not exist in itself. But this conclusion has nothing to do with the characteristics peculiar to things in themselves. Nor does it turn on the notion of transcendental ideality. It is an immediate inference from the conclusion –which Kant's argument does establish– that the object called the totality of appearances does not exist at all.¹

If the concept of the world is self-contradictory, it follows (at least for non-Hegelians) that it doesn't refer to anything real. The world is not a thing. Rejection of world-reificationism, not of

¹ Gram, M. (1967) p. 512.

transcendental realism, is what follows from the fact that the concept of the world is self-contradictory.¹

I have found two different replies to this criticism in the literature, one by Carl Posy and the other one by Allison. Both of them are incorrect.

Posy notes that “distrust of abstracta like the sum of all empirical objects”, is in principle, perfectly compatible with the view that objects of experience exist independently of our representative faculties. So why does Kant feel justified in moving from the proof that the concept of world is self-contradictory to a denial of the transcendental reality of e.g. everyday objects? Posy has an answer to this question. The transcendental realist *must* be committed to the claim that the world exists as an empirical object, having a definite magnitude, and since the proofs of the Thesis and of the Antithesis (granting all that has to be granted) do show that the world is neither finite nor infinite, they constitute a rejection of transcendental realism. The obvious question that Posy now has to answer is, of course: Why must a transcendental realist necessarily be committed to the claim that the world has a definite magnitude? And he answers this way:

The realist’s right to speak of the world in the “Antinomy” and to construe it as an object on par with other empirical objects follows from the two notions of an unconditioned which are forced on those realists caught by the first antinomy. If the unconditioned is the entire series of past durations [Antithesis], then that series is the temporal world, and homogeneity guarantees that it exists as an observable empirical object. If

¹ Gram is not the only one who suggests that anti-reificationism, rather than transcendental idealism, is the only conclusion to which Kant would have been entitled had the antinomy been genuine. Michael Hallett (1984) p. 226, e.g., says: “Indeed, it is a crucial premise of the Antinomy that the world *can* be treated as an ordinary physical object of the same kind as my pen, that is, as an object which, by the law of excluded middle, either has or has not ordinary spatial and temporal properties. (Note the involvement of excluded middle in Kant’s proofs. The proofs are often cast as *reductios*; we start from $S(a)$, derive a contradiction and conclude $\sim S(a)$, the assumption being that a , the world, must possess either S or $\sim S$). It is just this assumption that the world is an object in the ordinary sense that Kant regards as the source of the trouble which the Antinomies present.” A.W. Moore (2001) pp. 88 and 89 says: “[Kant] denied that there was any such thing as the physical world *as a whole* [...]; the debate about whether the physical world was infinitely old; the debate about whether it was infinitely extended in space; and the debate about whether it was infinitely divisible. These arose only on the assumption that there *was* a physical world, existing as a whole (that is, capable of being given to us in experience). Granted this assumption, it was a perfectly legitimate question whether this world was infinite or finite on each of the specified respects; and it had to be one or the other. But then, on Kant’s view, there was bound to be irresolvable controversy. [... Therefore] we should drop the common underlying assumption. In each of the specified respects the world –as a whole– was *neither* infinite *nor* finite. It did not exist.” Graham Priest, (2002) p. 96 says: “Since [the arguments of the antinomies] are perfectly sound, it follows that both the Thesis and the Antithesis are false, but the only way that statements of the form ‘ S is P ’ and ‘ S is not P ’ can both be false is that S does not exist.”

the unconditioned is a first element [Thesis], then again it must be observable. We would then have a discretely ordered series of durations with a first and last element (namely the beginning and now). This series would constitute the world. But the series can be enumerated by a successive empirical synthesis, and so it would exist empirically. In both cases *the world* is identical to the “whole series of representations” whose existence is demanded for the realist, by the dialectical syllogism.¹

The crucial point of this passage is when Posy says “homogeneity guarantees that the world exists as an observable empirical object”. The word “observable” here is not to be understood in a narrow sense. An observable empirical object is simply an object of experience, like my pen, Mars and the Solar System. So Posy is simply insisting on a point we have already noted: realists tend to assume that the world is homogeneous with the objects we come across with in experience, only much bigger and older. But although this point might be statistically interesting it does not express a logical necessity. That is to say, even if one agrees that most –or even all- transcendental realists view the world as homogeneous with empirical objects, this doesn’t mean that transcendental realists, as such, are logically bound to view the world this way. As far as logic is concerned, transcendental realists could view the world as a *totum analyticum*, and since objects of experience are *tota synthetica* (and therefore have a definite magnitude), the world would be regarded as heterogeneous with respect to empirical objects. The Antinomy would therefore either not arise (the realist would refuse to accept the Thesis argument), or else after its emergence the realist would be entitled to reject the view that the world is homogeneous with the empirical objects. So realists are not logically bound to reject transcendental realism and accept transcendental idealism.

Allison also deals with the present criticism. He explicitly refers to Gram’s objection and then goes on to say:

Although this line of criticism seems plausible, it suffers from a failure to consider the essential role of transcendental illusion in Kant’s analysis. In the present context, this illusion explains why it seems perfectly natural, even unavoidable, to take ‘world’, or its equivalents, as a naming expression. Continuing with the linguistic analysis, the term seems to function more like ‘army’ than, say ‘the average family’ in designating an actual collection, indeed, one that is “given” (in an empirical sense) and with respect to which the determination of magnitude seems perfectly in order.²

¹ Posy, C. (1893) p. 91.

² Allison, H. (2004) p. 390.

This reply by Allison is invalid. To see this, note that Kant, while recognizing the “unavoidability” of the illusion, takes himself to be adopting a philosophical position in which the illusion and its invalidity have already been detected. Now, it is to *this* critical philosopher, and not to any dogmatic champion of either the Thesis or the Antithesis, that Gram’s criticism is directed. It is towards this critical philosopher, who claims that he’s no longer committing the errors that the illusion leads him to commit, that Gram addresses his criticism when he says that all that he was allowed to infer was that the totality of appearances does not exist (and *a fortiori* that it does not exist in itself), *not* that the totality of appearances does not exist *outside representations because it is nothing but the totality of representations*. Another way to appreciate the incorrectness of Allison’s reply is to note that Kant’s doctrine of transcendental illusion is not independent of his doctrine of transcendental idealism. The relation between these doctrines is a subtle one. One can succinctly express it by saying that the doctrine of transcendental illusion is Kant’s explanation of how the fallacious arguments of the *Dialectic* arise. And at the most basic level of explanation, Kant says: the fallacies emerge when two conditions are given: 1) appearances are conflated with things in themselves, and 2) one assumes the truth of the principle of the unconditioned. And in fact the principle of the unconditioned is not said to be *illusory* by itself but, as Grier puts it, “precisely on the condition that one is a transcendental realist”.¹ Thus you cannot follow Kant in thinking that transcendental illusion obtains if you have not accepted his view that categories can only be applied to phenomena and not to things in themselves. But you cannot trace the distinction between phenomena and things in themselves without being a transcendental idealist: according to Allison himself, the distinction between phenomena and things in themselves “constitutes the heart of transcendental idealism”. So when you are asked to accept what Kant has to say about the transcendental illusion, you are thereby asked to accept what he has to say about the “conflation” between appearances and things in themselves, i.e. you are asked to be a transcendental idealist.

The result is that Allison’s appeal to the doctrine of transcendental illusion in his argumentative defense of Kant’s indirect proof of transcendental idealism is not a legitimate step. Allison’s response to Gram’s criticism is summarized by the claim that “we cannot simply deny that ‘world’ is a referring expression without also exposing the underlying illusion that naturally leads us to view it as such”.² If in this sentence we understand the word “naturally” in a non-Kantian way, we must accept Gram’s point: it might be true that we “naturally” (i.e. “usually” or “naïvely”) assumed that the world is a “thing” but, thanks to the discussion of the Antinomies, we realized that such an

¹ Grier, M. (2002) p. 180.

² Allison. (2204) p. 390.

assumption leads to contradictions, and must therefore be rejected. But this conclusion is not enough to establish Kant's transcendental idealism, because the denial of the claim that the world exists is not equivalent to the claim that the world is identical with the series of representations, or that it is a series of appearances as opposed to things in themselves. Now, Allison assumes that we should understand the word "naturally" in the Kantian sense. But the problem with this is that Kant's doctrine of the illusions to which we are "naturally and unavoidably" led *presupposes* his doctrine of transcendental idealism.¹

1.4 Why I Am Not a Transcendental Idealist

We have seen that transcendental idealism is not the only possible way out of the (alleged) antinomy. If the concept of the spatiotemporal world had been shown to be contradictory, we could draw the conclusion pointed out by Gram: there is nothing out there corresponding to our concept of the world, there is no such thing as the world. A perhaps weaker, but importantly different conclusion would be that our commitment to the spatiotemporal world's existence is similar in kind to our commitment to the existence of a square circle. If we reject the existence of 'truly contradictory objects' the two conclusions are more or less the same: if the antinomy is genuine, the concept of the world is contradictory, and hence there is no object corresponding to it. But we are not allowed to carry out such a rejection by our previous discussion.

¹ In the (2007) debate, Allen Wood makes a criticism of Kant's "indirect proof" which is more general than that of Gram, although both criticisms are clearly related. Wood's point is that if the (mathematical) antinomies are generated by the conjunction of the principle of the unconditioned and transcendental realism, why not reject the principle of the unconditioned and keep transcendental realism? If this option is as plausible as the one Kant took, i.e. to reject transcendental realism and keep the principle of the unconditioned (now re-interpreted under the transcendental idealist's lights as a regulative principle), then Kant's claim to have provided a proof of transcendental idealism is unjustified (pp. 7-9). Allison replies by denying the availability to the realist of the option envisaged by Wood. The principle of the unconditioned, according to Allison, "[...] cannot be discarded, since it is inseparable from the theoretical use of reason. Consequently, transcendental realism makes it impossible to avoid being deceived by the illusion [...]" (Ibid, p. 27). What is clearly not available, in my opinion, however, is Allison's reply. Given that he thinks that the principle of the unconditioned cannot be discarded, he must agree (as he does, following Grier) that this principle can be interpreted either in a realist or in an idealist fashion. But if the principle can be interpreted either way, what does Allison mean when he says that *it* can't be discarded? There are only three options here. Either *it* – the principle – is being interpreted by Allison realistically, idealistically, or neutrally. In fact, the first and the third options coalesce into one, which is the only correct one at the moment: for the purposes of the indirect proof, we are assuming the truth of transcendental realism, so the principle (neutral in itself) is being interpreted realistically. But then Wood's envisaged option is still available: Interpret the principle as a regulative one, but keep realism. Where is the logical impossibility? If, on the other hand, Allison is already interpreting the principle idealistically, he is incurring in circularity again.

Seen as an unjustified solution to a false problem, transcendental idealism might be thought to be an easily dismissible doctrine. It is not. True, we saw in subsection 1.1.1 that the other motivation that Kant had for the doctrine (namely, that it is, according to Kant, the only way we have of explaining the possibility of *a priori* knowledge) cannot be turned into a valid argument in support of transcendental idealism either. But this only shows, in conjunction with the findings of the previous subsection, that transcendental idealism is not *justified* by the two arguments that Kant provides for it. This is not an unimportant claim, but it is not the only important claim to be made about the doctrine. We need to see whether transcendental idealism was anyway a *plausible* solution to the antinomy, even if it was not the only possible one. We need to establish this because in the following chapters we will be concerned with other antinomies which, unlike Kant's, do follow from natural presuppositions. And in the final chapter of this investigation (chapter IV) I will defend the view that there are structural and conceptual similarities between the two sorts of antinomies, even if the statement of those antinomies examined in the following chapters is more convincing than Kant's statement of his alleged antinomy. Thus, as I will defend, we will be facing a problem *similar* in some respects to the one that Kant says we face with the antinomy. It is therefore worth asking whether Kant's "solution" to his "problem" was a plausible one: when we face a more convincing but similar problem, we might be tempted to adopt a solution which is similar to Kant's in the relevant respects. In what follows I should like to argue that such a temptation must be resisted. Better said, I would like to suggest that the temptation is illusory. Kant's transcendental idealism is a highly implausible doctrine. It faces enormous problems in its own right. It can hardly be seen as a solution to other (real or imaginary) problems.

The claim that Kant's transcendental idealism is an implausible doctrine has been put forward innumerable times by philosophers and interpreters who belong to the most varied traditions. Almost as many times, however, has the doctrine been defended, if not quite as it stands in many passages of the *Critique*, at least as it can be interpreted on a sympathetic reading. It would therefore be naïve to think that such a long-standing debate could be settled in this subsection. My aim in what follows is simply to point out what I consider some of the most pressing objections to the doctrine, even when it is interpreted sympathetically.

Under a "harsh" reading, it is clear that the doctrine is self-stultifying. We said in subsection 1.1.1. that one of the claims that identify transcendental idealism is the Non-spatiotemporality claim: things in themselves are not spatiotemporal. But the other identifying claim is the Unknowability claim: things in themselves are unknowable. How are these claims to be made compatible with one another? How can we say that things in themselves are non-spatiotemporal if we also want to say

that we don't know them? Kant's claim is not the skeptical one that we don't know whether things in themselves are spatiotemporal or not. It is an ontological denial.¹ Perhaps the doctrine can be interpreted so as to allow two different ways in which we can be said to know something. Kant's intention would be to say that we cannot know things in themselves in the way we ordinarily know objects of possible experience (with the use of our faculties for knowledge, sensibility and understanding), whereas, of course, if pressed, he would allow for other ways of knowledge in which we can be said to know that things in themselves are not spatiotemporal. The problem with this is that Kant *needs* to make some claims about things in themselves, and those claims are not only of the same kind as ordinary claims about empirical objects, but, more importantly, given the Non-spatiotemporality thesis, those claims are hardly intelligible. Kant needs to say that things in themselves *exist* (BXX, XXVI, 308-9) . He also needs to say that they *affect* sensibility (§32 of the *Prolegomena*), thus partly *causing* the phenomena (B522) or serving as their ground (B723-4).² It is very, very hard to understand what existence, causation and affection should mean if they are not supposed to be understood in a temporal sense. If to this we add that Kant sometimes says that we cannot "make intelligible" the categories (like *existence* and *causality*) unless we think of them as applying to temporal objects (A240-2/B300-1), the self-stultifying nature of the doctrine is evident.

Now, perhaps Kant doesn't really need all these claims about the things in themselves. Perhaps transcendental idealism can do without the existence of things in themselves, their affection of sensibility and their causation of phenomena. That Kant made those claims is undeniable, but perhaps the doctrine can be saved from the charge of inconsistency if the existential commitment

¹ Cf. P.F. Strawson (1966) p.38: "The doctrine is not merely that we can have no knowledge of a supersensible reality. The doctrine is that reality is supersensible and that we can have no knowledge of it." This formulation of the doctrine is easily seen to imply Kant's ontological denial as soon as we recall Kant's claim in the Transcendental Aesthetic that space and time are "nothing but" the pure forms of sensibility. (Cf. A26/B42 for space and A 32/B49 for time). Reality being supersensible, and space and time being nothing but forms of sensibility, it follows that reality is not spatiotemporal. That transcendental idealism entails an ontological denial has not only been pointed out by "harsh" readers, by the way. Cf. Gardner, S. (1999) pp. 99-101.

² Perhaps the passage in which Kant most clearly makes all these claims together is the one from the *Prolegomena* referred to in the text: "In fact, if we view the objects of the senses as mere appearances, as is fitting, then we thereby admit at the same time that a thing in itself underlies them, although we are not acquainted with this thing as it may be constituted in itself, but only with its appearance, i.e. with the way in which our senses are affected by this unknown something. Therefore, the understanding, just by the fact that it accepts appearances, also admits to the existence of things in themselves, and to that extent we can say that the representation of such beings as underlie the appearances, hence of mere intelligible beings, is not merely permitted but is unavoidable." Academy edition: [4: 311-5]. Our edition: p. 68.

with regard to the things in themselves is seen within Kant's system as problematic, if not straightforwardly jettisoned.¹

The problem with this proposed reconstruction, in my view, is that what it leaves us with is not even the *kind* of doctrine Kant was after. In other words, the reconstruction dangerously allows as a possibility exactly what Kant was trying to deny. If according to the reconstruction the commitment to the existence of the things in themselves is merely a problematic one, then transcendental idealism thus reconstructed can only claim that it is merely possible that things in themselves exist. Hence it must be possible that things in themselves do not exist. If they do not exist, *a fortiori* they do not affect sensibility thereby grounding the phenomena. If this is so, subjectivity as described by Kant would not only provide the formal framework of experience (space, time, the categories, etc.), but subjectivity alone would be sufficient to explain the whole of experience, including the existence of material objects like chairs and trees, galaxies and atoms. This is the possibility that under the reconstruction in question transcendental idealism could not legitimately deny. Now, since the reconstruction explicitly admits its departure from Kant's text, those who reconstruct the doctrine in this way might gladly admit all these consequences of the reconstruction.² But it is perhaps better to strive for an interpretation of the doctrine which does not allow for this possibility of subjective self-sufficiency. It is not an accident that most of the well-known passages where Kant affirms the existence of the things in themselves, and where he appeals to them in order to account for the *material* aspects of experience, are passages that stem either from the *Prolegomena* or from the second edition of the first *Critique*, where Kant's main concern was to distance himself from what he called the "material idealism" of Berkeley.

The interpretative challenge, then, is to keep the claims that non-spatiotemporal things in themselves exist and that they affect sensibility, but to do this in an intelligible way. This challenge has been taken up by those who propose to "deontologize" transcendental idealism, i.e. to

¹ Cf. W.H. Walsh [1981] (1987) pp. 123-4: "Thus the familiar objection that Kant allows the existence of things-in-themselves but inconsistently adds that they are entirely unknown to us is certainly valid as it stands, but holds against a position Kant took up carelessly and inadvertently, one he might well have avoided. [...] Had Kant been content to work with the *idea* of an independent thing in itself, leaving open the question whether anything corresponds to it, he could have stood by all his main conclusions without being exposed to the charge of inconsistency." Along these lines is, I think, the reconstruction defended by Hoke Robinson (1994), especially p. 435: "On the current view, however, the thing in itself is only something I think [a problematic] God would 'see'."

² Indeed, transcendental idealism thus reconstructed provides an excellent explanation of why post-Kantian philosophy evolved the way it did. What would happen if one keeps Kant's "main conclusions" (say, the idea that the objects we know "must conform" to our cognitive capacities) but at the same time one dismisses the *problematic idea* of things existing independently of the mind? Perhaps the answer to this question is to be found in the writings of Fichte, Schelling and Hegel.

understand the distinction between appearances and things in themselves not as an ontological distinction between two kinds of entities, but as a distinction between two ways of considering the same objects: *as* phenomena and *as* things in themselves. As Henry E. Allison, a prominent defender of this interpretation, has put it, when we speak of things in themselves “the concern is rather with the familiar objects of human experience, *considered as they are in themselves* [...], where the *an sich selbst* functions adverbially to characterize *how* a thing is being considered rather than the kind of thing it is or the way in which it exists.”¹ The two problems mentioned above, that of the existence of things in themselves and of their affection of sensibility, are supposed to be solved by this interpretation because we would have as strong an existential commitment with regards to things in themselves as we have it with regards to phenomena. Consequently, as regards the problem of affection,

[...] no entities are assumed (in the account of affection) other than the spatiotemporal objects of human experience. The point is only that, insofar as these are to function in a transcendental account as material conditions of human cognition, they cannot, without contradiction, be taken under their empirical description.²

This interpretation has as much textual evidence in its support as the “harsh”, ontological interpretation of Kant’s distinction between phenomena and things in themselves,³ and if it can accommodate the claims of existence and affection of things in themselves we have every good reason to think that this is the kind of doctrine Kant was after, whether or not he was always successful in his formulations.

Under this interpretation, however, transcendental idealism becomes an unintelligible doctrine in perhaps the clearest possible way. Transcendental idealism, thus interpreted, by no means relinquishes the claim that things in themselves are *not* spatiotemporal. How are we, then, to understand the idea that we are talking about the *same* objects but that they are spatiotemporal under one consideration and non-spatiotemporal under another consideration? In what sense is this table in front of me *identical* with an object which is not spatiotemporal? Are we giving up the commitment to the principle of identity of indiscernibles? If the logical inconsistency was hard to

¹ Allison, H. E. (2004) pp. 51 and 52.

² Ibid p. 68.

³ The preface to the second edition has some of the passages that most strongly support this sympathetic reading. Crucially, see B XIX-XX. As should be clear from my references to Kant’s text in my characterization of the “harsh” interpretation, however, if not from my use of scary quotes, I do not believe that the “harsh” interpretation is quite the depiction of a straw man.

spot under the ontological reading of the doctrine, the *deontologizing* move makes the very formulation of transcendental idealism a patent absurdity.¹

I do not think that these are conclusive arguments against transcendental idealism. My reference to the different interpretations of the doctrine is everything but exhaustive. I hope, however, that these considerations are enough to make the reader understand why, in a section of this investigation which is far ahead (section 4.4), I express my extreme satisfaction at the fact that transcendental idealism is not enforced upon me by the existence of an antinomy more convincing than the one Kant tried to derive from the concept of the spatiotemporal world. The problems pointed out in this section are indeed the reasons why I am not a transcendental idealist.

¹ Cf. Guyer, P (1987) p. 334; Robinson, H. (1994) p. 421, and Van Cleve, J. (1999) pp. 146-50. Allison (1996) addresses the criticism as formulated by Robinson. He understands it as the problem of *how* are we to characterize *the object that appears* if it is to be considered in “such disparate manners” –one which attributes spatiotemporality to it and one which denies spatiotemporality to it-. Here is how Allison attempts to solve the problem: “the object is to be characterized as a “transcendental object = x”. In this respect, the *concept* of the transcendental object is not equivalent to the *concept* of the thing considered as it is in itself. [...] What this suggests, however, is not the rejection of any connection between the thing in itself and transcendental object, but rather the necessity of a further distinction between the thing in itself *simpliciter*, which for us (as finite discursive intellects) can be thought merely as a transcendental object = x, and the thing *considered* as it is in itself, which is thought through pure categories. The basic difference is that the former must be characterized as “= x” because it remains inaccessible to all the resources of a discursive intellect, while the latter, as involving independence merely from sensible conditions, can at least be thought problematically” (p. 16). As far as I can see, however, the problem is not solved at all. First, because the “problematicity” in question is of a very radical kind: how can we *think* (even problematically) of a thing in itself as being *identical* to an empirical object if the former is not spatiotemporal but the latter is? Second, and more importantly: if *what appears* (note the singular) is inaccessible to us (“= x”) and non-spatiotemporal, but the empirical objects are (note the plural) neither inaccessible nor non-spatiotemporal, in which sense is the distinction between the transcendental object and the empirical objects *not* a distinction between two kinds of objects, hence an ontological distinction?

Charles Parsons (1992) provides a different reply to this problem, a reply which he partly attributes to the German commentator Gerold Prauss. After stating the difficulty, Parsons responds this way: “The only possible reply to this objection is the one suggested by Prauss: When one considers *this chair* as it is in itself, “this chair” already refers to an empirical object. So that its consideration as an appearance is presupposed. So long as there is *some* distinction between empirical objects and representations, this way of understanding talk of things in themselves is available.” (p.90) It must be noted, however, that the only sense in which Parsons allowed to distinguish between objects and representations previously in that same paper (p. 85) was the sense of an empirical distinction, i.e. as a distinction that “can only be made in some way within the sphere of representations”. Hence, Parsons’ reading of the distinction does not seem to be the transcendental one that Kant seems to be interested in, at least when he asks how are *empirical objects* possible, and he replies that by our being affected by things that are *not* empirical objects. Now, if transcendental idealism should be interpreted in such a way that we should reject this latter talk about things beyond possible experience being the ground of things within experience, then the resulting version of transcendental idealism does not seem to me to be the kind of doctrine that Kant was after, and, more importantly, the non-spatiotemporality thesis would now be even harder to affirm: if there is one thing that is out of any serious discussion, it is that Kant thought that the objects of possible experience are, precisely, spatiotemporal objects.

Conclusions

We have answered with an unequivocal 'No' to the most important questions that we asked in this chapter. There is no such thing as a first antinomy of pure reason with regard to the spatiotemporal world. Even if the antinomy was genuine, transcendental idealism would not follow from it, so transcendental idealism was not the only way out of the alleged antinomy. Moreover, transcendental idealism was not even a plausible way out of the antinomy since it faces enormous problems in its own.

I hope, however, that in the course of the following chapters of this investigation, and especially in the fourth one, the reader comes to appreciate the First Antinomy as an important episode in the history of philosophical grappling with the concept of what is absolutely unconditioned. It was no doubt an erred attempt. But this is of little importance, since –to advance things a little- I don't think we are in a much better position more than two centuries after Kant.

Chapter II. A Collection of all Collections: A First Approach to its Problems

In our previous chapter we looked at Kant's First Antinomy, which, if genuine, would be the situation that we have to affirm that the spatiotemporal world is both finite (Thesis) and infinite (Antithesis). We saw that there is no genuine antinomy, since the purported proofs of both the Thesis and the Antithesis are open to criticism. The argument for the Thesis, when carefully examined, turns out to presuppose what it should conclude –viz. that an elapsed eternity, being a completed infinity, is impossible–, as Bertrand Russell pointed out. On the other hand, the argument for the Antithesis rests on the presupposition of a definition of “beginning” according to which anything that begins must be preceded by a time in which that thing did not exist. In the case of an absolute beginning of the world, that definition implied that the beginning of the world must be preceded by empty time. But this presupposition is not necessary, as Jonathan Bennett, taking up an argument from Leibniz, has so neatly exposed.

In this chapter I would like to move on to a series of antinomies which look more convincingly genuine. I mean the paradoxes which, under certain presuppositions, are generated in set theory by the concepts of overly comprehensive totalities. Not that these paradoxes are absolutely unavoidable by any other means than by rejecting the existence of the totalities in question, which is the standard approach to these problems. There are plausible attempts at avoiding the paradoxes by other means, as we will see in the last chapter of this investigation, when we refer to alternative set theories (like Quine's system “New Foundations”) and to non-set-theoretical attempts to conceptualize overly comprehensive totalities (like Grim's and Rescher's “Plenum Theory”). The adoption of these alternative theories, however, does not come without a price, a price we might not be willing to pay. And it seems to me that we are entitled to refuse to pay that price. If this is correct, the standard approach to these issues is worth looking at carefully, as is the investigation of its philosophical significance.

What is the connection between Kant's fake antinomy and the more compellingly genuine paradoxes of set theory? This will be the subject matter of our next chapter, but some preliminary hints can be advanced. At first glance, the connection seems to be quite feeble, not only because the paradoxes of set theory rest on stronger arguments than Kant's "proofs", but also because the paradoxes of set theory arise from a certain presupposition which seems to be forced upon us by *logic*¹; but Kant thought that logic, being a formal science, had secured its status as a science since the time of Aristotle, and this only at the cost of being unable to make any advance. Here, again, Kant was simply mistaken, as the development of logic(s) from the nineteenth century onwards bears witness. Add to these considerations the quite evident remark that the concepts which generate the two problems (Kant's *spatiotemporal world* and, say, the *set of all sets*) are by no means identical, and then you have enough excuse to forgo any deep investigations into the connection between the antinomies and the paradoxes. But I will argue that there is a relation between Kant's spatiotemporal world and the overly comprehensive collections which generated the problems of set theory. For Kant sometimes put the point in these terms: what generates the antinomies, and indeed all our metaphysical cul-de-sacs, is the idea of *the Unconditioned*. If this is correct, then the sentence: "Try to frame the idea of an absolutely comprehensive totality, i.e. of a totality which in terms of comprehension is absolutely unconditioned, and you'll get involved in contradictions", is both a legitimate paraphrase of Kant's ideas and a statement of what the paradoxes of set theory suggest (to some of us, anyway), i.e. that there is no such *definite* thing as Everything.

To elaborate further on these thoughts, however, would be to precipitate matters. Let us concentrate for the moment on the paradoxes of set theory and leave the question of the connection between the two problems for our next chapter.

I will proceed as follows. First, I will show how the concept of a set of all sets cannot be consistently upheld while at the same time retaining a theorem of set theory called Cantor's theorem. This is the so-called Cantor's Paradox. Secondly, I will show how the concept of a set of all sets generates a further paradox by leaving the door open to our framing the concept of a set of all those sets that do not belong to themselves; but this concept, as Bertrand Russell discovered, is self-

¹ I have in mind, of course, the presupposition that given any condition there is a collection containing exactly those objects that satisfy the condition. Now it has been forcefully defended that *Georg Cantor* himself, the founder of set theory, never worked under such a presupposition (See Hallett (1984), p. 38 and Lavine, S. (1994) pp. 55 and 77). The paradoxes that arise for the "logical conception of a collection" shouldn't therefore be characterized as paradoxes of *Cantorian* set theory; the logical conception, together with the paradoxes which it engenders, is attributable rather to Peano, Frege and Russell. This historical fact will play an important role in the later sections of this chapter, and indeed in subsequent chapters.

contradictory¹. This is Russell's Paradox. After presenting these paradoxes in a way which does not presuppose that the reader has any knowledge of set theory, I will move on to a philosophical discussion of them. I will present Russell's mature attitude to the paradoxes in a positive light. According to this attitude what the paradoxes suggest is that there is no such thing as everything. Inspired by this attitude, I will go on to suggest that it would be wrong to downplay the philosophical significance of the paradoxes, for example, by relying on the formal distinction between different kinds of collections (sets and classes), which are supposed to behave differently. Nothing is changed, by the introduction of this distinction alone, with regard to the problematical status of the concept of an absolutely all-comprehensive totality. Finally, I will discuss some issues related to the proposal that the universe of set theory should be considered as a potentially infinite totality.²

2.1. The Paradox of the Largest Cardinal (also called "Cantor's Paradox")

Once one understands the language in which it is couched, the so-called Cantor's Paradox is very simple, and it is the most straightforward way to say what we want to say. Cantor proved that given any set α , there is another set which has more members than α , namely the Power-set of α , symbolized $P\alpha$. It follows that if a set of all sets existed (we could refer to this Universal Set by 'U'), there would be a set with more members than U, namely PU. But this is a contradiction since the set of *all* sets is, by definition, the most inclusive of all sets. Our task in this section is to state this paradox as clearly as possible, without presupposing that the reader has any knowledge of set theory.

Let us refer to the set which consists of exactly the members 0, 1, 2 and 3 by the following string of symbols: {0, 1, 2, 3}. This set has four members. Again, if we want to refer to the set which contains exactly the members a, b and c, we refer to it by: {a, b, c}. This set has three members. We are not concerned with the order of the members in the set. Thus, our first set is the same as the set {1, 0, 2, 3} and our second set is the same set as {c, a, b}. If we want to say that a certain element belongs to a certain set, we use the symbol ' $\in$ '. Thus if we want to say that 1 belongs to our first set, we write: $1 \in \{0, 1, 2, 3\}$. Or, if we call our first set ' α ', we can say that 1 belongs to α by writing: $1 \in \alpha$. Again, if we call our second set ' β ' we can affirm truly that $c \in \beta$.

¹ By the self-contradictoriness of a concept I mean that a contradiction can be derived from the presupposition that there is an object that corresponds to this concept.

² The following two sections (2.1 and 2.2) are, as I said in the text, written without presupposing any knowledge whatsoever of set theory by the reader. Those readers who already know what these paradoxes consist in can skip the two sections without loss, and start reading from p. 58 onwards.

An important concept in set theory is the concept of a *subset* of a given set. A subset of a given set is also a set, namely a set which is said to be *included* in the original set because there is no member which belongs to the subset and which does not belong to the original set. For example, the sets $\{0\}$ and $\{0, 1\}$ are both subsets of our first set in the previous paragraph: $\{0, 1, 2, 3\}$. So are the sets $\{0, 1, 2\}$ and $\{0, 3\}$. What these subsets have in common, and what makes them all subsets of the set $\{0, 1, 2, 3\}$ is that any member which belongs to them also belongs to the original one. More formally, we can say that a set α is a subset of the set β if and only if for any object x which is a member of α , x is also a member of β . If we use the symbol ' $\subseteq$ ' to express that the set on the left-hand side of the symbol is a subset of the set on the right-hand side of the symbol, we can then define the subset-relation (sometimes also called the inclusion-relation) thus: $\alpha \subseteq \beta \leftrightarrow_{\text{def}} \forall x (x \in \alpha \rightarrow x \in \beta)$.¹ This is read, to repeat: α is a subset of β if and only if any member of α is also a member of β .

What happens if α has no members? That is to say, what happens if α is the so-called null-set, symbolized by ' $\emptyset$ '?² Well, in that case it is still true that any member of α is also a member of β , no matter what set we want to refer to by the symbol β . In classical deductive logic, the conditional $P \rightarrow Q$ is true even when P is false. Applying this rule, in set theory you say that when $\alpha = \emptyset$, the conditional inside the parenthesis of our definition is 'trivially' true: since $\emptyset$ by definition has no members, it can't be the case that something is a member of α (i.e. of $\emptyset$) without being a member of β . In other words, it is a consequence of the definition of a subset that the null-set is a subset of every set: $\forall x (\emptyset \subseteq x)$. Another important consequence of the definition of the subset relation, this time an obvious and immediate consequence (since any member of x is, of course, a member of x), is that every set is a subset of itself: $\forall x (x \subseteq x)$. These two facts, that the null-set is a subset of every set and that every set is a subset of itself, explain why Russell says that subsets are "ways of selecting [a set's] members (including the extreme cases where we select all or none)"³.

Those subsets of a set which are *not* identical with the set itself are called *proper subsets*, symbolized by ' $\subset$ '. Thus we can affirm truly that $\{0, 1\} \subset \{0, 1, 2\}$. And although it is also true that $\{0, 1\} \subseteq \{0, 1, 2, 3\}$, it is false that $\{0, 1, 2, 3\} \subset \{0, 1, 2, 3\}$. That is to say, every proper subset of x is also a subset of x , but not every subset of x is also a proper subset of x : x itself is a subset of x which is

¹ Cf. The definition of the "inclusion relation" in Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 26.

² There can't be more than one null-set, *the* null-set: in set theory there is an Axiom of Extensionality, which says that the identity of a set is determined by the set's members, so that two sets are identical if and only if they have the same members. Thus any other set which had the characteristic of having no members would be identical with the null-set. See theorem III in Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 39.

³ Russell, B. [1919], pp. 84-5.

not a proper subset of x . Think, to return to our second example above, of the set $\{a, b, c\}$. Following our definition, we can see that its subsets are exactly the following sets:

- $\emptyset$,
- $\{a\}$,
- $\{b\}$,
- $\{c\}$,
- $\{a, b\}$,
- $\{b, c\}$,
- $\{a, c\}$,
- $\{a, b, c\}$

These are, in effect, all the possible “ways of selecting the members” of the set $\{a, b, c\}$, “including the extreme cases where we select all or none”. All but the last of these subsets, namely $\{a, b, c\}$ itself, are the so-called proper subsets of the original set.

This also shows why one must clearly distinguish between the membership-relation and the subset-relation, i.e. between $\in$ and $\subseteq$. So far we have established quite clearly that every set is a *subset* of itself, but we haven’t said anything about a set being a *member* of itself –whatever that may mean. We are sure now, for example, that $\{a, b, c\}$ is a subset of itself. But it is blatantly false that $\{a, b, c\}$ is a member of itself. $\{a, b, c\}$ has only three members, namely a , b and c . Again, it is crucial to distinguish between an element x and a set whose only member is x . That is to say, $x \neq \{x\}$. It is true that $x \in \{x\}$, but it is false that $\{x\} \in \{x\}$ because $\{x\}$ only has one member and it is x , not $\{x\}$.¹

When we gather all the subsets of a given set into a new set, what is formed is called the Power-set of the original set. If our original set is again $\{a, b, c\}$, then its Power-set is the set in which we gather exactly all the subsets we just listed above. If we call β the original set and we denote the Power-set of β by ‘ $P\beta$ ’, then we have:

$$\beta = \{a, b, c\}$$
$$P\beta = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}.$$

Again, if we think of our first example:

¹ “In our terminology, while a set always *includes* [$\subseteq$] itself and its subsets, it *contains* [$\in$], in general, neither itself nor its subsets”. Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 27.

$$\alpha = \{0, 1, 2, 3\},$$

then the Power-set of α is:

$$P\alpha = \{\emptyset, \{0\}, \{1\}, \{2\}, \{3\}, \{0, 1\}, \{1, 2\}, \{2, 3\}, \{0, 2\}, \{1, 3\}, \{0, 3\}, \{0, 1, 2\}, \{1, 2, 3\}, \{0, 1, 3\}, \{0, 2, 3\}, \{0, 1, 2, 3\}\}.$$

Now, these basic examples suggest that the Power-set of a given set will always be larger than the original set. The Power-set of a given set seems to be of a greater size than the original set, i.e. it seems to have more members. If our original set has three members, we see that its Power-set has eight members. If our original set has four members, its Power-set has sixteen members. But is there a specific relation between the number of members of a given set (i.e. its *size*) and the number of members of its correlative Power-set?

In order to answer this question, first note that $8 = 2^3$, and that $16 = 2^4$. And if you collect together all the subsets of a set of 20 members (which you're free to do, provided that you have the time and patience necessary to do it), you'll find out that they are in total 2^{20} . Likewise, the Power-set of a set of 100 members has 2^{100} members. And in general, if a set has n members, its Power set has exactly 2^n members. (Hence the name: Power-set). This is no coincidence. Think of the fact that, for every member x of a set α , there are just two possibilities if you consider it in relation to any of the subsets of α , i.e. in relation to those sets y which constitute $P\alpha$: either x is in y or it is not, i.e. either $x \in y$ or $x \notin y$. So these two possibilities have to be realized as many times as there are members of the original set. If the set has five members, then the subset has $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^5 = 32$ members.¹

Given that 2^n is clearly a larger number than n , we could be easily tempted to say that we have already established what we wanted: namely, that the Power-set of any given set is always larger than the original set, because the Power-set has 2^n members whenever the original set has n members. This is actually true, but what we need is a *proof* of this fact. Moreover, our proof must be such that it conforms to the criterion of size comparison that is essential to set theory, namely that

¹ I have found the clearest explanation of this fact in Mary Tiles' (1989) p. 58: "If we have a collection of n things, then the total number of possible ways of selecting subcollections, of whatever size, from it will be 2^n because, for each of the n elements, we can either include it in our subcollection or not, and n such two-way choices determine a subcollection. So for two elements we get four possibilities:

a	0	0	1	1
b	0	1	0	1
	$\emptyset \{b\} \{a\} \{a,b\}$			

where $\emptyset$ is the empty-set, 0 means the element gets excluded and 1 means that it gets included." Cf. as well Courant, R. [1941] pp. 85-6 and Moore, A.W. (2001) p. 148.

criterion according to which two sets are the same size, i.e. have the same number of members, when there is a one-to-one correlation between the elements of one of the sets and the elements of the other.

This set-theoretical criterion of size comparison yields unproblematic results whenever we are dealing with finite numbers. Think of those cases in which we are able to enumerate all the elements of a given set, and thus state the exact number of elements. For example, if we want to establish the size of the set of objects in our drawer, we would usually enumerate one by one the elements of that set and thus find out that the number is, say, twenty; and if we want to find out whether another set is of the same size as the set of objects in our drawer, we would enumerate the elements of the second set, and thus find out whether it has twenty elements or not. *In principle*, this sort of size comparison, i.e., enumeration, could be made for any set with a finite number of members, although for physiological or technological reasons we might not be able to carry out the actual counting of elements of a set when their number is too big. And the result of this enumeration-criterion must coincide with the result of applying the correlation-criterion we mentioned above: indeed, an act of enumeration is (among other things) an act of correlation, since whenever we are counting a set of n elements we are correlating each of those elements with the set of natural numbers from 1 up to n . When a set has n number of members, we will say from now on that the *cardinal number* of this set is n , or that its *cardinality* is n , and when two sets have the same cardinality we will say that they are *equinumerous*.¹

Now, it is important to realize that we use the correlation-criterion even in cases when we do not know the exact number of elements in a set. For example, there are cases in which we can say that the number of knives on a very big table is equal to the number of forks on it: if the table is elegantly disposed, but we are somehow unable to carry out the operations which would let us know the exact number of forks and knives on the table (enumeration, multiplication, etc.), we might still be able to recognize at a glance that to every knife on the table there corresponds a fork and vice versa. Thus we know that the cardinal number of these sets is the same, whatever that is. Again, in a monogamous society in which there are no widows nor widowers, we immediately know that the number of wives is the same as the number of husbands, because we know that to every husband there is a corresponding wife and vice versa.

¹ Cf. Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 44.

Now, although what I've been saying looks fairly unproblematic and even intuitive, the correlation-criterion of size comparison famously led Cantor to *prima facie* puzzling results when he dealt with *infinite* sets instead of finite ones. Think of the set of natural numbers:

$$\{0, 1, 2, 3, 4, \dots\}$$

And now think of the set of even numbers:

$$\{0, 2, 4, 6, 8, \dots\}$$

Although the set of natural numbers contains all the elements that are in the set of even numbers *and infinitely many more elements* as well (namely all the uneven numbers), there is a one-to-one correlation between the two sets: to every natural number there is a corresponding even number and vice versa. According to the correlation-criterion, then, the set of even numbers has as many elements as the set of natural numbers. These sets are the same size. If you think that this is puzzling, think now of the set of multiples of one million. By the same argument as above, this set is also equinumerous with the set of natural numbers. Thus all these three sets are equinumerous: the set of natural numbers, the set of even numbers, and the set of multiples of one million.

We don't need to discuss the issue of the intuitiveness of these results. Those who developed set theory knew very well that the results of this new discipline required us to question some time-honored beliefs, like the belief that any whole must be greater than its parts ("*Totum parte majus*").¹ They challenged the generality of this rule; in particular, they thought that although it is a valid rule in the case of finite sets, its application is no longer valid in the case of infinite sets. An infinite set is, in some sense, a whole which is *not* greater than some of its parts: it is equinumerous with some of its proper subsets. Admittedly, if the only acceptable criterion of size comparison were the so-called 'subset criterion', i.e. the one according to which set α is larger than set β when all of β 's members are in α but not all of α 's members are in β , then it would follow that the set of all natural numbers must be larger than the set of even numbers, and that this in turn must be larger than the set of all multiples of one million. But the subset criterion is not the one that is used in set theory. The criterion used in set theory is, as we said, the correlation criterion; and this is in principle a plausible criterion. That it immediately leads to counterintuitive results does not necessarily mean that it leads to contradictions.

What is important for our present purposes is that, even when we use the correlation-criterion of size comparison, there is a *proof* of the fact that, given any set α , there is a set with larger cardinality

¹ Cf. Deiser, O. (2010) p. 92.

than α , namely $\mathbf{P}\alpha$. Since there is a proof of this fact, the proposition which expresses this fact is called a *theorem*, and since the proof was found by Cantor himself, the theorem is called *Cantor's theorem*.

The proof has to establish not only that $\mathbf{P}\alpha$ has *at least as many members* as α (i.e. that if it is not equal, it is larger, which would be symbolized: $|\alpha| \leq |\mathbf{P}\alpha|$). This is actually easy to prove, since we can correlate every element of α with the so-called "singleton" which contains exactly that element of α . For example, if α is $\{0, 1, 2, 3\}$, we have seen that its Power-set has, amongst other members, those singletons which are formed by taking each element of α at a time; thus we have that $\mathbf{P}\alpha$ has amongst its members the sets: $\{0\}, \{1\}, \{2\}, \{3\}$. But then we can correlate one-to-one the elements of α with those singletons. Since we are sure that $\mathbf{P}\alpha$, by its very definition, contains at least these singletons, we can be sure that $\mathbf{P}\alpha$ contains at least as many members as α . Thus $|\alpha| \leq |\mathbf{P}\alpha|$. But, as I said, what has to be proved is that $\mathbf{P}\alpha$ is *strictly larger* than α . That is to say, we have to prove that although it *is* the case that $|\alpha| \leq |\mathbf{P}\alpha|$, it is *not* the case that $|\alpha| = |\mathbf{P}\alpha|$. Only this will establish that $\mathbf{P}\alpha$ has a strictly greater cardinality than α , i.e. $|\alpha| < |\mathbf{P}\alpha|$. The proof of this fact is by contradiction. Suppose that α and $\mathbf{P}\alpha$ are equinumerous ($|\alpha| = |\mathbf{P}\alpha|$) . Then there is a one-to-one correlation between the members of α and the members of $\mathbf{P}\alpha$. So every member x of α is correlated with a member y of $\mathbf{P}\alpha$ and vice versa. Call this hypothetical correlation ' f '. It will look something like the following:

Members of α	Members of $\mathbf{P}\alpha$ (i.e. subsets of α)
X_1	$y_1 = f(x_1)$
X_2	$y_2 = f(x_2)$
X_3	$y_3 = f(x_3)$
.	.
.	.
.	.
X_n	$y_n = f(x_n)$
.	.
.	.
.	.

Now, since the members of $\mathbf{P}\alpha$ are subsets of α , and since a subset of α is either α itself, or the null-set $\emptyset$, or a set which has some but not all the members of α , then it is clear that for every member x of α there are exactly two possibilities: either x is contained in the element y of $\mathbf{P}\alpha$ which f assigns to it, or it is *not* contained in that element.¹ Consider now the set of those elements of α which are *not* contained in the element of $\mathbf{P}\alpha$ with which f correlates them. Call this set ' β '. The set β must be a genuine subset of α since, if it has any members at all, they are elements of α ; and if β has no members, then β is simply the null-set, which again is a genuine subset of α . So β is a subset of α (i.e. $\beta \subseteq \alpha$). But then β must be contained in $\mathbf{P}\alpha$ (i.e. $\beta \in \mathbf{P}\alpha$), because by definition $\mathbf{P}\alpha$ contains all and only the subsets of α . But since *ex hypothesi* there is a one-to-one correlation f between the members of α and the members of $\mathbf{P}\alpha$, there must be a member of α , call it ' w ', which is correlated with β by f . That is to say, $\beta = f(w)$. But now we can ask: Is w a member of β or not? Given that f correlates w with β , then w *can't* be a member of β , because β is so defined that it contains exactly those elements of α which are not contained in the set which f assigns them. So $w \notin \beta$. Now if w is *not* an element of β , this means that w is not contained in the element of $\mathbf{P}\alpha$ which f assigns to it. But then, by the definition of β , w *must* be in β because β contains exactly those elements of α which are not members of the set which f assigns to them. Thus w is in β , i.e. $w \in \beta$. We have arrived at a contradiction: w is and is not in β . Under the presupposition that the number of elements in α is the same as the number of elements in $\mathbf{P}\alpha$, we get that $w \in \beta \wedge w \notin \beta$. The presupposition must therefore be false. And since we previously established also that the number of elements in $\mathbf{P}\alpha$ is *at least as big* as that the number of elements in α , we have already established our two desired results: $|\alpha| \leq |\mathbf{P}\alpha|$ and it is *not* the case that $|\alpha| = |\mathbf{P}\alpha|$. This means that $|\alpha| < |\mathbf{P}\alpha|$. In other words, no matter how many members there are in a given set α : its Power-set, $\mathbf{P}\alpha$, contains more members. This is Cantor's theorem.²

Now we really have all the necessary elements to understand the next statement of Cantor's paradox:

According to Cantor's theorem, the set $\mathbf{P}s$ of all the subsets of any given set s has a greater cardinal than has s itself. Consider now the set of *all* sets, call it U . Its "power-set", $\mathbf{P}U$, i.e. the set of all subsets of U ,

¹ If we return to our example, where α is $\{0, 1, 2, 3\}$, then f could correlate 1, for example, with $\{0, 1\}$, in which case f is correlating an element of α with a member of $\mathbf{P}\alpha$ which *does* contain that very element; but it could happen that f correlates 1 with $\{0, 2, 3\}$, or with the null-set $\emptyset$, or with $\{2\}$, and in any of these cases f would be correlating an element of α with a member of $\mathbf{P}\alpha$ which *does not* contain that very element.

² Cf. Russell, B. [1919] pp. 85-6, Courant, R. [1941] pp. 84-5, and Priest, G. (2002) p. 119.

has then a greater cardinal than U itself, which is paradoxical in view of the fact that U by definition is the most inclusive of all sets.¹

We have, therefore, established what we wanted to establish in this section. Acceptance of a set of all sets is incompatible with Cantor's Theorem. One of them must go. Or else we must abandon the claim that given any set α there *exists* its Power-set, $\mathbf{P}\alpha$. We would then be allowed to look with indifference at Cantor's result that $\mathbf{P}\alpha$ is always strictly larger than α . For nothing would force us to affirm the existence of $\mathbf{P}\alpha$ in the first place. Unfortunately, this way out is not available in set theory. The existential assumption about a $\mathbf{P}\alpha$ for any α is an *axiom* of set theory, and such a fundamental one to understanding the nature and behavior of the real numbers (and the continuum) that it is hard to imagine that such an axiom could be abandoned. Thus the Power-set axiom won't go. Cantor's Theorem won't go either, because it is a theorem and we just saw its proof. It is the set of all sets which must go, it seems.

Now, the concept of the set of all sets is not only paradoxical on account of Cantor's theorem. Let us look at Russell's paradox in order to see that the same conclusion, namely that we must give up the concept of a set of all sets, can be arrived at by other means.

2.2. Russell's Paradox

Probably the best way to understand Russell's paradox is to start by making reference to the so-called 'naïve set theory'. Naïve set theory is just set theory when it is practiced under the assumption of the *Axiom of Comprehension*.² But what is this Axiom of Comprehension?

Given its foundational purpose, set theory must work with a concept of *set* which allows sets to be *as arbitrary as possible*. We might expect that this desideratum of generality should be expressed in a rigorous definition of the concept of *set*. In fact, Cantor provided only a few characterizations of

¹ Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 7. I slightly modified their notation, since the authors symbolize the Power-set of s as $\mathbf{C}s$.

² Actually, as Boolos (1971) p. 15 points out, naïve set theory is characterized by the acceptance of the Axiom of Comprehension plus the Axiom of Extensionality which we mentioned in footnote 4.

this concept throughout his life¹; the one that came closest to expressing this desideratum of generality is probably the next one, given in 1895:

By a 'set' we understand every collection to a whole M of definite, well-differentiated objects m of our intuition or our thought. (We call these objects the 'elements' of M).²

Characterizations of this sort might lead us to think that the concept of set is such that, given any conceivable condition whatsoever, there must be a set which contains all and only those objects which satisfy that condition. This was actually Frege's and Russell's view. The Axiom of Comprehension expresses exactly this:

For any condition $\phi(x)$ on x there exists a set which contains exactly those elements x which fulfill this condition.³

According to this axiom, if we consider the condition of *being red*, we are allowed to assert that there is a set which contains exactly those objects which are red. This set is symbolized as follows:

$\{x \mid x \text{ is red}\}$

This is read: "The set of objects x such that x is red".⁴

When we consider such conditions as *being red*, there seems to be no reason why there wouldn't be a set which contains exactly those objects which are red. But even when the arbitrary condition under consideration is such that no object could possibly fulfill it, the Axiom of Comprehension implies that there is a corresponding set whose elements are exactly the objects which fulfill that condition. In those cases, the set in question is of course the null-set, $\emptyset$. Think, for example, of the condition of *not being identical with itself*. The Axiom of Comprehension implies the existence of the set: $\{x \mid x \neq x\}$, i.e. the set of those objects x such that x is not identical with itself. This set is the null-set (i.e. $\{x \mid x \neq x\} = \emptyset$).

So far, everything seems to be good with the Axiom of Comprehension. But think now of the condition that would be expressed by the formula: $x \in x$. That is to say, think of the condition of *belonging to itself* or of *being a member of itself*. There seems to be no problem with this condition. With regard to many things, in particular to many sets, we seem to be allowed to say that they

¹ See Hallett, M. (1984) pp. 33-4 and 45.

² Quoted by Hallett, M. (1984) p. 33. Also quoted by Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 15, and by Moore (2001) p. 9.

³ Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 31. In symbols, the Axiom is the next one: $\forall z_1, \dots, \forall z_n \exists y \forall x (x \in y \leftrightarrow \phi(x))$, where $z_1, \dots, z_n$ are the free variables of $\phi(x)$ other than x , and y is not a free variable of $\phi(x)$.

⁴ See Ebbinghaus, H-D. (2003) p. 3: "A transition, by collecting-into-a-whole, from a property E to the set $\{x \mid E \text{ is true of } x\}$ is called a *Komprehension*."

belong to themselves; and with regard to certain other sets, we seem to be allowed to say that they *do not* belong to themselves. For example, the set of planets is not itself a planet, and therefore does not belong to itself. Again, the set of human beings is not itself a human being, and therefore does not belong to itself. On the other hand, the set of things mentioned in the present chapter *is* a member of itself: I just mentioned it, and it therefore belongs to the set of things mentioned in this chapter. Another example of a set which *does* belong to itself is the set S defined as follows: “ S contains as elements all sets definable by an English phrase of less than twenty words”.¹ Thus, it seems that both ‘ $x \in x$ ’ and its negation, i.e. ‘ $x \notin x$ ’, are legitimate expressions of our language, i.e. expressions which we are allowed to use in order to define conditions.

But then we seem to be committed to the next statements:

- 1) The proposition ‘ $x \notin x$ ’ expresses a condition in our language.
- 2) Given any conceivable condition whatsoever, there must be a set which contains all and only those objects which satisfy that condition (Axiom of Comprehension).

These two statements imply the next existential statement:

- 3) There is a set which contains exactly those sets which are not members of themselves.

That is to say, we are forced to affirm the existence of the set: $\{x \mid x \notin x\}$. But this set, which is sometimes called the Russell-set R , is a paradoxical set. To see this, consider whether R belongs to itself or not. If it belongs to itself, then it is a member of that set whose members do not belong to themselves, so R does not belong to itself. That is to say, if R belongs to R then it does not belong to R , i.e. $R \in R \rightarrow R \notin R$. Conversely, suppose that R does not belong to itself. Since it does not belong to itself, it must therefore belong to that set which, by definition, contains exactly those sets which do not belong to themselves, which is no other set than R . Thus, $R \notin R \rightarrow R \in R$. This is the paradox. We asked whether R belongs to itself or not, and it turned out that if it does it doesn’t and if it doesn’t it does. ($R \in R \leftrightarrow R \notin R$).²

Since we have already dealt with Cantor’s theorem, we are in a position to see this Russellian paradox from a different perspective. Recall that in order to prove Cantor’s theorem we arrived at a contradiction when we presupposed the existence of a one-to-one correlation f between a set α and its Power-set $\mathbf{P}\alpha$. The contradiction was arrived at when we considered the subset of α called ‘ β ’,

¹ This example is taken from Courant, R. [1941] p. 87. Other examples are given by Tiles (1989) p. 116: “For example the set of all sets containing more than two elements must belong to itself, whereas the set of all sets containing less than two elements does not”.

² See Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) pp. 5-7, Courant, R. [1941] p. 87, Boolos, G. (1971) p. 15, and Moore (2001) p. 115.

defined by the condition that x belongs to β if and only if x is *not* contained in the subset of α to which f correlates it, i.e. $\beta = \{x \mid x \in \alpha \wedge x \notin f(x)\}$. We saw that this subset cannot be correlated with any element w in α , and this result contradicted the presupposition that there *was* such a one-to-one correlation. Now, not only Cantor's contradiction, but Russell's contradiction as well, can be derived by considering a set like β . To see this, think of f as the *identity* relation, i.e. the function that correlates every object with *itself*. The set $\beta = \{x \mid (x \in \alpha) \wedge (x \notin f(x))\}$, becomes now $\beta' = \{x \mid (x \in \alpha) \wedge (x \notin x)\}$. In order to get the Russellian set R from β' you just jettison the condition that β' must be a subset of α : Russell's set is the set of *all* sets that do not belong to themselves, and thus its elements must not also be elements of any previously obtained set α . Thus, we delete the first conjunct in the specification of the condition of β' (i.e., we delete $x \in \alpha$). What we get, then, is Russell's set R : $\{x \mid x \notin x\}$. And, as we just saw in the previous paragraph, this set is paradoxical because if it belongs to itself it doesn't and if it doesn't it does.¹

But what does Russell's Paradox have to do with our present concern? It has *a lot* to do with it. Two of the ways to avoid the Russellian paradox imply that there is no set of all sets. One of these ways is to reject statement 1) above. This was Russell's own solution. For him, it didn't make sense to say that something belongs to itself, nor of course that something fails to belong to itself. It follows that there is no set of all sets, for if such a set existed, it would have to contain itself as a member (since it is a set!). That is to say, if one accepts that something can contain itself as a member, one must accept as well that a set can fail to contain itself as a member; but as long as one retains statement 2) above (Axiom of Comprehension), Russell's paradox arises. The Russellian solution is, then, to refuse to countenance as meaningful any expression to the effect that a set is a member of itself. Since we would have to be committed to such an expression if we accepted the existence of a set of all sets, such an existential commitment is abandoned. We will come back to Russell in the next section of this chapter, but for the moment let us look at a second way of escaping this paradox.

The second way to avoid Russell's paradox consists in modifying statement 2) above in such a way that we are not allowed to derive from it the existence of the set R . But recall that the motivation of the Axiom of Comprehension was an important desideratum of set theory: to allow sets to be as *arbitrary as possible*. One needs therefore a harmless version of the Axiom of Comprehension: an axiom which does not imply contradictions but which at the same time allows set theory to keep playing its foundational role. Ernst Zermelo, who had independently discovered "Russell's" paradox

¹ Cf. Priest, G. (2002) pp. 128-9.

around the same time as Russell¹, found precisely this required axiom. It is called the *Axiom of Subsets*:

For any set a and any condition $\phi(x)$ on x there exists the set that contains just those members x of a which fulfill the condition $\phi(x)$.²

That is to say: *Given that set a exists*, we can then affirm the existence of a subset of a which is the set of all objects which fulfill any (arbitrary) condition. This is just like the Axiom of Comprehension, but without its bad consequences. It is like the Axiom of Comprehension because, as we can see, the condition by which the subset of a will be defined can be as arbitrary as we want. Thus you can certainly think of the condition that x does not belong to itself, $x \notin x$; and our new axiom makes sure that there is a *subset of a* which contains all those objects which do not belong to themselves. But what our new axiom certainly *does not* allow you to infer is the existence of a set-which-is-not-a-subset-of- a which contains *all* objects x that do not belong to themselves: given the definition of *subset*, the subset of a must contain, if it contains any members at all, only those elements which are also members of a . Therefore, when we use the new axiom, nothing at all is implied about the existence of those sets which are not subsets of a . Nothing forces us, therefore, to affirm the existence of R , the Russellian set: $\{x \mid x \notin x\}$. All that we are allowed to infer, given that we have a , is the existence of the set $\{x \mid (x \in a) \wedge (x \notin x)\}$.

But now, again, what does the adoption of the Axiom of Subsets have to do with our present concern, namely the set of all sets? Well, again: *a lot*. And this is very simple. Suppose that there is a universal set, a set which contains all sets. Call this set ' U '. Given that we have the Axiom of Subsets, we are forced to affirm the existence of a set of all those subsets of U which fulfill any arbitrarily chosen condition. In particular, there must be a set of those subsets of U which are not members of themselves: $\{x \mid (x \in U) \wedge (x \notin x)\}$. But this set is no other than R , i.e. the Russellian set (in the previous symbolizations of this set, $\{x \mid x \notin x\}$, the condition that $x \in U$ was omitted, but if the set U exists, that condition is implied by the very definition of U). Thus, given the Axiom of Subsets, the existence of the paradoxical Russellian set is implied by the presupposition that there is a set of all sets. This is the reason why in modern set theory, in which there is an Axiom of Subsets, there is also a theorem according to which there is no set of all sets.³

¹ Cf. Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 5.
² Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 36. In symbols: $\forall z_1, \dots, \forall z_n \quad \forall a \exists y \forall x ((x \in y) \leftrightarrow (x \in a) \wedge (\phi(x)))$, where $z_1, \dots, z_n$ are the free variables of $\phi(x)$ other than x , and y is not a free variable of $\phi(x)$.
³ See Theorem 6 in Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) pp. 40-1. Cf. also Michael Potter's (2004) p. 43.

2.3 How to React in Case of Paradox

What did Russell take these paradoxes to show? He expresses his view in the next passage of *Introduction to Mathematical Philosophy*. It is a long passage, but I think it deserves to be quoted at large because it summarizes very well all that we have been discussing so far, connecting Cantor's and Russell's own paradox, while at the same time directing our attention to the discussion we are about to get involved in:

The class consisting of all objects that can be counted, of whatever sort, must, if there be such a class, have a cardinal number which is the greatest possible. Since all its sub-classes will be members of it, there cannot be more of them than there are members. Hence we arrive at a contradiction [given Cantor's theorem].

When I first came upon this contradiction, in the year 1901, I attempted to discover some flaw in Cantor's proof that there is no greatest cardinal [...]. Applying this proof to the supposed class of all imaginable objects, I was led to a new and simpler contradiction, namely the following: -

The comprehensive class we are considering, which is to embrace everything, must embrace itself as one of its members. In other words, if there is such a thing as "everything," then "everything" is something, and is a member of the class "everything". But normally a class is not a member of itself. Mankind, for example, is not a man. Form now the assemblage of all classes which are not members of themselves. This is a class: is it a member of itself or not? If it is, it is one of those classes that are not members of themselves, *i.e.* it is not a member of itself. If it is not, it is not one of those classes that are not members of themselves, *i.e.* it is a member of itself. Thus the two hypotheses – that it is, and that it is not, a member of itself- each implies its contradictory. This is a contradiction.¹

In other words, the concept of the set of all sets is just the concept of a collection which we might refer to with the word "everything" or with the phrase "the class of all imaginable objects". The concept of the set of all sets is philosophically significant because it is, *prima facie* at the very least, the concept of an *absolutely comprehensive totality*. The paradoxes seem to show that there is no such thing.

It is at this point, of course, that one might be tempted to bring to bear that distinction between sets and classes which was introduced, amongst others, by John von Neumann and Paul Bernays. There are important differences between those systems of set theory which admit classes. For example, in some systems (including Paul Bernays') the distinction is traced between two kinds of

¹ Russell, B. [1919] p. 136.

entities, sets and classes, stipulating that to every set there corresponds a class but not vice versa (the paradoxical classes do not have a corresponding set); in von Neumann's system, on the other hand, there are only classes, some of which (which you can call "sets") behave exactly like the sets of ZF while others (viz. "proper classes") do not behave like them because they are not membership-eligible. Let us look at the latter proposal first.

Von Neumann's system of set theory was developed according to the guiding idea that the problematical aspect of the paradoxical sets is not so much their size but their 'elementhood': that is to say, the problem with them is not that they are overly comprehensive, but that they can be elements of other collections. So von Neumann's system makes a distinction between those collections or classes which *can* be elements of other collections and those that *cannot* be elements of other collections. All collections in von Neumann's system are called 'classes'; but those classes which can be members of other classes are called 'sets' and these behave exactly like the sets of the Zermelo-Fraenkel set theory we have been considering so far; on the other hand, those collections in von Neumann's system which *cannot* be members of other collections are called 'proper classes'. In particular, in von Neumann's system, although there is no set of all sets, there *is* a class of all sets. It is called 'V'. And given that this class V cannot be a member of any other collection, it cannot, in particular, be a member of V itself, so Russell's paradox is avoided. Cantor's paradox is avoided because the collection which includes all the subsets of V is again a proper class, not a set; but Cantor's theorem only applies to sets, not to proper classes. Thus the class of all sets seems to be a legitimate collection which does not generate the paradoxes created by the set of all sets.¹

One important aspect of von Neumann's system is that it can be regarded as an elaboration of Cantor's own ideas.² Cantor was never committed to the view that every kind of collection or totality is a set. So he didn't really accept something like an *unrestricted* axiom of comprehension. Accordingly, his conception of a set was not the so-called "logical conception of a collection"³. For him, a *set* is something which is mathematically accessible and manageable: a number can be assigned to it, and it is increasable. More importantly, a set is a collection which can be *enumerated*:

¹ For an exposition of this extension of set theory into a set-theory-with-classes, see Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) pp. 119-53, on which my account in this paragraph heavily relies. It must also be noted that von Neumann spoke primarily of functions and only derivatively of classes. But this issue, as noted by Priest (2002) p. 163, is not so important in our present discussion.

² This has been forcefully argued for by Michael Hallett, to whose (1984) the next two paragraphs are highly indebted.

³ This is the conception to which I referred in note 1 of this chapter. Lavine's (1994) is another important source for the ideas contained in this paragraph.

Cantor's conception of a collection has been called the "combinatorial" conception of a collection.¹ He certainly thought that there were infinite sets which, as sets, were accessible to mathematical investigation, increasable and *enumerable*. Not by us, of course, but by an idealized subject whose powers of enumeration are not as limited as ours. But the fact that he believed in the existence of infinite sets doesn't imply that every kind of collection was a set in his view. There were, according to him, two kinds of totalities: on the one hand, the *finite* and the *transfinite* (i.e. infinite but still mathematically manageable) totalities, both of them being increasable and accessible to mathematical investigation; on the other hand, the second kind of totality was constituted by what is 'truly' or 'absolutely' infinite, i.e. by what is inaccessible, unincreasable, unbounded, etc. The Cantorian *doctrine of the Absolute* apparently underwent certain changes throughout Cantor's life; in particular, while at the beginning of his career he spoke of *the Absolute* as that which is in every respect beyond mathematical treatment (he wrote that 'The Absolute can only be acknowledged and admitted, never known, not even approximately'²); at a later stage, though, and as a response to the claim that his theory engendered paradoxes, he seems to have shifted to a doctrine of *absolute totalities*, which, although again were to be regarded as unincreasable, they were nonetheless to be treated as a *standard* of what is beyond mathematical manageability (and in this respect they were assigned a mathematical role, being now devoid of the mystical aura which surrounded *the Absolute* at the beginning of Cantor's career). Quite apart from this change, however, the main thing is that Cantor didn't see the so-called "paradoxes of set-theory" as paradoxes that arose for *his own* theory. In his view, those paradoxes arise only if one treats *absolutely infinite* totalities as sets, i.e. as mathematically accessible and manageable totalities; but this was precisely what his theory did not do. In 1899 he wrote in a letter to Richard Dedekind:

If we start from the notion of a definite multiplicity (a system, a totality) of things, it is necessary, as I have discovered, to distinguish between two kinds of multiplicity (by this I always mean *definite* multiplicities).

For a multiplicity can be such that the assumption that *all* of its elements 'are together' leads to a contradiction, so that it is impossible to conceive of the multiplicity as a unity, as 'one finished thing'. Such multiplicities I call *absolutely infinite* or *inconsistent multiplicities*.

As we can readily see, the 'totality of everything thinkable', for example, is such a multiplicity; later still other examples will turn up.

If on the other hand the totality of elements of a multiplicity can be thought of without contradiction as 'being together', so that they can

¹ Lavine (1994) p. 77.

² Quoted by Hallett (1984), p. 13.

be gathered together into '*one* thing', I call it a *consistent multiplicity* or a "*set*". (In French and in Italian this notion is aptly expressed by the words "*ensemble*" and "*insieme*").¹

As we can see, '*collection*', for Cantor, is not the same as '*set*'. All sets are collections but not all collections are sets: there are absolutely infinite collections which are not sets. And this certainly looks like von Neumann's distinction between sets and proper classes. Moreover, Cantor himself insisted upon the non-elementhood aspect of inconsistent totalities, i.e. on the fact that they cannot be members of anything:

[...] an inconsistent multiplicity, because it cannot be understood as a *whole*, and thus as a *thing*, cannot be used as an *element* of a multiplicity.

Only *complete things* can be taken as *elements* of a multiplicity, only *sets*, but not *inconsistent multiplicities*, in whose nature it lies, that they can never be conceived as *complete* and *actually existing*.²

Thus, it is by their failure to constitute unities that inconsistent collections cannot be members of any other collection. And this seems to make perfect sense: only "*unities*" can be iterated. If something is not a unity (a "*one*"), it cannot be iterated, which is to say that it cannot be a member of a collection.

How should we evaluate this attitude towards the paradoxes? Undoubtedly, it should be conceded that the distinction between two kinds of collections allows us to form the concept of a collection of all sets. In Cantor's case, we are allowed to form the concept of the multiplicity of all sets. In von Neumann's case, we are allowed to form the concept of the (proper) class of all sets. But this distinction can only constitute a solution to the paradoxes if it is shown that the paradoxes cannot be restated once the distinction has been introduced. Both Cantor and von Neumann take up the challenge when they say that the derivation of the paradoxes is blocked by the fact that proper classes (inconsistent multiplicities) cannot be members of any collection at all. The obvious question now is a very simple one: Why not? Why can't our newly introduced collections be iterated? Cantor went as far as taking up *this* challenge and responded: because they aren't unities or "*ones*". But does this really work? Think of a paradigmatic example of an inconsistent multiplicity or proper class:

¹ "Letter to Dedekind" in van Heijenoort, J (1967) p. 114.

² "Letter from Cantor to Jourdain, 9 July 1904", a translation of which appears as Appendix A in Lavine's (1994), pp. 98-9. Cf. Hallett (1984) p. 172 : "The analogue of what Cantor calls an '*absolute or inconsistent collection*' is a '*proper class*'."

the collection of all ordinals.¹ Is it a unity or not? Well, we are talking about *it*, and we are assuming that there are criteria for *its* identification. It is not the same thing as any of its elements and it is not to be identified with any other multiplicity, consistent or inconsistent. For example, the collection of ordinal numbers must be differentiated not only from the collection of natural numbers (which is a set in Cantor's view), but also from "the totality of everything thinkable" (which is not a set in Cantor's view, but an inconsistent multiplicity). *Pace* Cantor, then, inconsistent multiplicities are as much a "unity" as consistent ones.² This means that no reason has been given yet to the effect that inconsistent multiplicities or proper classes cannot be iterated. And this, in its turn, means that the distinction between different kinds of collections, by itself, is not enough to block the derivation of the paradoxes.³ Either there is a collection of all collections or there is not. If there is, it is one thing, and thus it must be a member of itself. But then we are back to square one, for in order to block Russell's paradox our axiom of Comprehension-for-collections must be such that it does not allow the derivation of the problematical collections. Since von Neumann's Comprehension is Zermelo's Separation axiom, we can still derive the existence of a Russellian collection from the collection of all collections.⁴

Similar considerations apply to those systems of set-theory-with-classes for which sets and classes are completely different sorts of entities.⁵ Those who propose these systems would insist that sets and classes are entities which respond to two different notions: the concept of *class* is a logical one, whereas the concept of *set* is a combinatorial (mathematical) one. Classes are not therefore subject to the combinatorial operations to which sets are subject. And while some of these systems allow *some* classes to be members of other classes⁶, the common denominator is that the would-be-paradoxical classes are not membership-eligible. The class of all sets can't be a member of any other collection. The obvious question is now the same as before: Why not? Note that the suggestion that

¹ This collection is paradoxical on account of what is now called the Burali-Forti Paradox. We haven't discussed this paradox, but it is in essence the next one: Ordinal numbers are generated by the general principle that, given *any* set of ordinals, there is one ordinal which succeeds them all. Think now of the set of all ordinals, Ω . The principle for the generation of ordinals forces us to say that there is an ordinal which succeeds Ω . But this is impossible, since Ω by definition contains *all* ordinals. Contradiction. Cantor was aware of this paradox and his solution was, of course, to deny that Ω is a set, or a "finished thing". Cf. Moore's (2001) p. 126. I will present this paradox in further detail in the following chapter, pp. 108-110.

² Alex Oliver (1998) criticizes the identification of Cantor's inconsistent multiplicities with von Neumann's proper classes because, he says (p.43), "a proper class is as much a 'unity' or 'finished thing' as a set.". But, as we have just seen, Cantor is not better off.

³ This predicament has been noted by Moore, A.W. (2001) pp. 150-1 and by Priest, G. (2002) pp. 164-5.

⁴ There are systems which modify the Comprehension axiom in such a way that the derivation of the Russellian set is effectively blocked. We will deal with one of these systems, namely Quine's NF, in our fourth chapter. In the present chapter I am confining myself, as I said in the introduction, to ZF and related systems.

⁵ See Parson's, C (1974), Maddy, P (1983) and Potter, M (2004) pp. 299-316.

⁶ This is Maddy's proposed system.

some classes are not unities would sound even more arbitrary as before. In effect, in these systems to every set there corresponds a class but not vice versa. This means that every set is *uniquely* represented by a class. The set of cats is represented by the class of cats, not by the class of dogs. Clearly, then, the non-problematical classes are unities or “ones”. Why aren’t the problematical classes unities as well, then? We are, after all, talking and thinking about them, and we want to distinguish them from one another as well as from their members. To say that they are not unities would be arbitrary, if not nonsensical.¹ And this again just means that we haven’t been provided with the means to block the derivation paradoxes.

2.4 Potential Universe, Indefinite Being.

A more sympathetic approach to Cantor’s distinction between consistent and inconsistent multiplicities, as well as to the distinction between sets and (proper) classes, would take the following consideration as its starting point: the reason why the problematical collections cannot be iterated is because they are always in the process of “becoming”. This suggestion would echo Cantor’s idea that an inconsistent multiplicity can’t be thought of as “one finished thing”. This would be an interesting and indeed historically ironical result, because what propelled the development of set theory was precisely the conviction that infinity, *pace* the Aristotelian tradition, *can* consistently be regarded as actual and not only as potential; it turns out that in the face of the paradoxes, the *Universe of set theory* -the *totality of totalities*-, along with other similarly problematical collections, can only be regarded as a potentially infinite collection (although one which contains some actually completed infinities as members and subsets). Ironical or not, it does look as if an Aristotelian conception of the Universe of sets is the only way to save what is valuable in the Cantorian grappling with the Absolute while at the same time avoiding the paradoxes; moreover, it has been pointed out that set theory can play the theoretical role it is expected to play (the foundation of classical mathematics) without assuming the actual existence of an absolutely comprehensive totality: Parmenidean aspirations aside, set theory can do very well with a potential, ever-growing universe instead of an actual one.²

¹ For further discussion of the claim that we tend to reify collections, i.e. treat them as individual entities in their own right, whenever we speak about them and differentiate them from one another, see footnote 3 in page 110 of this work.

² Given the axiom of extensionality, a Cantorian set cannot be said to grow or change its size in any way. This means that what an Aristotelian could see as an ever-growing universe, the theorist who refuses to admit the existence of any other collections apart from definite ones (sets) is forced to see as a plurality of different “universes”. Cf. Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 118: “When we try to reconcile the image of

There is at least one difficulty with this proposal.¹ It is the following one. Cantor had a principle which has been called (by Michael Hallett, 1984, pp. 7 and 28) the “Domain Principle”. According to it, any potential infinity presupposes a corresponding actual infinity, which is the fixed domain of variability of the potential infinity. Now, if the universe of set theory is a potentially infinite as opposed to an actually infinite collection, then either there must be a corresponding universal set, which we have seen to be paradoxical, or else the universe of set theory would be a collection for which Cantor’s Domain Principle would not apply. For there could be no other collections “outside”, so to speak, of the collection of all collections, and thus there could be no fixed domain of variability of a potentially infinite universe.

The Domain Principle has some appeal. Consider the most basic kind of infinity, that of the natural numbers. We can, if we wish, think of this totality as potential and not actual. We can say that all we mean by calling it an infinite totality is that no matter how large n is, $n+1$ would always be a legitimate natural number. It seems, however, that this cannot be the whole story, since we certainly do not think that the natural successor of some as yet unreached n could be, say, $n.39803545487656\dots$. For $n.39803545487656\dots$, unlike $n+1$, is a member of the set of real numbers. So these two infinities must be definite. Determined *already*, so to speak, and given all at once. But this is what is meant by the concept of an actual infinity. Hence Cantor’s Principle:

There is no doubt that we cannot do without variable quantities in the sense of the potential infinite; and from this it can be demonstrated the necessity of the actual-infinite. In order for there to be a variable quantity in some mathematical study, the ‘domain’ of its variability must strictly speaking be known beforehand through a definition. However, this domain cannot itself be something variable, since otherwise each fixed support for the study would collapse. Thus this ‘domain’ is a definite, actually infinite set of values.²

the ever-growing universe with our desire to talk about the truth or falsity of statements that refer to *all* sets we are led to assume that some ‘temporary’ universes are as close an approximation to the ultimate unreached universe as we wish. In other words, there is no property expressed in the language of set theory which distinguishes the universe from some ‘temporary universes’.” A.W. Moore ((2001) p. xiv) also notes that there is “something strikingly Aristotelian” about the universe as conceived by modern set theory, namely its potential, as opposed to actual, infinitude.

¹ I discuss a second difficulty in the next chapter, section 3.2, pp. 94-6. It is the difficulty that the conception of the universe as potentially infinite, which usually goes hand in hand with the claim that the universe is an ever-growing entity, seems to be incompatible with a realist conception of mathematical objects. I leave this discussion for the next chapter because there we will have in view another collection theory (namely, plenum theory) which conceives of the universe not only as ever-growing but as “self-amplifying”. The idealism charge will be clearer and more pressing having both set theory and plenum theory in mind.

² Quoted by Hallett (1984), p.25.

Another way to express this point is to say that the collection over which the variables of a particular theory range must behave according to the axiom of extensionality: it must be a fixed domain or set.¹ What if this particular theory happens to be set theory itself? Its variables must, if the Principle applies here as well, range over a set. So there *must* be a set of all sets. This is, in effect, the conclusion that Graham Priest draws:

According to Cantor's Domain principle [...] any variable presupposes the existence of a domain of variation. Thus, since in ZF there are variables ranging over all sets, the theory presupposes the collection of all sets, *V*, even if this set cannot be shown to exist in the theory.²

That there must be a paradoxical set of all sets is anything but a problematical result for Graham Priest, who thinks that there are "truly contradictory objects".³ But this dialetheist position could be avoided if the commitment to the existence of a set of all sets is not forced upon us by logic. This way out would be available if Cantor's Domain Principle is not allowed to hold sway at least in the particular case of the universe of set theory. Let us explore this possibility in what follows.⁴

Note, first, that *Cantor himself* did not apply "Cantor's Principle" in the present context. For he took the paradoxes to confirm his view that certain multiplicities are *not* sets but unfinished totalities, collections that cannot be regarded as complete and actually existing.⁵ Thus he held the view that there are totalities which are not determinate in the way sets are. To these totalities the principle does not apply.

¹ Graham Priest provides the following example to illustrate Cantor's point: "For example consider the claim 'Let *z* be a root of the equation $ax^2 + bx + c = 0$.' This is true if *z* may be complex: false if *z* must be real." (2002) p. 125. In effect, the equation $ax^2 + bx + c = 0$ is equivalent to the equation $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$, as is well known. But if we circumscribe to the real numbers, this equation won't have a solution in many cases, since $\sqrt{b^2 - 4ac}$ will be meaningless when $(b^2 - 4ac)$ is negative: if you multiply a negative number by itself the result is positive. So in the realm of real numbers you can't get a solution to such a simple equation as $x^2 = -1$. The general meaningfulness of the equation therefore presupposes that the range of *x* be the set of complex numbers (i.e. the real plus the imaginary numbers). See Courant, R. [1941] pp. 89 ff.

² Priest, G. (2002) p. 158.

³ Priest, G. (2002) p. 166.

⁴ One could also, of course, try to criticize Cantor's Domain Principle by arguing that we have other means to account for the "fixed support of the mathematical study" which do not appeal to the existence of actually infinite domains of quantification. As far as I know, not only full-blown conventionalists but also philosophers inspired by Wittgenstein's later views on mathematics would follow this line of criticism; and whatever the inconvenient aspects of these positions might be, any non-Hegelian would be happier to accept those inconveniences than to accept that there is a 'truly contradictory object'. A similar point is made by Weir, A (1999) p. 124.

⁵ Dümont J and Mau F (1998) note that "where Priest recognizes another example of a true contradiction caused by the Absolute, Cantor takes the contradiction as a compelling reason to deny the existence of the Absolute as a 'unity' or as 'one finished thing'." p. 295. This is true, and it shows that Cantor does not apply his own principle where Priest does.

If the protest arises that admitting exceptions to the Domain Principle would be an *ad hoc* attitude, the straightforward reply is available that, for all its appeal, the principle is invalid -to say the least-. Richard Cartwright, who calls the Domain Principle the "All-in-One principle", has pointed this out:

The general principle appears to be that to quantify over certain objects is to presuppose that those objects constitute a "collection", or a "completed collection" –some one thing of which those objects are the members.¹

But in fact

It is one thing for there to *be* certain objects; it is another for there to be a *set*, or set-like object, of which those objects are members.²

A set is an object in its own right, whose behavior is described by set theory. The only reason that could be adduced for the existence of such an object whenever any other object exists seems to be the claim that there must be a set corresponding to any description.³ But, let us remind ourselves, this is just the unrestricted axiom of comprehension, which immediately yields paradoxes. This *reductio ad absurdum* of the principle is, in the eyes of a dialetheist like Graham Priest, the opposite of a reason to reject it. The rest of us are free take the *reductio* as a very good reason. We can still regard the "Domain Principle" as a generally reliable rule, if we please. But not as a principle forced upon us by logic.⁴

¹ Cartwright R L (1994) p. 7.

² Cartwright (1994) p. 8.

³ Confronted with the accusation that he falls pray of the All-in-One Principle, Priest (2002) claims that if the accusation attributes to him a failure to see that the needs of quantification of any sentence of the form "*The so and so ...*" are already served as long as the *so* and *so* exist, the accusation is false. He admits that no further entity is required in order to fulfill the referential needs (ontological commitments) of such a sentence. But he adds (pp. 281-2): "The totality [which has the *so* and *so* as members] is presupposed as soon as we talk, not of the possible references of the variables, but of the sense of the sentence containing the quantifier. [...] it is not the ontological commitment of the sentence that is at issue here; it is the sense of the sentence; the determinacy of this does require there to be a determinate totality [...]"

⁴ Alan Weir offers a third alternative. Taking as a starting point the fact that the package which consists of naive set theory plus classical logic yields contradictions, Weir challenges the traditional diagnosis according to which the culprit of the contradictions is naïve set theory. In his view, it is classical logic which has to be revised. Weir does not claim, though, as Priest does, that logic would have to be thus modified as to admit some contradictions as true. His proposal is rather that the structural rules of logic, as opposed to its operational rules, would have to be modified. See his (1998 pp. 767 ff.) and (2006 pp. 341 ff.) An evaluation of this proposal of Weir's is beyond my capabilities. The view Weir wants to defend, however, includes the idea that the absolutely all-embracing domain is a member of itself, and I will express my dissatisfaction with this idea in the following chapter.

If what I've been saying is correct, the idea of a universe of sets which is not itself a set is *prima facie* a plausible one. At least, it should not be rejected on account of its "violation" of Cantor's Domain Principle. But, as we are about to see in the next chapter, this doesn't mean that all the problems for the concept of an all-embracing collection have been cleared out.

Conclusion

In this chapter we had a first approach to the problems surrounding the idea of an all-embracing totality. We saw that if this totality is a definite collection or set, and hence the set of all sets, paradoxes arise. We also saw that this is a significant philosophical problem, since what is at issue is the question whether there is such a thing as everything. We said that the concept of an all-embracing collection could be still recovered if we think the totality in question as a *potentially* infinite, and hence not a *definite* collection. In fact, all we have argued for so far is that, in thinking of the universe as a potentially infinite collection, we need not worry that we are violating Cantor's Domain principle. For this principle, or rather its defense, by invoking the idea that "no fixed support for the study" would be available if the collection we are trying to grasp is not a definite one, is hardly distinguishable from unrestricted axiom of comprehension, which is what generates the paradoxes in the first place.

This doesn't mean, however, that all problems have been addressed, let alone solved, so far. For if the universe is still thought of as a collection that contains absolutely everything, it must contain itself. Now, as we will see in the next chapter, there are ways of affirming self-membership while avoiding the traditional paradoxes. But do we really understand what it is for something to be a member of itself?

Let us see.

Chapter III. Two Ways of Saving Everything From Paradox

-Okay, Dude. Have it your way.

The Stranger in *The Big Lebowski*

We have dealt so far with two sorts of antinomies. Kant's alleged antinomy concerns the concept of the spatiotemporal world, understood as the totality of spatiotemporal objects. In set theory, we have seen paradoxes arise with respect to the concept of the absolute totality of sets. In moving from one problem to the other, we have pointed out that Kant's antinomy is not genuine, while those antinomies of set theory are, if not unavoidable, to be avoided by abandoning some assumptions. In particular, we saw that the concept of an all-embracing collection still makes sense as that of an ever-growing totality. This way out might be unpalatable to philosophers who would admit no exception to Cantor's Domain principle, and which think therefore that all (collective) being must be definite being. But since this principle is hardly distinguishable from the unrestricted axiom of comprehension, which appears to be the culprit of the paradoxes¹, the objection to an Aristotelian conception of the universe of sets which is based on the Domain Principle can be set aside.

If there is something really naïve, however, it is to think that the Parmenides Within can be dealt with so easily. What if we still want to keep a *definite* entity which is the collection of all collections, no matter how complicated our principle for the derivation of other collections gets? One of the systems which we present in this chapter takes this route. I am referring to the system called "New Foundations", which stems from the work of Quine (1937). The other system, which also attempts to theorize an all-embracing collection, is not an alternative set theory but an alternative to set theory. It is Nicholas Rescher's and Patrick Grim's "Plenum Theory". Plena are not definite entities, and hence not sets. Thus, the way in which these systems resemble one another and differ from one another is philosophically interesting: while for both systems there is such a thing as Everything, and

¹ Unless we want to see the culprit not in "naïve set theory" (characterized by its commitment to the unrestricted axiom of comprehension) but in the *logic* used in the derivation of the paradoxes, as I pointed out in the last note of the last chapter, in reference to the work of Alan Weir.

it is conceived as a collection, according to one system *Everything* is a definite being, while according to the other it isn't.

My approach to these systems is somewhat skeptical. As far as I can see, Quine, on the one hand, and Rescher and Grim, on the other, do succeed in showing that the paradoxes that traditionally beset totalization are not unavoidable. This doesn't mean, however, that these authors show that the concept of *Everything* is unproblematic. On the contrary, given that their concept of *Everything* is that of a collection of collections, nay that of the collection of *all* collections, the object referred to by this concept must be a self-containing object, a self-membered collection. And a dissatisfaction with the idea that something contains itself as a member is, if not universally shared, by no means universally disowned either.

3.1. Avoiding The Diagonalized Collection

3.1.1. Quine's New Foundations: An Alternative Set Theory

We saw in chapter II that the combination of the following two presuppositions leads to paradox:

- A) For any condition ϕ expressible in the language of set theory, there is the set of exactly those things that fulfill ϕ . (Unrestricted Axiom of Comprehension).
- B) Both ' $x \in x$ ' and ' $x \notin x$ ' are legitimate (meaningful) expressions of the language of set theory.

The paradox was, as we saw, that A) and B) together imply that there is a set of exactly those things that are not members of themselves: $\{x \mid x \notin x\}$, the so-called *Russellian set*. Since this set is a member of itself if and only if it isn't, we have a paradox, which is known as Russell's paradox. We also saw that Russell's way out of the problem was to reject presupposition B), and thus to deny that ' $x \in x$ ' and ' $x \notin x$ ' are meaningful expressions of the language of set theory. Now, if there was a set of all sets it would obviously have to be a member of itself. Russell's solution therefore entailed that it is not even meaningful to say that there is a set of all sets.

But Russell's rejection of presupposition B) is not the only available way out of Russell's paradox. Note that one thing is to *deny* that there are self-membered sets (and hence deny that there is a set of all sets), and quite a different thing is to discard as *meaningless* any statement to the effect that there are self-membered sets (as the set of all sets would have to be). If one simply denies that there are self-membered sets, but avoids any commitment with regard to the meaningfulness of set

theoretical expressions, one must still be committed to the denial of the existence of a set of all sets, but one is no longer allowed to reject as meaningless the assertion of the existence of a set of all sets.

Keeping presupposition B) is thus essential to the defense of the belief in the existence of a set of all sets. But presupposition B) in conjunction with presupposition A) is a one-way ticket to paradox. Accordingly, if one is interested in defending the existence of a set of all sets, one will naturally look for ways in which to avoid the acceptance of the unrestricted Axiom of Comprehension. This situation represents indeed a great challenge to the set theorist. Given that set theory is supposed to provide a conceptual foundation for mathematics in terms of logic alone, it is an important desideratum for set theory to be able to identify its objects (i.e. sets) by means of a condition which should be as arbitrary as possible. Rejection of presupposition A) means that one has to try to meet this desideratum of generality by other means than by what seemed to be the most natural and intuitive one. Zermelo-Fraenkel set theory faces this challenge as well, for it contains a restriction of the axiom of comprehension. And, as we saw in chapter II, it seems that ZF indeed meets the challenge, since it manages to avoid the paradoxes while at the same time succeeding in providing the conceptual foundations for classical mathematics. But ZF's way of doing things is of no avail at the moment, since we know that the particular way in which ZF restricts Comprehension also implies the nonexistence of the set of all sets. The surrogate for unrestricted Comprehension in ZF is the axiom of separation, *Aussonderung*, also called the axiom of subsets: *Given a set α there is the set of those subsets of α that satisfy any arbitrary condition whatsoever.* This was a clever way to meet the desideratum of generality while at the same time avoiding the paradoxes, but it also entailed that there is no set of all sets. If there was, the axiom of separation would guarantee the existence of the Russellian set, which is paradoxical. So Zermelo's way of restricting Comprehension cannot be what we are looking for at the moment.

A system of set theory which would fulfill all the three desiderata at issue (enough generality to provide a foundation for mathematics, avoidance of the usual paradoxes, existence of a set of all sets) was proposed by Willard Van Orman Quine, and it is known as New Foundations (or simply NF), due to the title of the 1937 paper where Quine first presented it.¹ In order to understand the way in which this system succeeds in meeting the current challenges, we must refer to another aspect of Russell's attitude to the paradoxes.² According to Russell, in order to avoid paradoxes (and

¹ "New Foundations For Mathematical Logic" Quine (1937).

² There was, in fact, no unique way in which Russell responded to the paradoxes; at least three different theories have been identified by which Russell attempted to provide a paradox-free foundation for

not only the set theoretic ones), it is crucial to realize that the objects of which we can meaningfully speak often belong to different *logical types*. The sets that generated the paradoxes are such that their very definitions, or the expression of some fact implied by these definitions, neglects the proposed type-distinction. The Russellian set, e.g., is defined as the set of all sets that are not members of themselves, and the correct expression for that set is ' $\{x_n \mid x_n \notin x_n\}$ '; but this expression makes as little sense, for Russell, as ' $\{x_n \mid x_n \in x_n\}$ '. Whenever our expressions contain the membership symbol, that symbol must be flanked by terms which refer to objects of consecutively ascending logical types. If bottles of wine are of logical type 1, the *set* of bottles of wine is of logical type 2, the *set of sets* of bottles of wine is of type 3, and so on. Whenever this distinction is not respected by our expressions, and hence whenever the terms flanking the expressions *could not* be subindexed in the consecutively ascending order, those expressions are, according to Russell, meaningless.

We can see more clearly now what we said at the beginning of this section. Russell's solution to the paradoxes was *not* to reject the unrestricted axiom of comprehension. He kept the presupposition that to any condition there must be the set of things that satisfy that condition. But, of course, this condition must be a legitimate one, i.e. one allowed by the rules of the language of the theory in question. But given the consecutiveness restriction, the expressions ' $x \in x$ ' and ' $x \notin x$ ', when properly written as ' $x_n \in x_n$ ' and ' $x_n \notin x_n$ ', are clearly seen to be illegitimate strings of signs. Think now of the universal set. Every object x belongs to it. The most natural expression which refers to that set is: ' $\{x_n \mid x_n = x_n\}$ '. This expression is *prima facie* allowed, since an object is of course of the same logical type as itself, so the type-numbers subindexed to the variables at the left and at the right of the identity sign must be identical. Now, one expresses the idea that there is an object to which every object belongs this way: ' $\exists y \forall x (x \in y)$ '. But there is no way of expressing this idea while at the same time respecting Russell's restrictions, since the y in question has no logical type, and a *fortiori* no logical type which would be consecutive to that of x , for every x . And, again, the existence

mathematics: the "limitation of size" doctrine, the "zigzag" theory and the "no class" theory. As has been pointed out, the main systems of set theory of the twentieth century are to a great extent descendants of one or other of these Russellian responses. For example, the doctrine of the limitation of size, according to which a collection is a set if and only if it is not equinumerous with the collection of all sets, is said to be one of the main ideas behind the development of ZF. Quine's NF, which we are about to consider, is said to be a descendant of Russell's "zigzag" theory, since the latter was a type-free system of logic which attempted to avoid the paradoxes by means of a restriction of the Axiom of Comprehension, but without sacrificing a collection of all entities. And the "no class" theory, according to which classes are nothing but logical fictions, is the one doctrine to which Russell himself subscribed after abandoning the previous two. It is, however, not unjustified to characterize the ideas that I am about to expound in the text (namely, the distinction between logical types) as "Russellian", since even when Russell considered classes as logical fictions he insisted upon his early distinction between logical types to which these fictional entities belong, and even contemplated a much more complex system of these types. See A. Urquhart (1988), *passim*.

of a set of all sets would imply that there is at least one set which belongs to itself (namely, itself), and the expression of this fact is, as we have seen, illegitimate.

Now, the *explicit* subindexing of the variables which appear in the expressions of set theory was only a requirement for an ideal exposition of that theory. In general, Russell had no problem with an expression like ' $x \in y$ '. This was, in Russell's view, a legitimate expression, even if the variables were not actually indexed, because *it is possible to subindex them* in such a way that the consecutiveness restriction is respected, i.e. in such a way that x and y are of ascending logical types. The opposite is the case with ' $x \in x$ ', since every object must have the same logical type as itself, and so *it is not possible to subindex* the variables of ' $x \in x$ ' in such a way that x has a different type than itself. Respecting this minimal requirement, i.e. the requirement that the variables contained are *able to be subindexed by consecutive numerals whenever they occur at the left and at the right of the membership sign*, the expressions of set theory were perfectly in order according to Russell. This meant not only that, say, ' $x \in y$ ' is a legitimate expression; it meant, as well, that when a formula is legitimate although not fully written with all its sub indexes, it could actually be an expression of infinitely many formulas. ' $x \in y$ ' could stand for ' $x_1 \in y_2$ ', but also for ' $x_2 \in y_3$ ' and for ' $x_3 \in y_4$ '. In general, ' $x \in y$ ' is a legitimate expression which could stand for any x and y as long as the logical type of y is $n + 1$ whenever that of x is n . As such, ' $x \in y$ ' is, although legitimate, *typically ambiguous*.

The realization of this aspect of Russell's theory of types is crucial in the understanding of Quine's system. Those typically ambiguous formulae whose variables can be supplemented with subindexes in such a way that Russellian legitimacy restrictions are respected are called *stratified* formulae by Quine.¹ Stratified formulae, that is to say, are those that would be meaningful according to Russell. But Quine does not go as far as saying that only stratified formulae are *meaningful*. He only says that stratified formulae are the only ones which, according to his proposed axiom of comprehension, would guarantee the existence of the set of things which satisfy the condition expressed by the formulae:

The alternative departure from the theory of types that suggests itself, finally, is this: that we reconstrue unindexed variables as truly general variables (rather than as typically ambiguous), but yet keep the restriction on [the axiom of comprehension: $\exists y \forall x (x \in y \leftrightarrow Fx)$] that the particular formulas put for ' Fx ' must be stratified.²

¹ Quine (1937) pp. 90-1.

² Quine (1963) p. 288.

Russell's idea was that what Quine calls unstratified formulae are not meaningful and therefore not legitimate in the language of set theory. Hence, although Russell kept an unrestricted axiom of comprehension, paradoxical sets were not allowed in his system since certain facts about them, when not their very definitions, could only be expressed by means of unstratified formulae. Quine, on the other hand, *restricts* the axiom of comprehension. His axiom of comprehension says: For any condition acceptable under Russellian standards, there is the set of exactly those things that fulfill that condition. Those formulae which are *not* acceptable under Russellian standards are not declared meaningless; it is just that *we are not entitled by the new axiom of comprehension* to assume the existence of sets whose defining conditions are expressible only by means of unstratified formulae. But –and here lies the crux of the whole matter– one thing is to say that we are *not entitled* to derive the existence of an object from our acceptance of a certain presupposition, and quite a different thing is to say that such an object *does not exist*, let alone that the sentence expressing its existence is *meaningless*.

The difference between Quine's NF and Russell's type theory is thus enormous, at least with regards to our present concerns. True, Quine avoids Russell's paradox by fairly Russellian considerations. First, the condition by which the Russellian set is specified is, as we saw, unstratified. Hence, in NF you cannot *derive the existence* of such a set by NF's restricted axiom of comprehension. On the other hand, we can reject the *existence* of such a set precisely because we know that it generates a paradox, namely that it belongs to itself if and only if it doesn't. So far so good. No one will really miss Russell's set. But what about the set of all sets? We saw that some facts about it can only be expressed by unstratified formulae, for example that it is a member of itself or that it contains every object of every type. This means that the existence of the set of all sets *cannot be derived from the axiom of comprehension using those unstratified formulae* to define the set. But this does not mean that the set of all sets does not *exist*. Since Quine does not see unstratified formulae as meaningless, and hence for all we know they could very well express a fact, there is in principle no semantic problem with the belief in the existence of a set of all sets. Moreover, given that the condition ' $x_n = x_n$ ' is stratified, we *are* in fact entitled by NF's axiom of comprehension to derive the existence of the set that contains exactly those things that are self-identical; but this is just the set of all sets.

But what of Cantor's paradox? We saw in chapter II that this was the most straightforward way to deny the existence of a set of all sets. According to Cantor's theorem, given any set α there is another set with more elements than α , namely the Power-set of α . If there was a set of all sets, U , the Power-set of U would have more elements U . But given that by definition U is the most inclusive

of all sets, there can't be a more inclusive set than U . Cantor's theorem, then, is incompatible with the existence of a set of all sets. What does Quine's system make of Cantor's theorem?

In Quine's system "Cantor's theorem" is simply not a theorem. But didn't we present step by step the *proof* of that theorem in chapter II? Yes, we did. But that proof is valid only under the presupposition either of an unrestricted axiom of comprehension or of *Zermelo's* restricted version of this axiom (*Aussonderungsaxiom*). But we have seen that Quine's way of restricting comprehension is precisely *not* Zermelo's. And Quine's restricted version of the axiom is such that the proof of Cantor's theorem is no longer available. Let us look at this in more detail.

The strategy for proving Cantor's theorem was a *reductio ad absurdum*. We first assumed that to any given set α there is its Power-set (the set of all its subsets), $\mathbf{P}\alpha$. Then we assumed that these two sets were equal in size, i.e. that they had the same number of elements. This was done according to the general criterion for size comparison in set theory: two sets are equal in size if there is a bijection between the two of them, i.e., in this case, if there is a one-to-one correspondence between the elements of α and the elements of $\mathbf{P}\alpha$. It was, as we saw, from the assumption of this one-to-one correspondence that a contradiction followed. For if there is a function f which correlates one-to-one the elements of α with those of $\mathbf{P}\alpha$, each element x of α will be assigned by f an element of $\mathbf{P}\alpha$ referred to by the expression ' $f(x)$ '. But we were asked by Cantor to form the concept of the set of those elements of α that are not members of the element of $\mathbf{P}\alpha$ with which they are correlated by the function f . That is to say, we were asked to frame the concept of the set $\{x \mid (x \in \alpha) \ \& \ (x \notin f(x))\}$. Let's call this set the "*diagonalized set*", or simply "*D*". Being a subset of α , D must be a member of $\mathbf{P}\alpha$. So it must be correlated by f with a member of α . But with what member w of α could D be correlated? Not with any member of α which is also a member of D , for in that case D would have a member, w , which is an element of the set with which it is correlated by f ; and this is impossible given the definition of D . But neither can D be correlated by f with an element w of α which is *not* a member of D . For in that case the definition of D would imply that w is, after all, an element of D . Since these are the only two possibilities for members of α (i.e. either $w \in D$ or $w \notin D$), and each implies its opposite, it follows that D is not correlated with any element w of α , which contradicts the presupposition of a bijection with which we started. Having derived this contradiction from the presupposition of a one-to-one correspondence between the elements of α and those of $\mathbf{P}\alpha$ Cantor concluded, as we know, that $\mathbf{P}\alpha$ must have more members than α , for any α .

But this proof, as we said above, is no longer available in Quine's system. The existence of the diagonalized set D , which is crucial in the derivation of the contradiction, is no longer guaranteed by

the axiom of comprehension of NF. This axiom guarantees the existence only of those sets whose specifying conditions use exclusively stratified formulae. But the diagonalized set D is not such a set since its defining formula is not stratified. The defining formula of D is $\{x \mid (x \in \alpha) \ \& \ (x \notin f(x))\}$; this would be a stratified formula if “ $(x \in \alpha) \ \& \ (x \notin f(x)) \ \& \ (f: \alpha \rightarrow P\alpha)$ ” were stratified. But this latter formula is not stratified because, on account of being to the right of a (negated) membership symbol, $f(x)$ would have to be assigned a type numeral consecutive to that of x ; but, on account of being a member of $P\alpha$, and hence correlated with $x \in \alpha$ by f , $f(x)$ has to be assigned the same type numeral as x (i.e. $\langle x, f(x) \rangle \in f$, where function f is supposed to be a set specified by conditions which use exclusively stratified formulae). Thus, given that the defining formula of the diagonal set D requires that a term be assigned two different type numerals, it is an unstratified formula, and therefore the existence of D is not guaranteed by the axiom of Comprehension used in NF.¹ Having no guarantee of the existence of D , in Quine’s system we have no means of saying that it is absurd to presuppose that there is a one-to-one correlation between the members of α and those of $P\alpha$, for any α .

But does this rule out the existence of D ? Even if one cannot prove the existence of D by means of NF’s axiom of comprehension, how can we be so sure of the non-existence of that set? We have insisted upon the difference between being unable to prove that a fact obtains (or fails to obtain) and the actual obtaining (or non-obtaining) of the fact itself. So how can we be sure of the non-existence of D ?

Quine was, of course, fully aware of this question. His answer was that sometimes we can, and sometimes we can’t, be sure of the non-existence of the diagonalized set. And the case with which we are now concerned, namely that of the set of all sets, is one of those cases in which we can be sure that the diagonalized set D does not exist. To see this, let us recall the connection, pointed at in chapter II, between Russell’s and Cantor’s paradoxes. We said that in order to get Russell’s paradox from Cantor’s paradox, the only thing that we need to do is to think of the function f as the *identity* function. It is the function that correlates each object with itself. So the diagonalized set D , $\{x \mid (x \in \alpha) \ \& \ (x \notin f(x))\}$, becomes in this case $\{x \mid (x \in \alpha) \ \& \ (x \notin x)\}$. But recall as well that the set α in question is not just any α ; it is U , the set of all sets. So D is in this case identical with $\{x \mid (x \in U) \ \& \ (x \notin x)\}$. Now, given that if there is a set of all sets then every set belongs to it, the latter expression of D is equivalent to $\{x \mid x \notin x\}$. But this is just the Russellian set, which we know to be paradoxical. So in Quine’s system we can, after all, deny the existence of the diagonalized collection on the grounds that it generates a paradox. In Quine’s words:

¹ Cf. Forster, T. (2006), end of section 3.

Since in the present instance α is U [i.e. the universal set] and the correlate of a subclass is that subclass itself, the class D becomes the class of those subclasses of U which do not belong to themselves. But [NF's axiom of comprehension] provides no such class D . Indeed, D would be $\{x \mid x \notin x\}$, whose existence is disproved by Russell's paradox.¹

The crucial aspect of Quine's avoidance of the set theoretical paradoxes is, then, a particular way of restricting the axiom of comprehension, namely to stratified formulae. And we have seen that at least the paradoxes most directly linked to the concept of the set of all sets, i.e. Russell's and Cantor's paradoxes, are successfully avoided by Quine's system.

Quine's system is an alternative to the ideas defended in the previous chapter. We said there that the set theoretic paradoxes are philosophically significant because they arise whenever we have the concept of a definite totality which is thought of as absolutely comprehensive. Quine would agree, I think, with the claim that if there were set theoretical paradoxes, some of them (especially Cantor's) could be interpreted as threatening a concept which some philosophers (for whatever reasons) are keen on salvaging. Quine, like Russell, sometimes refers to the set of all sets as the class "to which absolutely everything belongs"², the "class of absolutely everything"³. But Quine, unlike Russell, thinks that no logical problem has been shown to be inherent to this concept. Not by the so-called set theoretical paradoxes, at any rate. If one is careful enough, those paradoxes can be avoided, as we have seen.

Now, a usual minimum requirement for the acceptance of a set-theoretical system is that it does not entail a contradiction, i.e. that it be a *consistent* system. The consistency of Quine's NF hasn't been proved, but no inconsistency has been discovered in it either. And the usual inconsistencies are, let us repeat, avoided. Thus, with regards to this consistency requirement, Quine's system is on a par with that system to which it provides an alternative: the standard Zermelo-Fraenkel set theory⁴. But the consistency requirement is not the only one that is important in the evaluation of a system of set theory. A system of set theory is also evaluated according to whether it does the job

¹ Quine (1937) pp. 92-3, n. 10. I am modifying Quine's notation in a number of ways. The class whose existence is said by Quine to be disproved by Russell's paradox is actually $\hat{y}(y \notin y)$. This expression refers to the class of all *sets* which are not members of themselves, as opposed to $\{x \mid x \notin x\}$, the class of all *classes* that are not members of themselves. But the distinction between $\hat{y}(y \notin y)$ and $\{x \mid x \notin x\}$ is more than a verbal one only for those systems for which the distinction between sets and classes is more than verbal, and since NF is not one of those systems, we can neglect that distinction for the time being. Other modifications are that I use α instead of k , D instead of h , and U instead of V . The modifications, of course, are only made in order to accommodate the language of Quine's passage to that which I have been using throughout the text.

² Quine (1937) p. 92.

³ Quine (1963) p. 46.

⁴ Cf. Forster, T. (2008) p. 106.

that it is supposed to do. This job is to provide the logical foundations of (classical) mathematics, i.e. to show how all the concepts and operations of (classical) mathematics can be explained in terms of the concepts, axioms and theorems of the set theoretical system under evaluation. Judged by these pragmatic standards, Quine's system has not really fared better than the set theory of Zermelo.¹ It is for this reason that the system ZF is referred to as the *standard* system of set theory. But this fact should not mislead us. *Our* concern in this text is philosophical rather than technical. The one question with which we have been dealing throughout our investigations is whether the concept of an absolutely all-embracing totality is a self-contradictory one or not. To put it in Russellian words, our question is whether (it is consistent to say that) there is such thing as *Everything*. The reason why we are dealing with set theory is not that we have a mathematico-combinatorial interest in sets, but rather the fact that the paradoxes that were discovered at the dawn of this discipline led some thinkers to the conclusion, precisely, that there is no such thing as *Everything*. Seen under this light, Quine's system, independently from its pragmatic drawbacks, represents indeed a relevant response to the concerns of our previous chapter. Only under certain presuppositions –namely an unrestricted axiom of comprehension or Russell's views on the meaninglessness of ' $x \in x$ '– is the concept of a set of all sets self-contradictory or meaningless. But these presuppositions are not obvious or incapable of revision.

Later in this chapter we will come back to the question of the meaningfulness, or lack thereof, of a formula like ' $x \in x$ '. For the moment, let us take as established that Quine's system offers a *possible* (if not necessarily pragmatically *plausible*) way out of those paradoxes which looked so unavoidable in chapter II by any other means than by an ever-growing universe.

3.1.2. Plenum Theory: An Alternative to Set Theory.

One way in which one might try to keep a concept akin to that of *Everything* is, as we just saw, an alternative set theory in which the concept of a set of all sets cannot be proved to generate contradictions. But there are other ways of salvaging *Everything* from paradox. In view of some of the technical difficulties of NF referred to in the previous subsection, Nicholas Rescher and Patrick

¹ NF has been shown to be incompatible with the axiom of choice and incapable of delivering a proof of all instances of mathematical induction, results which represent a disadvantage of NF with respect to ZF. A related system, NFU, i.e. NF plus the assumption that there are individuals or Urelemente, has been shown (by Jensen R. [1968]) to be both consistent relative to ZF and to be able to cope with these difficulties, but this hasn't deterred ZF from being the system most standardly used in mathematical and foundational work. See Bar-Hillel (1967), p. 171, Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 163. For positive accounts of NFU, see Forster, T. (1995) and (2006), and Holmes, M. Randall (2006, revised in 2010).

Grim have proposed that what we should strive for is not so much an alternative set theory but rather an alternative *to* set theory. “There is some irony”, they say, “in the fact that sets have been called on to play a philosophical role for which they are essentially unsuited.”¹ And by this they mean not only that in set theory the oversize collections which are “intuitively plausible”² are bound to generate problems; they mean, as well, that it is incorrect to attempt to categorize every plurality in terms of a specific kind of plurality. When we realize, moreover, that sets are expected to conform to the axiom of extensionality (two sets are identical *if and only if* they have the same members), and hence that set theory simply cannot make sense of the idea that a set *acquires a further element* or *loses one of its elements*, i.e. when we realize that it is absurd to say that a set *changes its size over time*, it is easy to see the point of these authors when they claim that “set theory is not more adequate to handle [certain kinds of] collectivities than eclipses”.³ What if the universal collection is not an entity with a *definite* size?

Amongst the collectivities whose behavior Rescher and Grim attempt to describe, the ones that are relevant for our present concern are those that they call “plena”:

A *plenum* as here understood is a collectivity that contains distinct entities corresponding to each of its sub-collectivities, where sub-collectivities follow the same pattern as subsets:

$$(\forall s) (s \subset P \leftrightarrow (\forall x) (x \in s \rightarrow x \in P))$$

The mark of *plena* is that *every sub-collectivity s of a plenum P is such that there is a member of P that exists in unique correlation with s* .⁴

Plena are thus collections similar to sets in some respects but utterly different from sets in other respects. They are similar to sets in that they are collections and in that we can talk about their sub-collections as those collections whose members, if any, are also members of the original collection. But plena have, for each of its subcollections, a member which is correlated with that subcollection. As the authors sometimes put it, each subcollection of the plenum is *represented* by a member of the plenum. Moreover, this correlation can very well be that of identity. When this happens, the authors say that we have a *membership* plenum.⁵ And given that, as we know very well, every collection is a subcollection of itself, this means that membership plena are members of themselves.

¹ Rescher and Grim (2011) p. 4.

² Rescher and Grim (2011) p. 59.

³ Rescher and Grim (2011) p. 98.

⁴ Rescher and Grim (2011) p. 61.

⁵ Rescher and Grim (2011) p. 62.

Certain totalities, and particularly those totalities with which we have been concerned throughout this investigation, should be characterized in terms of plena rather than sets, according to Rescher and Grim. We said in chapter II that if the concept of the set of all sets is self-contradictory, this meant that the concept of a totality of all objects or things is self-contradictory (as long as we do not adopt *a priori* a nominalist attitude, according to which problems regarding abstract objects like sets wouldn't endanger our ontology). Rescher and Grim do agree that if these totalities are conceptualized in terms of sets we are led into a dilemma: either we have to give up all conceptualization of these collectivities on account of their paradoxical nature, or we have to appeal to systems of set theory which are technically unsatisfactory. But the dilemma is escapable. Why should *Everything* be thought of in terms of sets? There is no *a priori* reason for this:

Collectivities also include the self-amplificatory collectivities we have termed plenary totalities: the universe as a whole, the totality of things, all propositions, states of affairs, facts or truths. Plena such as these quite explicitly explode beyond the limits of classical logic and set theory.¹

Let us defer the question of what "self-amplificatory" means in this passage to the next section of our present chapter. For the moment, let us concentrate on the claim that the concept of a plenum is the one that is going to allow us to conceptualize oversize totalities "and –in the end- to [conceptualize] the totality of everything at large."² But given that *Everything* is thought of as a *collection*, and that we are also allowed to use the concept of the subcollections of this collection, the most natural question is How can the theory of plena avoid a "plenary" version of Cantor's paradox?

Rescher and Grim do have an answer to this question. Their strategy is, first, to prove that, as a consequence of the characteristic by which a plenum is defined (namely, that every subcollection of a plenum has a correlate which is a member of the plenum), every plenum must have what Rescher and Grim call an "indeterminate collectivity" as one of its sub-collectivities³; having proved this claim, they show that Cantor's theorem is not valid for plena, thus paving the way for a paradox-free conception of plena such as *Everything*.

By an indeterminate collectivity Rescher and Grim mean a collection C such that it is not determined, for every object x, whether x belongs to the collectivity C or not. How can this be the

¹ Rescher and Grim (2011) p. 73.

² Rescher and Grim (2011) p. 60.

³ Rescher and Grim (2011) p. 60.

case? Well, think of a Russellian collectivity C^R of all non-self-membered collectivities. We know that it is a member of itself if and only if it isn't. And we know that $C^R \in C^R$ is true if and only if it is false. Rescher and Grim see these cases as genuine "gaps" in membership status and in truth-value.¹ C^R neither belongs nor fails to belong to itself –which shouldn't be interpreted as saying that it has a third membership status called "indeterminate", the claim being rather that C^R *fails determinately to have* a membership status. Its membership status (with respect to itself) is *oscillatory*: from the assumption of its belonging to itself it follows that it does not belong to itself, and from the latter assumption we are sent back to that of its being a member, and so on ad infinitum. This logical indeterminacy is not, however, the only kind of membership indeterminacy that Rescher and Grim envisage. They give the example of the collection described as "Atoms in certain mass of unstable transuranic material (Californium 252 Cf, for example) that will still be there 30 days hence."² According to our best physical theories, nothing in the history of the universe up to that moment determines which are the elements of that collection. It is not that we are unable to *know* which are those elements. The indeterminacy in question is not epistemic in any respect; it is not that we are unable to *establish* which are the members of the collection; it is rather an indeterminacy "in the things themselves."³

Rescher's and Grim's task is thus to show, firstly, that every plenum has an indeterminate collectivity as a subcollectivity; secondly, that this fact is enough to save the concept of a plenum (in particular, *Everything*) from a tailor-made version of Cantor's paradox. Interestingly enough, the first task is accomplished by an argument which parallels the proof of Cantor's theorem –i.e. by "diagonalizing out" the desired indeterminate collectivity-, while the second task consists precisely in showing how the existence of that "diagonalized" indeterminate collectivity implies that Cantor's proof does not work for plena. Let us look at this in more detail.

We said that every *subcollection* of a plenary collection C is correlated with a *member* of C by a bijection (a one-to-one function which can be that of identity). So every element of PC , the Power-collectivity of C , has a unique correlate which is a member of C . As Rescher and Grim say, each member of PC has a *partner* in C . Since this correlation is called a "transform", the element c of C

¹ These ideas are influenced by some approaches to the Liar paradox, as is admitted by Rescher and Grim, p. 13.

² Rescher and Grim (2011), p. 21.

³ Rescher and Grim (2011), p. 14. note that the logical indeterminacy which we mention in the text (that of the Russellian collectivity) is not epistemic either. They envisage at least six kinds of membership indeterminacy: *epistemic*, *semantic* (vagueness), *stochastic*, *ontological*, *counterfactual* and *logical*. This is not the place to explain all these sorts of indeterminacy; for our purposes it is enough to note that, for these authors, indeterminacy is not necessarily a subjective matter.

that is the partner of a unique element p of $\mathbf{PC}$ is symbolized as: $\rightarrow\text{Tr}(p)$. Equally, the element p of $\mathbf{PC}$ which is partnered with a unique element c of C is symbolized by: $\leftarrow\text{Tr}(c)$. Now, what about C^* , the collection of all those members of C that are not elements of that subcollection of C with which they are partnered? By definition, C^* is a subcollection of C ; it must therefore be a member of $\mathbf{PC}$. But then it must have a partner in C , symbolized as: $\rightarrow\text{Tr}(C^*)$. We ask now, is $\rightarrow\text{Tr}(C^*)$ a member of its partner C^* or not? Should we say that $\rightarrow\text{Tr}(C^*) \in C^*$ or rather that $\rightarrow\text{Tr}(C^*) \notin C^*$? As we can see, $\rightarrow\text{Tr}(C^*)$ is a member of its partner C^* if and only if it isn't, just as in the proof of Cantor's theorem the diagonalized set D , which was a member of $\mathbf{P}\alpha$, had a correlate in α that belonged to D if and only if it didn't. The membership status of $\rightarrow\text{Tr}(C^*)$ with respect to its partner in $\mathbf{PC}$ is logically oscillatory or indeterminate, just as the membership status of the Russellian collection with respect to itself is logically indeterminate. But instead of concluding from this logical indeterminacy that there is no such correlation that partners the members of C with those of $\mathbf{PC}$, Rescher and Grim conclude that C^* is an indeterminate (and hence a non-set) subcollection of C . And since C was thought of as any arbitrary plenary collection, the authors take their first task as accomplished: they draw the general conclusion that every plenum must have an indeterminate collectivity as one of its subcollectivities.¹

C^* is thus an element of $\mathbf{PC}$ and it is indeterminate. Now we must remember that amongst plena there are *membership* plena: the ones that *contain as members* each one of their subcollectivities. That is to say, if C is a membership plenum, C^* -an indeterminate collectivity- is not only a member of $\mathbf{PC}$ but also a member of C itself. But this means that we can no longer prove that $\mathbf{PC}$ is strictly larger than the original C . For nothing guarantees us now that there is a subcollectivity of C (and hence a member of $\mathbf{PC}$) which plays the role of the diagonalized set D in the proof of Cantor's theorem. The existence of the set D provided the contradiction required for the *reductio ad absurdum* of the presupposition that there was a one-to-one correlation between α and its Power-set $\mathbf{P}\alpha$. But what works for α and $\mathbf{P}\alpha$ need not work for C and $\mathbf{PC}$. Nothing assures us in the latter case that each element c of C has as its partner in $\mathbf{PC}$ a *determinate* collectivity. Both $c \in C^*$ and $c \notin C^*$ lack a determinate truth-value when $c = C^*$, and this means that the key contradiction which proved that there can be no bijection between α and $\mathbf{P}\alpha$ ($w \in D \leftrightarrow w \notin D$) is in this case simply the statement that characterizes C^* as a logically indeterminate collectivity ($c \in C^* \leftrightarrow c \notin C^*$):

Here again the strong rejection of the Law of Excluded Middle avoids contradiction. For a plenum of all sets, there is no guarantee that all its sub-collectivities are themselves sets. If they are not –if, in particular,

¹ Rescher and Grim (2011), pp. 69-70.

that which plays the role of the diagonal set D is not- then it need not hold that for any x either $x \in D$ or $x \notin D$. With an abandonment of the special situation of sets we lose such a presumption regarding membership. And without that presumption, we are no longer forced to say that the power set of a plenum is larger than the plenum itself. Indeed it cannot be, since plena are defined so as to contain distinct elements for each element of their power sets.¹

This is, then, the way in which Rescher and Grim argue that Cantor's theorem does not apply to plena, which means that a "plenary" version of Cantor's paradox is not available.² Now the collection of everything is not only a plenum but a *membership* plenum. In this case, the paradox is avoided in a much more straightforward manner: the diagonalised collection is of course a subcollection of the original collection, and since membership plena have by definition all their subcollections as members, the diagonalised collection is by definition a member of the membership plenum.

As can be seen, the strategy with which Rescher and Grim attempt to save *Everything* from contradiction resembles that of Quine's in important aspects, but also differs from it in other, equally important aspects. The two strategies are similar in that they work by defusing the proof of Cantor's theorem for the relevant collection, which then paves the way for the claim that there is in principle no inconsistency in thinking that the Power-collection of a certain kind of collection (non-Cantorian sets, plena) is of the same size as the original collection. And given that for the proof of Cantor's theorem the diagonalized set $\{x \mid (x \in \alpha) \ \& \ (x \notin f(x))\}$ is crucial, both strategies work by showing that the Power-collection of the relevant collection is not guaranteed to contain as a member any collection which plays the role of the diagonalized set. But NF and plenum theory manage to make the latter claim in different ways. In NF, the diagonalized set is not guaranteed to exist by the axiom of comprehension with which the system works, and the existence of such a set can be ruled out on the grounds that otherwise Russell's paradox would arise. Plenum theory, on the other hand, contains no restriction on the axiom of comprehension for plena. But no member of the Power-collection of the plenum is guaranteed to play the role of the diagonalized set D because *the role of D* can only be played by a set, whereas the Power-collection of a plenum has not only sets but also indeterminate collectivities as members. In the case of *membership* plena, the diagonalised collection is by definition a member of the plenum, and hence the existence of the crucial bijection (between a collection and its Power-collection) follows straightforwardly from our definitions.

¹ Rescher and Grim (2011), p. 77. See also their (2008) p. 435. I use the symbol D instead of Grim's and Rescher's S^* .

² Russell's paradox is defused in a much more straightforward manner, as we hinted at four paragraphs above in the text: the Russellian collectivity is not a set, and thus there is in principle no problem with the claim that its membership status (with respect to itself) is indeterminate.

So far we have been focusing on the main technical device by which NF and plenum theory manage to save *Everything* from contradiction. This technical device –namely the avoidance of the diagonalized set–, as we have just seen, allows us to identify both similarities and differences between NF and plenum theory. But the interesting similarities and differences between NF and plenum theory are by no means exclusively technical. These systems resemble and differ also in philosophical ways. To these we now turn.

3.2. “... and perhaps the mind ought to boggle...”

One idea in common between NF and Plenum Theory is that what in each case plays the role of *Everything* is said to be self-membered. This couldn't be otherwise: a collection of all collections must obviously contain itself as a member. But what exactly does this mean? Do we so much as understand the idea that a collection contains itself?

Let us start by looking at this idea from the point of view of set theory. George Boolos has put the problem of our understanding self-membership this way:

It is important to realize how odd the idea of something's containing itself is. Of course a set can and must *include* itself (as a subset). But *contain* itself? Whatever tenuous hold on the concepts of *set* and *member* were given one by Cantor's definitions of “set” and one's ordinary understanding of “element,” “set,” “collection,” etc. is altogether lost if one is to suppose that some sets are members of themselves. The idea is paradoxical not in the sense that it is contradictory to suppose that some set is a member of itself, for, after all, “ $\exists x (Sx \wedge x \in x)$ ” is obviously consistent, but that if one understands “ $\in$ ” as meaning “is a member of,” it is very, very peculiar to suppose it true. For when one is told that a set is a collection into a whole of definite elements of our thought, one thinks: Here are some things. Now we bind them up into a whole. Now we have a set. We don't suppose that what we come up with after combining some elements into a whole could have been one of the very things we combined (not, at least, if we are combining two or more elements).¹

¹ Boolos, G. (1971) in Boolos, G. (1998) p. 17-8. The expression “ $\exists x (Sx \wedge x \in x)$ ” means: there is an object x which is both a set and an element of itself.

Let us for the moment concentrate on Boolos's the temporal image.¹ If sets were temporal objects, i.e. objects that get formed in time, then the problem which Boolos is pointing at would be a real impossibility. For according to the axiom of extensionality, in order for a set to *be* (i.e. to be *the set it is*), all its members have to *be* there. Thus, in this temporal conception, a set cannot get formed until *all* its members are there. A self-membered set would be impossible, since it would have to wait forever to get formed or, alternatively, there would never be anything upon which the set could build its identity.

This temporal image cannot be the end of the discussion, however, and it has been shown to be deficient in at least two ways. In the first place, it is dangerously constructivist. Sets do not "get formed". We are at any rate not allowed to presuppose they do. Platonists think the exact opposite to that. For them, sets have a reality of their own, regardless of the constructivist capacities of any (real or idealized) subject. We should therefore avoid any language that suggests constructivism. In the second place, sets are abstract objects, and abstract objects are not usually thought of as having temporal relations; rather, the opposite claim is most naturally (although not necessarily) advanced, i.e. that sets, like all abstract objects, are atemporal.² Thus, again, those who want to reject self-membered sets should avoid a language that suggests that their argument depends on the idea that sets have temporal relations.

The conception of a set according to which a set cannot be self-membered³ need not be bound by constructivist or temporal metaphors, however. What is essential to this conception is the idea that there is a *priority* of the elements with respect to the set; one way to make sense of this is by a constructivist or temporal metaphor, but it is not the only way; we can, after all, make an explicit (Platonist) commitment to a metaphysical dependence of the set with respect to its elements.⁴ This is not to say, of course, that this idea of a metaphysical dependence is completely unproblematic. After all, when we try to make sense of this dependence, all we seem to be able to say is that a set would not have existed if its members had not existed; but this seems to conflict with the idea that

¹ As noted by Hallett (1983) p. 223, Boolos himself is careful enough to distinguish the temporal image –which is only a means of elucidating his ideas– from the formal account of the "iterative conception of a set", which is the one Boolos is defending here. See Boolos's (1989), also in his (1998). For an account which, unlike Boolos's, is according to Hallett too dependent on the temporal and constructivist metaphors, see Wang, H. (1977).

² See Lear, J. (1977) p. 88 n.5, Maddy, P. (1983) p. 117 and Dummett, M. (1991) p. 311.

³ Or have as a member a set which contains it as a member, etc. This conception is sometimes referred to as "the iterative conception of a set", but T. Forster (2008) disagrees with the nomenclature, since he thinks that other conceptions of a set, which admit self-membership, can make as much sense of iteration and of a recursive method for establishing the identity of sets, as the conception under which self-membership is excluded. Forster refers to the conception we are now discussing as the "cumulative hierarchy".

⁴ See Parsons, C. (1977) pp. 340-1.

there is an empty set, or sets which have the empty set as a member.¹ The defender of the so-called iterative conception could reply that the empty set would anyway either have to have a special status or force us to think of both individuals and one-element-sets as having a special status.² And in fact, it has always been a mystery how to accommodate the empty set to that Cantorian definition according to which a set is *a many thought of as a one, a plurality thought of as a unity*.³ That the empty set turns out to be exceptional in the metaphysical version of the iterative conception is hardly a specific problem of that conception.

It would be wrong to think, however, that any strong reasons have been provided in defense of the iterative conception, and hence against the admission of a set of all sets. True, if there is a set of all sets it belongs to itself, and this is incompatible with the iterative conception. But why should we accept the iterative conception? In ZF, given the axiom of foundation, no set x contains as a member. But at this point this cannot be seen as a *reason* in support of the iterative conception since in NF there simply is no axiom of foundation. Another way to put the question with which we are now concerned is, then, why should we accept the axiom of foundation?

It is sometimes said that

It conforms to a great extent with the intuitive notion of set that the only objects which can rightfully be called sets are the well-founded sets.⁴

But even those who say this admit that this is not the sort of statement with which we should be satisfied. For one thing, it would be on a par with a statement to the effect that "it conforms to a great extent with our intuitions that there *is* such thing as Everything; if the concept of the set of all sets is our theory's concept for Everything, and if the axiom of foundation conflicts with the

¹ Cf. Potter, M. (2004) p. 39.

² As Quine points out, the axiom of extensionality generates the dilemma that either there is only one individual and no empty set or there is an empty set but there are no individuals. For, if we have an axiom of extensionality, the identity of any objects A and B is determined by their members. Thus if A is a memberless object, it is identical with any other memberless object. If A is an individual (i.e. a memberless object), it is therefore identical to any other individual, and hence there is only one individual in the universe. And if A is a set (a memberless set), it cannot share with any other object the property of being memberless, so there are no individuals in the universe. Quine solves this problem by identifying any individual x with the set which contains that individual. $x = \{x\}$. But then, of course, $x = \{\{x\}\}$, since, if $x = \{x\}$ then $\{x\} = \{\{x\}\}$ and, given that the relation of identity is transitive, $x = \{\{x\}\}$. In this way you can argue $x = \dots\{\{\{\{\{\{x\}\}\}\}\}\}\dots$ iterating sets ad infinitum. About this issue, Quine says: "This result is *prima facie* unacceptable, since x is an individual and y [i.e. $\{x\}$] is a class. But actually it is a harmless result; none of the utility of class theory is impaired by counting an individual, its unit class, the unit class of that unit class, and so on, as one and the same thing." Quine (1963) pp. 31-2.

³ The one-element-sets are, of course, in conflict with this definition as well.

⁴ Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) p. 88.

existence of such a set, then what must go is the axiom of foundation, not the concept of the set of all sets”.¹ Since this (let’s call it Parmenidean) intuition is in principle on a par with the intuition of well-foundedness, appeal to intuitions is rather symptomatic of the fact that the discussion has reached a cul-de-sac. This is why most authors tend to agree that the criterion with which one ought to evaluate a set-theoretical system is not so much its “intuitive appeal” but its effectiveness in providing the conceptual foundations of mathematics. But here, again, it seems that both those who argue in favor of the “well-founded” conception, as well as those who argue against it, have arguments in support of their respective positions. On the one hand, it is widely agreed by both sides that *classical* mathematics can do very well without non-well-founded sets (let alone a universal set); but those who argue in favor of non-well-founded sets claim that some phenomena in computer science are best modeled by a set theory with an anti-foundational axiom, according to which there are as many non-well founded sets as possible.²

In the end, the best thing that the defender of the iterative conception can do is to point at the fact that we simply seem to be unable to understand what it means for a set to be a member of itself. As George Boolos says, “perhaps the mind ought to boggle at the idea of something’s *containing* itself.”³ But we can never be sure that this impossibility is more than an empirical fact about ourselves (and thus an anthropological or statistical fact).

Complaints about counterintuitiveness apart, however, the acceptance of both self-membership and the axiom of extensionality poses another problem. Quine himself knew this. He states the difficulty in the following way:

NF [...] can be further criticized for allowing self-membership, which beclouds individuation. The glory of classes, over against properties, is their clear individuation; classes are individuated only as clearly as their members. Under self-membership the individuation ceases to wind down.⁴

¹ I think the statement in quotation marks could be a paraphrase of what T.E. Forster says in his (1995) pp. 11-2. It should be said, though, that Forster is careful enough to point out that he is only stating the *motivations* of his work.

² See Aczel, P. (1988). p. xix. Some would even extend the applicability of non-well-founded set theories from computer science to mathematics: “There are several areas of mathematics where a type-free approach seems the most natural. For example, in recursion theory, it is quite natural to think of an algorithm which operates on algorithms, in particular which can operate on itself. This idea is in fact central to the general theory of algorithms; Kleene’s recursion theorem is based on this idea. Similar situations arise in group theory (where a group can act upon itself) and in category theory. Such mathematics is perhaps most naturally carried out in a type-free formalism like the lambda calculus”. A. Urquhart (1988) pp. 84-5.

³ Boolos, G. (1971) p. 15.

⁴ Preface to the 1980 edition of *From a Logical Point of View*, where Quine’s 1937 appears.

In effect, the axiom of extensionality tells us that if sets A and B have the same elements, they are identical. But if this is so, not even an idealized (infinite, transfinite, divine) subjectivity would be able non-circularly to identify –“individuate”- (the elements of) a self-membered collection. This does not seem to me to be a minor problem. Not one that could be shunned by setting aside as irrelevant all appeals to the capabilities of an idealized subject, anyway. Such appeals were never out of the question, as Shaughan Lavine points out:

Cantor thought of sets in terms of what God could do with them, and his Omnipotent Mathematician is still very much with us, even if theology has been taken out of the picture. Virtually every attempt to motivate the principles of combinatorial set theory still relies on some notion of idealized manipulative capacities, capacities of the Omnipotent Mathematician.¹

Thus, Quine seems to be saying that it is hard to conceive how even an Omnipotent Mathematician could individuate a self-membered extensional entity. True, that “it is hard to conceive” non-well-founded sets doesn’t mean that there is a contradiction in their concept. Moreover, the axiom of foundation has been proved to be independent of all the other axioms of set theory.² That is to say, the admission of self-membered collections is not logically incompatible with the axiom of extensionality. So those whose minds do not boggle are free to think that there can be self-membered sets, and hence that there can be a set of all sets.³ This is, indeed, a possible way of saving *Everything* from paradox.

There seems to be an alternative way, however. We have just seen that the axiom of extensionality makes it hard for some of us to think that a collection can be self-membered. Now, the axiom of extensionality is shared by ZF and NF. It is therefore less controversial at this stage to say that “only objects which conform to extensionality can be rightfully called sets” than to say the same with respect to the axiom of foundation. Rescher and Grim would agree with this. As we saw, plenary collectivities, *not being sets*, do not behave according to the axiom of extensionality. They don’t have to “carry that burden”.

¹ Lavine (1994) pp. 289-90.

² For references, see Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973) pp. 86 and 101-2.

³ Even Thomas Forster, a prominent contemporary defender of a “set theory with a universal set” admits that there is something “metaphysically unsatisfying” about the idea of self-membership, namely, that one seems to be unable to fulfill the Quinean requirement expressed by the phrase “No entity without identity!”. He proposes a recursive (and hence also “iterative”) method for identifying most self-membered sets, but he also admits that “some infinite [identification] plays have to be classed as wins for player Equal” (2008, p. 101) which, if I understand him correctly, means that in some cases, like that of the set of all sets, identity must simply be *stipulated*.

It is not altogether easy to see why the liberation from the yoke of the axiom of extensionality would elucidate the idea of a collection's self-membership. There is, however, one idea which does become automatically authorized, namely the idea that a collection can change its size, i.e. that it can either gain or lose elements, that it can grow or decrease. And plena are said to be growing, amorphous collectivities:

When viewed from the point of view of anything like sets, plena are *self-amplifying* and indeed explosive in content. Every time one looks to yet another sub-collectivity one automatically finds further members of the plenum itself. [...] Plena are so large that we can never bring them into view as a whole. Whatever one's first view, the plenum will contain more –ever adding elements corresponding to all elements of the power set of what is in view. Once one has included *those* in the plenum as well, it is clear that each sub-collectivity of that *larger* compendium will represent a further member of the plenum. The process continues and expands one's view ad infinitum.¹

We saw in section 4.1.2 that plena have a member corresponding to each one of their subcollectivities, and that when those members simply *are* all the subcollectivities what we have is a *membership*-plenum. This means that a membership plenum contains itself as a member. The fact that the mind boggles at this idea is unabashedly integrated into the theory:

Plena are quite literally mind-expanding, growing beyond any determinate conception at any time.²

The claim that plena “grow” is, as we said, allowed by the newly achieved liberation from the axiom of extensionality. It affords, moreover, an interesting connection with an idea at which set theory has been hinting since Cantorian times. The idea is that the philosophically interesting totalities, like the totality of everything, are non-set collectivities. As we just saw, this means amongst other things that these oversize totalities do not behave according to the axiom of extensionality. And this, in its turn, means that they are allowed to *grow*. Perhaps the passage in which Cantor came closest to the assertion of this idea is to be found in his famous letter to Dedekind, already alluded to in our previous chapter:

For a multiplicity can be such that the assumption that *all* of its elements ‘are together’ leads to a contradiction, so that it is impossible to conceive of the multiplicity as a unity, as ‘one finished thing’. Such multiplicities I call *absolutely infinite* or *inconsistent multiplicities*.

¹ Rescher and Grim (2011) p. 66.

² Rescher and Grim (2011) p. 67.

As we can readily see, the 'totality of everything thinkable', for example, is such a multiplicity; later still other examples will turn up.¹

There is a sense, of course, in which even "inconsistent multiplicities" must be treated as unities. If they are to be identified and differentiated from one another (e.g. the totality of sets should be differentiated from the collection of ordinals) we certainly must assume them to be unities. But Cantor is not really denying this. He explicitly says that he means the denial of unity to the inconsistent multiplicities in the sense that the elements of those multiplicities cannot be thought of as being "together", on pain of contradiction. Unity means here "one finished thing" and not "one thing" *simpliciter*. What Cantor is saying is not that inconsistent multiplicities are not, in an extremely abstract sense, "things"; what he is saying, rather, is that inconsistent multiplicities are not the kinds of things that sets are. And if inconsistent multiplicities are to be thought of as not being finished things, as not having gathered *all* its members, this means that inconsistent multiplicities are different from sets precisely in their being able to win (or lose) elements, i.e. in being able to change their size over time. Inconsistent multiplicities, that is to say, need not behave according to the axiom of extensionality. In this respect, inconsistent multiplicities and plenary totalities are similar. The plenary collectivities of Rescher and Grim, just as the proper classes of some systems of set theory, are direct heirs of Cantor's inconsistent multiplicities, just as the latter are, ironically, heirs of the Aristotelian conception of potential infinity.²

The claim that plenary collectivities grow is therefore not only allowed by the abandonment of the axiom of extensionality with respect to plena, but it also has a philosophically interesting ancestry. In our last chapter, we also said that we shouldn't worry about violating Cantor's Domain principle by admitting indefinite totalities. If the principle were universally applicable, there would have to be a *definite* collectivity corresponding to every plenum. But as soon as we enquired into the justification of the principle, all we came up with was that otherwise "each fixed support for the study would collapse"³, which makes the principle hardly distinguishable from the unrestricted –and

¹ Cantor [1899] "Letter to Dedekind" in van Heijenoort, J. (1967) p. 114.

² Rescher and Grim (2011) p. 107 do point out that Cantor's non-set collections and their plenary totalities are similar in some respects. As we said in the last chapter, amongst the authors who note the Aristotelian overtones of the Cantorian talk of inconsistent multiplicities and of some systems of set theory which allow proper classes, are A. Bar-Hillel, Y. and Levy, A. (1973) p. 118. and A.W. Moore (2001) p. xiv. The abandonment of the axiom of extensionality is more usually stressed, if at all pointed at, with respect to inconsistent multiplicities than with respect to proper classes. The reason for this is that systems of set theory with classes would either view classes as the same kind of entities as sets –only bigger–, or as corresponding to sets when they are small enough and not having a corresponding set when they are as big as the universe of sets. The former kind of system has stems from John von Neumann's and Gödel's work; the latter from Paul Bernays'. Either way, the definiteness of sets would seem to be a desideratum for at least *some* classes, which means that those classes are viewed as behaving according to the axiom of extensionality.

³ Cantor, as quoted by Hallett (1983) p. 25.

paradoxical- axiom of comprehension¹. We have, then, good reasons to suspect of the purported universal applicability of the principle. Moreover, since for Cantor some totalities were not definite, it is quite obvious that he didn't think of the principle as universally applicable either.

So let us take for granted that there is in principle no problem with the idea that plena are ever-growing or, as Rescher and Grim prefer to put it, *self-amplifying* collectivities. There is one reason, however, why we should be cautious with this claim, at least if we limit ourselves to the reasons provided in support of it by Rescher and Grim. A plenum can be said to grow because it is said to amplify itself. But we simply do not seem to be entitled to say that a plenum amplifies itself. According to the definitions provided by Rescher and Grim, every collection is a subcollection of itself and a membership plenum contains each of its subcollections as elements. From this it does immediately follow that a membership plenum contains itself as a member. It also follows that plena are ungraspably large, since if we "have in view" a given size for the plenum (a given cardinal number for its elements), that view *must* be a partial one: corresponding to *any definite size* there is the larger size of the Power-collectivity, but plena are said to contain elements corresponding to (maybe identical with) every element of the Power-collectivity. And if we bring, as we should, the latter elements into a new and enlarged conception of the size of the plenum, we must immediately, by the definition of a plenum, think of the plenum as having elements corresponding to every element of the Power-collectivity of what we just said we had in our *enlarged* view. So on ad infinitum. This is the meaning of the attribution of "explosiveness in content" to plena. But it does not follow from any of these ideas that the *plenum itself* has thereby gained a single new element. The latter would only follow if the size of a plenum is correlative to the conception that we have of that size "at any given time." But this is incompatible with metaphysical realism regarding plena, which Rescher and Grim endorse.² Thus, that plena are "mind-expanding" does not imply that they "grow" or that they are "self-amplifying". What is doomed to explosion is any given conception of the size of a plenum that we happen to have at any given time, not the size of the plenum itself.

It must be pointed out that this difficulty (the apparent incompatibility with realism regarding mathematical objects) is not only faced by plenum theory, but also by any conception of the totality of things as ever-growing. The difficulty is more prominent in the case of plena because they are said to be self-amplifying and the only reason provided for this claim is that every *conception* of the size of a plenum is bound to be partial and doomed to explosion, but, as we just pointed out, the

¹ And, as we saw, Graham Priest (2002, pp. 281-2) represents understands both principles in this connection. See section 2.4 of the previous chapter in this investigation.

² Rescher and Grim (2011) p. 66.

explosiveness of any given conception of the size of a collection does not imply that the *collection* in question acquires a single new element. Now, a parallel criticism can be made against any view according to which the universe is ever-growing (whether “self-amplifying” or not), at least when the reason for this claim is that we always have the means for defining new elements of the collection in question.¹ If mathematical objects are real independently of our conceptual capacities, the multiplicity which is constituted by them does not increase whenever we realize that a given conception of the size of that multiplicity is only partial. The objection is important because, as we saw, one way in which Cantor’s distinction between consistent and inconsistent multiplicities can be interpreted, a way which emphasizes Cantor’s claim that their difference lies in that inconsistent multiplicities cannot be *thought of* as “finished things” on pain of contradiction, is to think of inconsistent multiplicities as potentially infinite. And fits very well, and suggests, a conception of inconsistent multiplicities as changing their size (growing). But if the only reason for this claim is that we can always form a more comprehensive conception of the size of those totalities, the argument conflicts with realism.²

Let us take stock. We have seen that plena, not being subject to the axiom of extensionality, can consistently be said to change their size, and therefore to grow. This is indeed an important difference between plena and sets and, moreover, it can be related to the idea of the Universe of sets as a non-set collectivity -an idea which dates back to Cantor-. Still, by conceiving the Universe as a *collection*, Plenum Theory inherits rather than solves one of set theory’s problems. Being a collection of all collections, the collection of everything must contain itself as a member. Now, instead of elucidating this latter idea, plenum theory seems to make it more difficult to understand. Self-membership is (sometimes) implied by the characteristic mark of plena -that they have a member corresponding to each member of their Power-collectivity (the implication occurring when the correspondence is identity)-. This characteristic mark also implies “explosiveness in content”, i.e. that no definite conception of a plenum’s size is achievable. But that plena are self-membered and that our idea or conception of them is explosive in content does not imply that plena are self-

¹ In the previous chapter, I defended the conception of the universe as ever-growing and potentially infinite against the charge that it violates Cantor’s domain principle; but I also pointed out (p. 65 footnote 1) that there is a remaining difficulty, namely the present one.

² This is certainly not the only way to understand Cantor’s distinction, for one can also understand it as a radical abandonment of reificationism regarding inconsistent multiplicities: they are not to be thought of as individuals at all (whether potential or actual infinities), they are genuine multiplicities, “manies” and not “ones”. I discussed this view in the previous chapter (section 2.3), and will say a little more about it in the following one (section 4.3).

amplificatory or that they grow. To think that this implication obtains is to incur in a repudiated subjectivism. Not that Rescher and Grim say that the implication obtains. But, to my knowledge, they do not provide any (other) reasons for the claim that plena are self-amplificatory either. Thus, either they believe that plena are self-amplificatory for unspecified reasons, or else when they say that plena are self-amplificatory they simply mean that any given *conception* of plena is explosive in content. Either way, Rescher and Grim fail to elucidate the claim that a plenum is a collection which has itself amongst its members.

In the end, self-membership seems to be as mind-boggling in the case of plena as it is in the case of sets.

Conclusions

Two ideas have been defended in this chapter. The first of them is that there are at least two ways of saving the concept of *Everything* from the paradoxes of set theory. These two ways are developed in two corresponding formal systems that deal with collections. These systems are Quine's New Foundations (an alternative set theory) and Rescher's and Grim's Plenum Theory (an alternative to set theory).

That a concept is free from paradox does not mean, however, that we understand it clearly. This is the second idea that I tried to defend in this chapter. In particular, since both NF and Plenum Theory conceptualize Everything as a collection, it is a collection which must contain itself. It must be an element, a member of itself. Now, some minds perhaps do not boggle at this idea. Mine does. At least as much as it does e.g. at the idea that language speaks.

Chapter IV. Kant, Cantor, and the Unconditioned

So far I have dealt with what look like two disconnected problems: Kant's first antinomy in chapter I, and the so-called paradoxes of set theory in chapters II and III. In fact, there is at least one connection between them, namely that by taking set theory into account we can see clearly that Kant's alleged antinomy is not an antinomy at all. The argument for the thesis of the antinomy, as we saw, presupposes that "the true concept of infinity" is that an infinite collection cannot be completed, which is just another way to express the traditional Aristotelian belief that infinity can only be potential as opposed to actual. But set theory's starting point is precisely the assumption that infinite collections can be actual, and its success and explanatory power suggest that it is a correct account of infinity. If this is so, the only connection that seems to obtain between Kant and set theory is that the latter can be used to show that Kant's alleged antinomy is not an antinomy at all.¹ True, there are paradoxes that threaten set theory, at least when it is practiced under certain presuppositions, but, if correctly interpreted, they are *not* paradoxes of infinity. They threaten our conceptions of some infinite collections, but not of actual infinity as such. We seem to be able reject the assumption of the existence of the problematic collections, or think of them as collections beyond mathematical treatment, while still holding on to the claim that those infinities with which Kant was concerned are actual infinities, and hence Kant's problem not a problem at all.

This dismissive attitude, however, is not very interesting, and if what I will defend in this chapter is correct, it is only a very partial way to tell the tale. If there is one philosophical problem that set theory leaves unsolved it is whether we can think of the multiplicity of everything as an individual, i.e. as an object, unity or thing. Not only that. If by "thing" we mean *definite* thing, the paradoxes suggest that we cannot. And this is a problem, if not for all of us all the time, at least for the Parmenides within us all. Now, there are elements in the *Critique of Pure Reason* that allow us to characterize the first antinomy as conflict generated by the concept of an entity way more comprehensive than that of the spatiotemporal world. Kant says that the main task of metaphysics is to think of *the unconditioned*, and that the antinomies show that the transcendental realist cannot

¹ This attitude towards Kant's antinomies can be found in Russell [1914] (1926) p. 161, Bennett (1974) p. 119 and Hallett (1983) p. 226-38.

think of the unconditioned without contradiction (BXIX-XX). He failed to derive a paradox with respect to the unconditioned that he had in mind, we know that. But what about –if the barbarism be permitted- a *more radical unconditioned* than the one Kant had in mind? Can we not characterize the paradoxes of set theory as problems for *this* radicalized unconditioned? I think we can, and I will try to make this point in the present chapter. Whether the problems that beset the concept of an unconditioned totality of everything can be used to defend a dualism akin to Kant’s transcendental idealism (in analogy to Kant’s using the antinomies as a *reductio ad absurdum* against transcendental realism) is not a problem to be dealt with in this chapter, nor indeed in this whole investigation. I will however, point at an answer in the concluding remarks. I think they can’t.

I will divide the chapter in four sections. In the first one I present both the concept of a set of all sets and that of a set of all facts as radicalized versions of Kant’s concept of the world. In the second section I endorse the idea, defended by Gottfried Martin and Graham Priest, that there are structural similarities between Kantian and set-theoretical paradoxes (namely that the same tendencies of human thought seem to be generating the two conflicts). In the third section I present a paradox that arises if we take Cantor’s Domain principle to be universally applicable. I call this paradox the paradox of Everything. In the fourth section I go back to the *Critique of Pure Reason* to emphasize Kant’s claim that the antinomies are generated by the idea of the unconditioned, and argue that the Everything that generates our paradox can be seen as the Kantian Unconditioned.

4.1. A Totality of What?

Let us start by looking more carefully at Kant’s own definition of the concept of the *world*. This concept, he says,

[...] signifies the mathematical whole of all appearances and the totality of their synthesis in the great as well as in the small, i.e., in their progress through composition as well as through division. [... In using this concept, one] look[s] at the aggregation in space or time so as to bring about a quantity [...]. (A418-9/B446-7).

In order to avoid confusions, the first thing to note here is that the word “appearances” is not and cannot be used in the idealist way, as opposed to “thing in itself”. On the contrary, since Kant’s intention is to derive a contradiction from this concept of the world, and thereby to prove that the

transcendental realist must necessarily face an antinomy, the concept of the world must not only be neutral with respect to the realist-idealist controversy, but it must be rather particularly appealing to the realist. The phrase “the mathematical whole of appearances” must then mean something like *the mathematical whole of spatiotemporal objects and events*. In fact, when one looks at the proofs of the First Antinomy, one realizes that this is exactly what Kant means. The temporal side of the Antinomy is a conflict between the claim that there must be an absolute *beginning* of the world and the claim that there cannot be such a thing. But here the beginning is quite clearly understood as a first *event*, as we saw. And when one looks at the proofs in the spatial side of the Antinomy (A426/B454; A427/B455), it is easy to see that what is under discussion is whether there can be a limit to the totality of space-fulfilling *objects*. “Appearances”, then, must in this context be Kant’s word for *spatiotemporal objects and events*. Now, we are asked to frame the concept of the “mathematical whole” constituted by appearances thus understood. That is to say, Kant asks us to look at the world as the *compositum* or aggregate constituted by these objects and events when they are *added* to one another¹; but what we want is the whole of it, not just one of its parts. Kant’s concept of the *world* is then the concept of *the totality of spatiotemporal objects and events*. Let us take this for granted in what follows.

Kant’s concept of the *world* does not seem problematical at all. It is not, at any rate, beset by the problems Kant thought it was. This doesn’t mean, however, that there is not an antinomy generated by *a* –if not *Kant’s*– concept of the world. The word “world” is perhaps infelicitous in this regard. In ordinary language, it usually means the planet Earth, and this is definitely not at issue, neither in Kant’s antinomy nor in the discussion that follows. On the other hand, Kant was concerned with the spatiotemporal *Universe*, but we saw that he fails to show that an antinomy is generated by this concept. What I want to suggest in what follows is that we do have good reasons to think that a certain concept of the world does generate antinomies. This self-contradictory concept of the world must be, however, far more general than that of the spatiotemporal Universe, let alone that of the planet Earth. If we take away the qualification of *spatiotemporality* from Kant’s concept of the world, and ignore (for the moment) the reference to *events*, we arrive at a concept of the world as *the totality of objects*. If we stick to the word “world” we would have to say that this is the world *simpliciter*, not the spatiotemporal world (Kant’s world). Does this concept, i.e. the concept of the world as *the totality of objects*, generate any antinomy?

¹ Kant also asks us to look at this *compositum* –the world– as regards the division of its parts, but this is the subject matter of his Second Antinomy and it can be ignored for our present purposes.

The findings of chapter II strongly suggest that it does. We saw there that the concept of the set of all sets generated paradoxes under certain presuppositions. Now a set of all sets would be precisely the *Universe* of set theory –the universal set. And, the way standard set theory conceptualizes its objects, the universal set is just the totality of objects. To see this, recall that when the standard system of set theory admits objects which are not sets themselves (i.e. when it admits memberless elements: individuals, *Urelemente*), these objects are at the “bottom” layer of the hierarchy of sets.¹ That is to say, the system admits first all individuals (objects which are not sets), then it countenances sets of individuals, then sets of sets of individuals, then sets of sets of sets of individuals, and so on. So if the system admits individuals at all, individuals and sets are the only entities countenanced by the theory. Now, individuals get collected at the very first stage of set-construction. To distinguish between these stages, numbers are assigned to them, and the number assigned to each stage or layer is called its *rank*. The bottom layer is not really a stage of set-construction, since there are no sets at all in this layer but only individuals. Thus, the rank of the bottom layer is 0. Then, in the first stage of set construction, one gets precisely the first kind of sets: sets of individuals. The rank of this layer is 1. (Here one also gets, by the way, the set with no elements: the empty set). In the next stage of set-construction, one gets not only all the sets that one had before, but also all the sets that can be made out of the elements of the first layer. That is to say, the layer of rank 2 contains not only individuals (from rank 0), the null-set and all sets of individuals (from rank 1), but also sets of sets of individuals, like the Power-set of any of the sets of layer 1. The layer with rank 3 contains all the elements of previous layers as well as all the sets that can be made thereof. And so on. This structure of layers is called the *cumulative hierarchy of sets*.²

It should be clear from these elementary aspects of standard set theory that the set of all sets, if there were such a thing, would be precisely the totality of *all* the objects countenanced by the theory. But there is no set of all sets. Thus there is no totality of objects. And if the concept of the world refers to an alleged totality which would comprehend all the objects there are, then the paradoxes of set theory strongly suggest that there is no such thing as the world. In standard set theory there are sets but not a set of all sets. There are objects but there is not an object which contains all objects there are.³

¹ I am leaving aside systems which do not admit individuals, i.e. *pure* set theories. I don't deal with these systems because in them my main point is rather trivially correct. If the theory countenances only sets as objects, then the set of all sets is simply the set of all objects, and if the set of all sets is paradoxical, so is the set of all objects.

² See Potter, M. (2004) pp. 40-2.

³ That the reality of *sets* is compatible with the rejection of a *set of all sets* and other paradoxical entities has been pointed out, among others, by Michael Hallett. Cf. his (1983) p. 225: “One way of analyzing the set-

Does this result have anything to do with Kant's First Antinomy? The answer to this question is to some extent a matter of interpretation. Russell, for example, was perfectly aware of both facts we've been concerned with so far: that Kant tried unsuccessfully to derive a contradiction from the concept of the spatiotemporal world, and that the concept of the set of all sets generates some paradoxes. And in view of the paradoxes, Russell eventually came to believe that there is no collection of all collections: that there is no such thing as 'Everything'.¹ But Russell didn't draw a connection between Kant's and Cantor's problem. He even thought that Cantor's set theory had finally solved age-old puzzles which arose out of confusions shared by a philosophical tradition stretching from Zeno to Bergson. He saw Kant's "Antinomies of Pure Reason" under this light.²

Strictly speaking, what Russell says is right. Strictly speaking, it is not the same concept which is said by Kant and Cantor to generate an antinomy. The concept of the spatiotemporal world and the concept of the set of all sets are by no means identical. And there is strong support for the claim that Cantorian set theory enables us to say that Kant's First Antinomy is no antinomy at all. We already saw in chapter I that the Antinomy is not genuine anyway, not even in its own terms, since, for example, the proof for the Thesis clearly assumes what it should demonstrate. But in case one had inadvertently granted Kant's assumption; in case one had carelessly allowed Kant to assume that a completed infinity is something of a contradiction in terms, then Russell's point is essentially correct: Cantorian set theory shows that it is not obvious that a completed infinity is a contradiction in terms. And it is also true that certain claims, which a long-standing philosophical tradition would consider plain absurdities that follow from the assumption of the possibility of completed infinities, were made by Cantor the very foundations of his theory of sets. Traditionally, arguments such as the following were not uncommon: "If there were completed infinities, it would be in principle possible to measure them, add them to one another, say that some of them are bigger than others, that some of them are the same size, etc." These counterintuitive consequences were reason enough for the Aristotelian tradition to reject the possibility of there being completed infinities. No doubt, the proof of Kant's Thesis-argument belongs to this tradition. An *elapsed-eternity-up-to-any-given-moment* would be an iterable eternity, an infinity which could be augmented, and this is one of the

theoretic contradictions is by attributing them to naïve mathematical realism. The comprehension principle bids us to treat any collection as an individual object of the same kind as the elements which go to make it up, and thus as an object which must either belong or not belong to any object, etc. If this naïve position is adopted, the contradiction ensues as a matter of logical course. One need not as a consequence of this abandon set-theoretic realism, since systems free from the known contradictions can be construed which embrace the principle of bivalence, and blithely accept completed infinities and impredicative definitions, all or any of which might be taken as indicative of realism. But the unchallengeable nature of the logical step from comprehension to contradiction must force even a mathematical realist to abandon his naivety."

¹ Cf. Russell, B. [1919] p. 136.

² See chapter VI of his [1914].

absurd-seeming consequences of the acceptance of completed infinities. But in the previous chapter we saw that the counterintuitiveness of certain results of the assumption of the existence of completed infinities, e.g. the fact that the set of natural numbers would have to be seen as containing the same number of elements as the set of even numbers, does not necessarily mean that those consequences are *absurd*. Cantor showed that under non-traditional assumptions regarding the criteria for size comparison, a theory of sets as completed infinities which are *mathematically manageable* is not only conceivable but extremely promising as an attempt to provide the logical foundations for mathematics. True, Cantorian set theory faces serious difficulties under certain naïve presuppositions. But they are *not* difficulties with the concept of a completed infinity. They are difficulties with the concept of the set of all sets, or with the set of all ordinal numbers, or with the Russellian set of all sets that do not belong to themselves. But, as the development of standard set theory in the twentieth century shows, those difficulties can be coped with by rejecting both the naïve presuppositions which led to the paradoxes and the existence of the paradoxical sets. And there we have it: Zermelo-Fraenkel set theory. In the light of such a theory, Kant's claim that the concept of a completed infinity is a self-contradictory one, that the "true" concept of the infinity of a series is that according to which that series cannot be completed, does reveal itself as the heir of a mistaken tradition. This is what Russell correctly points out.

Having said that, certain connections between the Kantian and the Cantorian concerns can't be ignored. I cannot ignore them at any rate. We saw that a set of all sets is, for the set theorist, the set of all objects there are. And at least a plausible conception of the world is that according to which it (the world) is the totality of all objects there are. That Kant failed to derive an antinomy from the concept of the set of all spatiotemporal objects doesn't mean that there is not an antinomy generated by the concept of the set of all objects *simpliciter*. The concept of the spatiotemporal world is an unproblematic concept; but could we say the same with respect to the concept of the world as the totality of objects?

Not without hesitation, it seems. A.W. Moore shows how a set theoretical antinomy could be derived from this concept. He uses the concept of the universe, but he clearly uses that concept in the sense that we are using the concept of the world. The antinomy goes as follows. Thesis: The Universe belongs to itself. Proof: If it didn't, it would not be the set of *all* sets, for there would be at least one set –the Universe- which it did not contain. Antithesis: The Universe does not belong to

itself. Proof: If it did, it would not be the *set* of all sets; for a set contains only sets of lower ranks, and so no set can contain itself.¹

Later on in this chapter we will search for elements in the *Critique of Pure Reason* which will allow the recognition of Kant's concerns in the difficulties faced by Cantorian set theory. But for the moment let us take this point as established: If the world is the totality of *all* objects, and this totality is conceived of as a set in the sense of standard set theory, then we have good reasons to think that this is a self-contradictory concept, namely the existence of the set-theoretic paradoxes. And since the concept of the world as the totality of objects is quite clearly a generalization of Kant's concept of the world, the set theoretic paradoxes can be seen as a more radical version of Kant's problem. Moreover, given that the set-theoretical antinomies are more difficult to dismiss than the Kantian antinomy, it seems that it is precisely in the radicalization of the concept of the world that the key to the genuineness of the antinomy resides. Kant's concept of the world is *not* problematic. But a much more radical version of it *is*.

"But the world is the totality of facts, not of things", you might want to say at this point, quoting the *Tractatus*.² If this is so, then the existence of the set theoretic paradoxes generated by the concept of the set of all sets seems to leave the concept of the world untouched. This in its turn would mean that the attempt to resurrect the Kantian concerns in the Antinomies by means of a specific way of reading the set theoretic paradoxes is hopeless. To be sure, if our ontology is liberal enough to allow not only "individuals" but also abstract objects like sets as constituents of the furniture of the world, then the problematical nature of the concept of the set of all sets clearly threatens the concept of the *world*; but why should we care about this, if the correct conception of the world is not that according to which it is the totality of objects ("things"), but the totality of facts?

I won't discuss the question whether the world should be conceived of as the totality of facts rather than as the totality of things. What I want to do instead is to focus on an ingenious argument to the effect that Cantorian difficulties arise even when the world is conceived of as the totality of

¹ Cf. Moore, A.W. (1988) p. 217. As can be seen, the proof of the antithesis of this antinomy rests on the presupposition that a set cannot be a member of itself. This is the presupposition which is questioned by non-standard set theories. We have already dealt with those non-standard set theories in chapter IV and we said that there is something unsatisfactory about an idea they must incorporate: self-membership. I will briefly come back to the issue of self-membership at the end of this chapter.

² Wittgenstein, L. [1922]. 1.1.

facts (or of states of affairs, situations, etc.), at least as long as “totality” here is understood as a set in the sense of standard set theory.¹

Let us first point out that the concept of the world as the totality of facts or of states of affairs is by all lights also a radicalization of Kant’s concept of the world. We saw that the temporal side of the first Antinomy is the conflict between the claim that the world had a *beginning* in time and the opposite claim that it didn’t have a beginning. And a beginning in this sense is clearly a first *event* in time: the first thing that happens. A first fact. So the first antinomy, had it been genuine, would have also been a problem for those who think that the world is not the totality of spatiotemporal objects but of spatiotemporal facts (assuming that there had been valid arguments in defense of the antithesis as well). The antinomy was not genuine, of course, and thus the concept of the totality of spatiotemporal facts seems to be quite an unproblematic fellow. But can we say the same about its radical cousin, i.e. about the Wittgensteinian concept of the world as the totality of facts *simpliciter*?

The answer is, again, that if “totality” here means a Cantorian set, then the concept of the totality of facts is as paradoxical as that of the totality of things. Now, this understanding of “totality” as a Cantorian set in this context is not so easily negligible, for it is precisely the understanding suggested by the way in which both actual and possible worlds are characterized in some theories of modality. Consider, for example, the way Robert Adams characterizes a possible world in terms of its correlate “world-story”:

Let us say that a *world-story* is a maximal consistent set of propositions. That is, it is a set which has as its members one member of every pair of mutually contradictory propositions, and which is such that it is possible that all of its members be true together. The notion of a possible world can be given a contextual analysis in terms of world-stories. [...] In accordance with this analysis, we can say that the actual world differs from the other possible worlds in that all the members of its world-story (the set of all the propositions that are true in it) are true, whereas the stories of all the other possible worlds have false propositions among their members.²

As we can see, a possible world is understood here as the correlate of a set of propositions (i.e., of its world-story). And the actual world is the correlate of all true propositions. That is to say, the actual world is the totality of facts, of what is actually the case. “The world is all that is the case”, in

¹ The argument, which is due to Russell’s [1903] p. 527, appears with enormous clarity and conciseness in Bringsjord, S. (1985) p. 64., and it is also to be found in Grim’s (1984) pp. 206-8 and also in his (1991) pp. 91-4. Bringsjord also credits Martin Davis (*Meaning, Quantification and Necessity*, Routledge and Kegan Paul, 1981, Appendix 9) for an earlier version of this argument.

² Adams, R. [1974] (1979) p. 204.

accordance with the opening line of the *Tractatus*. Now, being a correlate of a set of propositions, any possible world, including of course the actual one, must also be a *set*, and quite a Cantorian one at that. To see this, note that both the world-story and its corresponding world must behave in accordance with the axiom of extensionality (sets are identical if and only if they have the same members). The actual world, in which you are engaged in logico-philosophical considerations right now, is different from the possible world in which you aren't; and this must be reflected in their correlative world-stories. Thus, the identity of worlds and world-stories in this conception is clearly determined by the identity of their members. This suggests that the (actual) world is conceived of as a *Cantorian set* of (actual) states of affairs.¹

If this is so, then we might accept the following definition of the world, proposed by Selmer Bringsjord:

w is a world =_{df.} *w* is a set of states of affairs such that (i) for every state of affairs *p* either $p \in w$ or $\neg p \in w$; and (ii) the members of *w* are compossible.²

The first condition is the so-called *maximality* condition. If we are talking about a given possible world *W*, it means that *W* contains one member from each pair of incompatible states of affairs. If we are talking about the actual world, it means that it contains *every* actual state of affairs, every fact. This idea, of course, also has its *tractarian* origins: "The world is determined by the facts, and by their being *all* the facts."³ The second condition means that there are not only atomic states of affairs but also those composed by two or more states of affairs.

Now, as I said before, and as Bringsjord himself goes on to point out, there is a powerful argument to the effect that this concept of the world as the maximal totality of facts is untenable. The argument is the next one. Suppose that there is such a maximal totality of facts. That is to say, suppose that the world (let us call it *W*) is the set of all facts. It contains Fact 1, Fact 2, etc., and it looks like this:

W: {*F*₁, *F*₂, *F*₃, *F*₄, *F*₅, ..., *F*₃₆₇₈₉₄₇₅₆₇₈₃,..., *F*_{*n*},...}

¹ Elaborating on Adams's ideas, Alvin Plantinga gives an account of possible worlds as states of affairs, characterizing the latter as totalities which have an inclusion relation which is a mixture of the Cantorian relations of inclusion and membership. Cf. Plantinga, A. [1976] (1979) pp. 258-9.

² (1985) p. 64.

³ Wittgenstein, L. [1922]. 1.11.

If it is a Cantorian set, then there is a Power-set to W , i.e. PW . This Power-set of W has as elements all the subsets of W , i.e. all the sets which have some (but not necessarily all) the members of W . So PW will look something like this:

$PW: \{\emptyset, \{F_1\}, \{F_2\}, \{F_3\}, \dots, \{F_1, F_2\}, \{F_2, F_3\}, \{F_5, F_{16}\}, \dots, \{F_2, F_{254}, F_{6750}, F_{98879}\}, \dots, W, \dots\}$

Now, we saw in chapter II that by Cantor's Theorem, PW must have more members than W . Under the presupposition that there is a one-to-one correspondence between W and PW a contradiction arises, and this was the *reductio ad absurdum* proof of Cantor's Theorem. To use the language of set theory, we would say that according to Cantor's Theorem, given any set α , there is a set with larger cardinality than α , namely $P\alpha$. So PW has a larger cardinality than W .

So far, the larger cardinality of PW with respect to W does not seem so obviously threatening to the existence of W . The world, W , is the totality of facts; it is true that logic forces us to admit that its Power-set PW has more members, but why should we care about this, if PW , unlike W , is *not* a set of *facts* but a set of *sets of facts*?

It is in the answer to this question that the strength of the present argument resides. The answer is that to each element of PW there corresponds at least one fact. For example, that it contains exactly the elements that it does. Or that it does not contain those elements that it does not contain. Or that it –exactly it– is an element of PW . Or that it –and it exactly– is not a member of all the sets of which it is not a member. Any of these facts will do. Thus, to every member of PW there corresponds (at least) one fact. But PW has, by Cantor's Theorem, more members than W . Thus, the set of facts corresponding to PW will have more members than W , which was supposed to be the set of *all* facts! Contradiction.¹

We must take stock before we proceed to the next section of this chapter. We saw that Kant's concept of the world is that according to which the world is the totality of spatiotemporal objects and events. We saw as well that despite the fact that Kant failed to derive a contradiction from this

¹ Grim and Bringsjord correctly point out that this argument is not only a threat to the concept of the *actual* world, but also to that of a *possible* world. For a possible world is here conceived of as the correlate of a maximal set of propositions. That is to say, a (possible) world is supposed to be a *maximal* set of (actual or merely possible) states of affairs. But given that the facts corresponding to the elements of the Power-set of any possible W would be actual states of affairs, and that according to Cantor's Theorem some of those facts cannot be included in the possible W in question, the truth of Cantor's Theorem implies that W is not maximal, i.e. that there is no such maximal set. Therefore, if possible worlds are conceived of in this Cantorian fashion, no possible and a fortiori no actual world is possible at all.

concept, some contradictions can be derived from a radicalized version of it, namely that according to which the world is the totality of objects simpliciter, or of facts simpliciter, at least as long as “totality” here means Cantorian set and “object” here means also sets. An all-comprehensive totality in this sense is not to be allowed, since to any given set there is a set with more members, according to Cantor’s Theorem. Were we, on the other hand, to conceive of the world not as the totality of objects but as the totality of facts, a contradiction can be derived from this notion anyway. For the Power-set of the set of all facts must have more elements than the set of all facts (Cantor’s Theorem), and since there would be a fact corresponding to each one of the members of the Power-set of the set of all facts, the set of facts about the elements of this Power-set would be a set with more facts than the supposed set of all facts, which is an evident contradiction.

I think that there is a lesson to be learned from the considerations of this section. The set theoretic paradoxes are generated by a concept which is a *radicalized* version of Kant’s concept of the world. The concept of the world as *a certain totality* (viz. a spatiotemporal totality) appears to be quite unproblematic. By comparison, the concept of the world as *the* totality of totalities, that is to say, as an *absolutely unconditioned* totality, seems to be precisely what is getting us into trouble.

4.2. Totalisation and Beyond

The view that there is a philosophically significant connection between the Kantian Antinomies and the paradoxes of set theory has already been advanced in the literature, and in what follows I would like to provide further interpretative support for it. It has been rightly pointed out that there are *structural* similarities in the *generation* of both Kantian and Cantorian antinomies; that is to say, that the arguments which prop up the Theses and the Antitheses of Kantian and Cantorian antinomies can be seen as expressions of the same tendencies of human thought. In this section I will emphasize this structural connection, and in the following one I will present a paradox that arises if we take Cantor’s Domain principle to be universally –unrestrictedly– applicable.

Let us first look at the suggestion that there is a structural similarity. We have seen that A.W. Moore presents an antinomy with respect to the concept of the set of all sets. It belongs to itself (for it is a set) and it does not belong to itself (for no set can contain itself as a member). In his *Kant’s Metaphysics and Theory of Science*, Gottfried Martin had already presented a very similar antinomy. Suppose that there is a set of all sets U . Then I can show you that U is not the set of all sets, since it

fails to contain at least one set, namely the set that is formed by taking any element s_x of U and U itself. This newly formed set $\{U, s_x\}$ cannot be contained in U , for how can U contain a set which can only be formed once U –and *a fortiori* the whole of U – is there?¹ The antinomy is then that U is and is not the set of all sets.² Martin goes on to say:

We are thus concerned with the concept of totality. I consider all sets and form the concept of their totality, namely the set of all sets. But then I regard this totality as an element in the formation of a new set. A concept of a totality is again used as an element. This conflict between concluding and beginning anew, between forming a totality and using this totality as a new element, is the actual ground of the [set-theoretic] antinomy. It is this conflict that gives the connection with the Kantian antinomies. Kant saw quite clearly that the antinomies rest on this antithesis between making a conclusion and going beyond the conclusion.³

To say it in a more Kantian way: both antinomies are generated by the same tendencies of reason. These are: *totalization* (Thesis) and *going beyond any totalization* (Antithesis). Elaborating on these ideas, Graham Priest calls the thetical tendency *Closure* and the antithetical one *Transcendence*. Priest's claim that both tendencies are as essential in the generation of the Cantorian paradoxes as they are in the generation of the Kantian ones is perhaps best understood if we look at two further paradoxes, the so-called Burali-Forti paradox and a paradox which emerges when we think of the *cumulative hierarchy of sets* as being itself a set.

To understand the Burali-Forti paradox one has to understand what is an ordinal number. To understand this, we shall start by pointing out that when we consider a set, we might want to focus not on its size or cardinality (the number of elements it contains) but on the way in which those elements are *ordered*. Now, not every kind of order will be relevant for our purposes; the crucial notion is that of a *well-ordering*. Not every ordering is a well-ordering. An ordering of a set is a well-ordering when, provided that the set in question is not the empty set, then it (the ordering) singles out one element of the set as its *first* element and, given any element in the set that has already been singled out, the ordering singles out *another* element of the set (if there is such another element) as its *immediate successor*. This sounds pretty complicated but it is not. Think of the following series, i.e. of the following ordered set:

¹ [1951] (1955) p. 54.

² Ibidem. This is also the page where Martin gives us the relic: "The connection [of the set-theoretic paradoxes] with the Kantian antinomies has often been noticed and I myself am greatly indebted to Ernst Zermelo for drawing attention to it in his lectures."

³ Ibid. p. 55. I slightly modified P.G. Lucas's translation in that I use "set" instead of "class". Martin's word is "Menge", i.e. the one that is usually translated in the Cantorian literature as "set".

$\langle \dots, -3, -2, -1, 0, 1, 2, 3, \dots \rangle$

This set is ordered but *not* well-ordered. True, for any member of the set, there is one which is its immediate successor, but there is no element singled out as the first member of the set. So it cannot be a well-ordered set. Or, again, consider the series:

$\langle 0, \dots, \frac{1}{4}, \dots, \frac{1}{2}, \dots, 1, \dots, 1\frac{1}{2}, \dots, 2, \dots \rangle$

This set is not well-ordered. The ordering of this set does single out one exact element as the very first element of the set, namely 0, but it doesn't single out an immediate successor of 0 -nor does it single out the successor of any of the other elements at that. On the other hand, the series of natural numbers, or any part of it in which the elements are arranged in the usual way,

$\langle 0, 1, 2, 3, 4, 5 \rangle$ or $\langle 0, 1, 2, \dots \rangle$ or $\langle 23, 24, 25 \rangle$, $\langle 98, 99, 100, \dots \rangle$ etc.

are well-ordered sets because the defining conditions are fulfilled.

Ordinal numbers are designed to measure the length of well-ordered sets. But note that *length* is then not the same as *size*. Size has to do with cardinality, length with order. Two sets of the same size can have different length. Consider the series:

$\langle 0, 1, 2, \dots \rangle$

And

$\langle 1, 2, \dots, 0 \rangle$

The first is N, the series of natural numbers. The second is the result of removing 0 from its privileged position in N as the first element of the set, and putting it instead as the first element to succeed all the remaining natural numbers.¹ Now, since there is a one-to-one correlation between the elements of the first set and the elements of the second (namely, the relation of identity, that relation which each thing has to itself), then by the Cantorian criterion of size comparison these two

¹ The claim that there is an immediate successor to an infinite series, i.e. an element which stands *immediately after the whole infinite series*, might strike us as odd. But in fact it is not an obviously contradictory claim. That there is an object x which is the immediate successor of a series does not necessarily mean that x is the immediate successor of a particular element (namely, the last element) of the series. If the series is infinite, there simply is no last element. The idea is then that if the series is infinite and x is the immediate successor of the series, x is the first object that comes after the *whole* series without being the first object that comes after any particular element *within* the series. Under the presupposition of the coherence of the idea of actual or completed infinities, there is no obvious logical problem with the idea that an object stands immediately after an infinite series.

sets have the same number of elements. They have the same cardinality, which Cantor famously symbolizes ' $\aleph_0$ ', Aleph-naught. But since the second series is slightly longer than the first one, for the first part of the second series is just as long as the first series but the second series has one element succeeding all the elements of its first part, then the lengths of the two series must be different. That is to say, these sets have a different well-ordering. Now, since the function of the ordinal numbers is to measure the length of well-ordered sets, a different ordinal number must be assigned to these two sets, even though the same cardinal number, viz. $\aleph_0$, will be assigned to them. The ordinal number assigned to the first set is ω ; that assigned to the second one is $\omega+1$. Or take the further example of the set:

$\langle 0, 2, 4, \dots, 1, 3, 5, \dots \rangle$

In this well-ordered set we first have all the even numbers and then all the odd numbers. But given the existence of the correlation between this set and the set of natural numbers, the cardinal number assigned to this set is again $\aleph_0$. Now, since we have an infinite sequence followed by another infinite sequence, the ordinal number assigned to this set is going to be $\omega+\omega$.

For our purposes, the crucial point is that ordinal numbers can deliver their function, i.e. to unequivocally measure the length of any well-ordered set, because they are supposed to be well-ordered *themselves*. That is to say, there is a first ordinal number, and given any ordinal number, there is another ordinal number which is its immediate successor. And if this is so no ordinal number precedes itself.

The Burali-Forti paradox is the next one. Consider the set of all ordinals, Ω . This class is well-ordered, so there must be an ordinal number which both measures the length of Ω and is the first to succeed Ω . But since Ω is the class of *all* ordinals the new ordinal must be in Ω and thus precede itself. But this is impossible since, as we said, no ordinal precedes itself.¹

We will shortly go back to the question of the relation between Cantorian and Kantian antinomies, but let us briefly turn our attention to a paradox regarding the *cumulative hierarchy* of sets. We saw that all sets in it contain elements of lower ranks. That is to say, sets contain *only* elements which were formed at *previous* stages of set-construction. The sets of the first stage of set-construction, if they are different from the empty set, are sets of individuals and contain *only individuals –not sets–*. In the next stage we get sets which can contain as members individuals and

¹ Cf. Giaquinto, M. (2002) pp. 37-40 and A.W. Moore's (2001) p. 125. The Italian mathematician Cesare Burali-Forti published this paradox in 1897, but Cantor had already discovered it two years earlier.

sets of the previous stage, i.e. of individuals, but we don't get sets which contain *sets of sets* of individuals. That is to say, at stage with rank 2 we can only form sets out of things with rank 1 or 0; *at stage 2* there are no sets *of stage 2* to be collected: the only sets to be collected are of previous stages. This means, among other things, that no set α can contain itself, or sets which contain α as a member (or sets which contain sets which contain α as a member, etc.) Now, a set which has this property, i.e. which contains only sets of lower rank than itself, is called a *well-founded* set. In the standard, Zermelo-Fraenkel set theory, all sets are considered to be well-founded, and the cumulative hierarchy of sets which we have been describing so far is accordingly supposed to be the hierarchy of well-founded sets.

But now consider the hierarchy itself. If it is a set all of whose members are well-founded sets, it must be well-founded. But if it is a well-founded set, it must belong to the hierarchy –i.e. to itself-, since the hierarchy is the set which contains exactly all the well-founded sets. But if it belongs to itself it is not well-founded!¹

Now, without having in mind a connection to Kantian concerns, Bertrand Russell identified a pattern that fits the set-theoretic paradoxes. He wrote:

Given a property ϕ and a function δ , such that, if ϕ belongs to all members of u , $\delta(u)$ always exists, has the property ϕ and is not a member of u ; then the supposition that there is a class Ω of all terms having the property ϕ and that $\delta(\Omega)$ exists leads to the conclusion that $\delta(\Omega)$ both has and has not the property ϕ .²

How do the paradoxes fit in here? Paradoxical collections have so far always been oversized collections: of all sets, of all non-self-membered sets, of all ordinals, of all well-founded sets. This indicates that these oversized collections play the role of Ω in Russell's pattern. Now, the paradoxes have always been proved by referring to some object that cannot be a member of the alleged Ω : an object in the Power-set of U but not in U , Russell's collection which cannot be a member of itself, the first ordinal to succeed the collection of all ordinals, the cumulative hierarchy itself. Now, when we are considering unproblematic collections u , a function δ allows us to refer to an object $\delta(u)$ that is not a member of u . In Cantor's case, the function δ gives us a member of the Power-set of α that is not a member of α . In Russell's case, if we consider an unproblematically delimited set of non-self-membered sets (like the set of all singletons), the function δ is simply the self-referring function: it

¹ This is Graham Priest's version of a paradox whose discovery is attributed to Dimitry Mirimanoff. Cf. Priest's (2002) pp. 119-20, who refers to Hallett (1983) section 4.4.

² "On Some Difficulties in the Theory of Transfinite Numbers and Order Types", quoted by Priest (2002) p. 130.

gives us a set, namely itself, which is not a member of the previously considered set (the set of singletons has more than one member). For every unproblematical series of ordinals α , $\delta(\alpha)$ simply refers to the least ordinal greater than α . And for every set s of well-founded sets, $\delta(s)$ simply refers to a set that contains s and which, of course, is not in s . The problem arises, as Russell points out, as soon as we apply the function to the oversize totality. That is to say, as soon as we think of $\delta(\Omega)$ in each case and of ϕ as the property which defines Ω . The set which is not in U must be in U because U is the set of all sets. The Russellian set must be an element of itself because it must fulfill the defining condition. The least ordinal greater than all ordinals, being an ordinal, must be in the series of all ordinals. And the set of all well-founded sets, being well-founded, must be contained in itself.

Graham Priest has put Russell's pattern in the following schema:

- | | |
|--|-----------------------------|
| (1) $\Omega = \{ y \mid \phi(y) \}$ exists. | <i>Existence</i> |
| (2) if x is a subset of Ω : a) $\delta(x) \notin x$ | <i>Transcendence</i> |
| b) $\delta(x) \in \Omega$ | <i>Closure</i> ¹ |

If a set of all sets exists (*Existence*), there must be a set in its Power-set which is not in it (*Transcendence*) but which, on account of being a set, must be in it (*Closure*). If a set of all non-self-membered sets exists (*Existence*), it can't belong to itself (*Transcendence*) but it must (*Closure*). If there is a set of all ordinals (*Existence*) the principle of generation for ordinals forces us to commit to the existence of least ordinal which is not in that set (*Transcendence*) but which, on account of being an ordinal, must be in that set (*Closure*). Finally, if the set of all well-founded sets exists (*Existence*), on account of being well-founded, it can't be a member of itself (*Transcendence*), but then, on account of having as members *all* the well-founded sets, it can't fail to be a member of itself (*Closure*).

We said a few pages ago that this excursus to the paradoxes of Burali-Forti and of the cumulative hierarchy should help us to understand why Graham Priest, following Gottfried Martin, thinks that there is a deep connection between Kantian and Cantorian antinomies. Now we are in a position to understand this connection. What Martin calls a conflict "between making a conclusion and going

¹ Priest, G. (2002) p. 130. Priest presents a more general but more complicated schema on p. 134. The latter schema he calls the Inclosure Schema. In Priest's Inclosure Schema, there is a further condition ψ that must be fulfilled both by Ω and by the subset x of Ω referred to in condition (2). The motivation for Priest's schema is to accommodate the paradoxes of definability, with which I am not concerned in this investigation. And given that when the schema is seen as characterizing the paradoxes of set theory the condition ψ is simply that $x = x$ (by Priest's own admission "the universal property"), it is better for our purposes to work with the simpler schema.

beyond the conclusion” is the opposition that Priest sees between *Closure* and *Transcendence*. These tendencies of thought are responsible for each side of a (Kantian or set-theoretical) antinomy. Take *Closure*, for instance. Priest says:

As Kant saw so well, given a notion like that of set or ordinal, reason forces us to conceive of the totality of all things satisfying it. Totalising is part of our conceptual machinery –like it or not.¹

That is to say, Kant had a doctrine according to which what motivates one side of the Antinomy – the *thetical* side- is reason’s urge to concluding any process of iteration by framing the concept of a *closed* totality that includes all objects iterated. On the other side there is reason’s opposite urge: the tendency that motivates the Antithetical side of Kant’s Antinomy. As Kant himself puts it, the concept of the unconditioned totality (in this case, the concept of the world) is “too small” for reason’s ambitions (A486-7/B515-6). Hence reason’s refusal to recognize any limit, its tendency to break through every barrier. But this anti-totalising tendency can also be said to motivate the “ δ ” function in the set-theoretical antinomies: the object that is “diagonalised out” is systematically defined so as to guarantee that it *does not* belong to a previously considered totality.²

4.3. A Farewell to the Unconditioned?

I would like to finish this chapter by pointing at some elements in Kant’s philosophy which provide the best way of expressing an idea which suggests itself after what we have been defending so far. If the concept of an absolutely all-embracing totality is self-contradictory, if the concept of *Everything*

¹ Priest, G. (2002) p. 162.

² Cf. Priest, G (2002) p. 130. Priest also presents what he calls, “with apologies to Kant”, the Fifth Antinomy (p. 100 and later, more formally, on p. 131). This antinomy is generated by thinking of any object *x* and taking the *thought of x* as a starting point. The thought of *x* is not the act of thinking but the content of thought, and *ex hypothesi* it is non-self-referential: it is about *x*, not about itself. Symbolize the thought of *x* by “*t(x)*”, the thought of the thought of *x* by “*t(t(x))*”, etc., and the series of thoughts so generated by “*T*”. The antinomy is that, since *T* cannot be generated as a member of the series, *T* cannot be thought of, but since we are obviously thinking of it, *T* can be thought of; alternatively, the antinomy is that *t(T)* is and is not a member of *T*. Now, I find Priest’s reasons for the statement of this antinomy rather unconvincing. As regards Priest’s formulation on p.100, it must be said that, even granting that *t(T)* cannot be generated as a member of *T*, it does not follow that “*T* cannot be thought of (Transcendence)”. Alternatively, on p. 131, while stating the antinomy in the arguably equivalent form of *t(T) ∈ T* & *t(T) ∉ T*, Priest says in support of Closure: “*T* is the sequence of thoughts generated by applications of *t* to any sequence of objects generated; *T* is a sequence of objects generated; so *t(T) ∈ T*.” But *T* is *not* the sequence generated by applications of *t* to *any* sequence of objects generated. Each time we pick an object *x* (like Priest does with the *Critique of Pure Reason*), a *particular* series *T* is generated. The thought of this particular series *T* might be said to belong to Thought, in general, but, as the Transcendence of this version makes clear, it does not belong to the particular series *T*.

is genuinely paradoxical, then, it seems, we must give up that totalizing concept altogether. Not that Kant himself drew this conclusion or at least an analogous conclusion from his alleged antinomy. He quite evidently didn't, and that was one of the flaws in his argumentation.¹ But the terms with which Kant expressed *his* problem allow us to depict the set theoretical paradoxes as a more radical version of the Kantian Antinomies.

Let us recall, first of all, that Kant thought that his Antinomies are generated by the principle which in chapter I we called the principle of the unconditioned: *If the conditioned is given, then the whole sum of conditions, and hence the absolutely unconditioned, is also given.* (cf. e.g. A409/B436 and A497/B525). This was actually just one of the necessary conditions for the generation of the Antinomies; the other one was the assumption of the truth of transcendental realism.² This second condition has been for us not only unproblematic but essential, for we saw that the Antinomy is supposed to be the *coup de grace* to realism, its *reductio ad absurdum*, "if perhaps someone did not have enough in the direct proof in the Transcendental Aesthetic" (A506/B534). So let us focus on the first condition, the assumption of the principle of the unconditioned. What is important for our present purposes is that there is, according to Kant, a particular *concept* corresponding to the principle of the unconditioned. We saw that Gottfried Martin pointed out that it is the concept of *totality* which generated both claims of an antinomy. He was right, but it must be emphasized that not every concept of a totality creates a paradox. It is only the concept of a very radical kind of totality, i.e. of an absolutely unconditioned totality:

So the transcendental concept of reason is none other than that of the totality of conditions to a given conditioned thing. Now since the unconditioned alone makes possible the totality of conditions, and conversely the totality of conditions is always itself unconditioned, a pure concept of reason in general can be explained through the concept of the unconditioned, insofar as it contains a ground of synthesis for what is conditioned. (A322/B379).

And a few lines later:

Now a transcendental concept of reason always goes to the absolute totality in the synthesis of conditions, and never ends except with the absolutely unconditioned, i.e., what is unconditioned in every relation. For pure reason [...] reserves for itself only the absolute totality in the use of

¹ In particular, it is precisely the availability of this conclusion which unmasks the failure of the "indirect proof" of transcendental idealism, as Gram quite correctly saw (Cf. (1967), p. 512). All this is, of course, under the quite generous supposition that the First Antinomy is genuine.

² This is correctly pointed out by Grier, M. (2002) p. 180.

concepts, and seeks to carry the synthetic unity, which is thought in the categories, all the way to the absolutely unconditioned *<bis zum Schlechthinunbedingten>*. (A326/B382-3).

In short: what generates the antinomies is the idea of the unconditioned. As used by Kant, the expression “the unconditioned” refers to the spatiotemporal world. But he failed to generate an antinomy from the concept of *the spatiotemporal world*. However, we have seen that there is a radical version of Kant’s concept of *the spatiotemporal world*, and that is the concept of *the totality of all things*. Given that this latter concept presents a more threatening paradox (the paradox of the largest cardinal), now that we have made reference to Kant’s concept of *the unconditioned* we can see the set-theoretical paradox as a radical version of Kant’s alleged antinomy. When we specify a set, we do this by specifying the condition that must be fulfilled by the objects that are the members of this set. The set of red things is the set of those objects that *are red*. The condition that determines this set is the condition “to be red”. The set of all *x*’s such that *x* is red is symbolised: $\{x \mid x \text{ is red}\}$. Now, we can think of the defining condition that appears on the right as the most general condition whatsoever: a condition that has to be fulfilled by absolutely all objects. For all those philosophers that are not Hegelians or dialetheists, this is the condition of self-identity. The set of all objects is then: $\{x \mid x = x\}$. This is a totality which, as regards comprehension, is absolutely unconditioned: there is no object that is not comprehended by this collection. Compare with Kant’s: $\{x \mid (x = x) \ \& \ (x \text{ is a spatiotemporal object or event})\}$. Clearly, the totality $\{x \mid x = x\}$ is more unconditioned than the Kantian one because by dropping the second conjunct of the defining condition the set in question becomes the more comprehensive set that there could possibly be. Hence, the concept of *the Universal set* can be seen as a radical version of Kant’s concept of *the spatiotemporal world*, and hence as a radical version of Kant’s concept of *the unconditioned*.¹

Kant’s own ideas about the concept of *the unconditioned* get fleshed out in his claim that transcendental idealism as a key to the solution of the alleged mathematical antinomies. The preface to the second edition of the *first Critique* is particularly eloquent about this matter. The unconditioned is that which, on the assumption of transcendental realism, “cannot be thought at all without contradiction”: the thing in itself. If on the other hand one adopts transcendental idealism and thinks of “the thing in itself as something actual for itself but uncognized by us”, then “the contradiction disappears” (BXX). This is just another way of saying that the Kantian Antinomies are supposed to be solved by the Kantian doctrine of transcendental idealism. Officially, the First and Second Antinomies get a “both-false” solution whereas the Third and Fourth Antinomies get a

¹ I wish to thank Peter Simons for suggesting this formulation of the claim that the set-theoretical concept is a radical version of Kant’s concept of the unconditioned.

“both-possibly-true” solution. There simply should be no question regarding the unconditioned finitude or infinitude of the spatiotemporal world, since that world, being phenomenal, is necessarily conditioned. Similar considerations apply to the Second Antinomy: being phenomenal, material objects are always conditioned by our cognitive capacities, which means that they are always capable of being further decomposed (so the Thesis is false) but also that it makes no sense to say that there is an actual infinity of constituting parts of the world (so the Antithesis is false too). On the other hand, freedom and a necessary being are said to be possible in the realm of things in themselves, while in the realm of appearances there is neither an uncaused cause nor a necessary being (A528-32/B556-60). These are, as I said, the official ways to make the contradictions disappear. Unofficially, certain passages suggest that even during his critical period Kant had in mind something like a “both-true” solution for the Second Antinomy.¹ Even more unofficially, but still within the confines of transcendental idealism, A.W. Moore advances an admittedly non-literal interpretation according to which a radical version of the First Antinomy would be solved by something like a Kantian “both-true” solution: the world considered in itself is *metaphysically* unconditioned, and the phenomenal world, although *mathematically* unconditioned, is metaphysically conditioned. That is to say, the world considered in itself would be a Kantian version of what Moore calls the concept of the *metaphysical* infinite, a concept associated with absoluteness, completion, perfection, etc. On the other hand the phenomenal world, even if conceived of as *infinite*, would be metaphysically conditioned in two respects: first, because it would not be a thing in itself; second, because the set theoretical universe is only potentially, not actually, infinite. Still, this set theoretic universe would be in a sense unconditioned, namely in that sense which, according to Moore, corresponds to the *mathematical* concept of the infinite – a concept usually associated with boundlessness, immeasurability, unsurveyability, etc.²

Fortunately, transcendental idealist “solutions” are not forced upon us. Not by logic at any rate. If the concept of a *square circle* is self-contradictory, then the most natural thing to do, at least in the non-Hegelian quarters of philosophy, is to say that there is nothing out there corresponding to that concept. The same for every other concept which reveals itself as contradiction-entailing. Kant seems to have been very certain about the emergence of the antinomies given realist

¹ See Kant’s 1783 *Metaphysical Foundations of Natural Science*: “Now the *composite of things in themselves* must certainly consist of the simple, for the parts must here be given prior to all composition. But the *composite in the appearance* does not consist of the simple, because in appearance, which can never be given otherwise than as composed (extended), the parts can only be given through division, and thus not prior to the composite, but only in it.” (Canonical pagination: pp. 507-8. Our edition: p. 45).

² Cf. Moore, A.W. (1988) pp. 205-18 and his (2001) pp. 84-95.

presuppositions.¹ Given his further statement that it is the idea of the unconditioned what generates the antinomies, the only conclusion to which he was entitled was that the transcendental realist should bid farewell to the concept of the unconditioned, *not* that transcendental realism is absurd.

It is surprising how few are the commentators that have been alive to this fact. An interesting case is William Hamilton, the nineteenth century Scottish philosopher. Although clearly wrong in taking the first antinomy to be genuine, his reasoning was nonetheless correct: *if* the concept of the world is the concept of the unconditioned, and *if* it turns out to be a self-contradictory concept, then we non-Hegelians should simply give up the concept altogether:

Kant has clearly shown, that the Idea of the Unconditioned can have no objective reality, -that it conveys no knowledge,- and that it involves the most insoluble contradictions. But he ought to have shown, that the Unconditioned had no objective application, because it had, in fact, no subjective affirmation; that it afforded no real knowledge, because it contained nothing even conceivable; and that it is self-contradictory, because it is not a notion, either simple or positive, but only a fasciculus of negations –negations of the Conditioned in its opposite extremes, and bound together merely by the aid of language and their common character of incomprehensibility. (The Unconditioned is merely a common name for what transcends the laws of thought –for the *formally illegitimate*).²

Taken beyond the limits of Kantian philosophy, these remarks are of great value. For we know not only that there is a more radical version of Kant's concept of the unconditioned, but also that we have good reasons to believe that the radicalized concept is paradoxical, in contrast to the fact that Kant provides no good reasons to believe that his concept of the world is paradoxical. So we know that there is *a* concept of the unconditioned which, on pain of self-contradiction, has to be abandoned: the concept of *Everything*. This means that at least in one respect the set theoretical paradoxes succeed in showing something that Kant did not succeed in showing. If Kant was not being disingenuous when he wrote that the concept of the unconditioned would generate antinomies for those who did not assume his transcendental idealism, then history seems to have proven him right, although he was only partly right and for the wrong reasons. He was partly right because the concept of *Everything* as a *definite* thing does present paradoxes, and this concept deserves, perhaps better than any of those Kant had in mind, to be identified as the concept of *the* unconditioned. If being is necessarily definite being, it does look as though the unconditioned cannot

¹ In the footnote to §52b of the *Prolegomena* he says: "I promise to answer for each proof I have given of both thesis and antithesis, and thereby to establish the certainty of the inevitable phenomenon of reason."

² Hamilton, William [1829], p. 17. On pp. 29-30 Hamilton suggests that the First Antinomy is genuine.

be thought of without contradiction. But this is a problem not only for the transcendental realist in Kant's sense; it is a problem for anyone who happens to think that there is such a (definite) thing as Everything.

Now, in our chapters II and III we have looked at different ways out of this problem. According to one of them, Quine's NF and related systems, we can still think of a definite totality which contains everything (a universal set) but the principle for the derivation of collections gets technically more complicated than the unrestricted axiom of comprehension. Now, there were two other problems with NF and related systems. A set of all sets must be a member of itself. Some people have complained that the notion of self-membership is counter-intuitive and mind-boggling. Although in the previous chapter I sided with those who complain in this way, I also pointed out that this cannot be a conclusive criticism, since our intuitions might not be the most helpful counselors when we are discussing mathematical collections –there is nothing "intuitive" about the claim that the set of all multiples of 1000 has the same number of elements as the set of all natural numbers. Complaints about counterintuitiveness apart, however, we also saw that there is a more important problem: self-membered sets cannot be non-circularly individuated, not even by an Omnipotent Mathematician.

Another option, which is still an attempt to think of Everything as an individual thing, is to think of that totality not as a set but as a plenum, which is an ever-growing, (not only self-membered but also) self-amplifying and explosive totality. As I pointed out in the previous chapter, this option not only inherits the circularity problems of the individuation of self-membered collections, but it also faces the problem that the claim that the universe changes its size on account of the explosiveness in content of any given conception of the universe's size is incompatible with a professed realism. This latter difficulty, as we said, is not exclusive to plenum theory but is shared by any conception that attributes the universe the property of ever-growing *on account of* our ability always to find new elements of some problematical collections. If the Cantorian distinction between consistent and inconsistent multiplicities is to be understood in this way, its conflict with realism should be pointed out.

The less problematical way out of the paradoxes seems to be, then, a radical abandonment of the reification of Everything. This doesn't mean that the concept of *everything* should be abandoned.¹

¹ See Cartwright (1994) and Williamson (2003) for the view that, as long as we give up reification, unrestricted universal quantification is possible and desirable.

By everything we mean the multiplicity of everything: many things and not one.¹ This option might even be made compatible with Cantor's distinction between consistent and inconsistent multiplicities, if we emphasize Cantor's claim that the assumption that inconsistent multiplicities can be gathered into "unities" leads to contradiction. Granted, this claim can also be fleshed out by an appeal to the conception of the universe as an "ever-growing" collection. But we just pointed out the problems of this interpretation in the previous paragraph. Now, Cantor's idea is simply that some multiplicities are not "ones" and hence not iterable things. This much seems to be true at least about the multiplicity constituted by every single thing.²

Conclusions

According to Kant, the antinomies of pure reason are generated by the idea of the absolutely unconditioned. He failed to show this, for his concept of the unconditioned was that of the spatiotemporal world, and this concept is not problematic in the sense he thought it was. But this

¹ Russell also had a distinction between classes as many and classes as one, as Cartwright (1994) p.8 points out.

² In chapter II, at the end of section 2.3, I made a criticism of the interpretation of Cantor's inconsistent multiplicities as not being unities at all. That criticism seems to be incompatible with my present proposal of understanding the multiplicity of all things precisely as a multiplicity and not a unity, i.e. with my present endorsement of a non-reificationist attitude towards the multiplicity of everything. The tension is important but it might not be insurmountable. At that point, my main argument was that we can't help treating even inconsistent multiplicities as unities because we want to talk about each one of them and differentiate them from one another (we don't want to identify the collection of all ordinals with the collection of all cardinals, for example). Now, if we really give up the reification of the multiplicity of all things, there would be no unity here to be differentiated from any other thing, and hence the earlier argument against Cantor's inconsistent multiplicities would not apply. This is what I mean in my last sentence in the text when I say that non-reificationism seems to be the best attitude "at least" when we are considering the multiplicity constituted by every single thing. If this line of argument is correct, then some of Cantor's inconsistent multiplicities would still be unities *pace* Cantor (and here the earlier argument would apply: they would be unities as long as we differentiate them from one another), whereas the multiplicity of everything would be a genuine multiplicity, not an individual object or unity (and, not being an object, "it" needn't be distinguished from anything). However, I admit that the present position can be challenged as being somewhat *ad hoc*. If the multiplicity of all things can be dealt with in a non-reificationist way, why couldn't the other inconsistent multiplicities be dealt with in the same way? Shouldn't non-reificationism be pursued in the most general possible way instead of only with regard to the multiplicity of everything? Indeed, I think that non-reificationism should be pursued generally, and it is even desirable that a non-reificationist account be given not only of Cantor's inconsistent multiplicities but also of all multiplicities. However, I cannot at the present moment develop this general anti-reificationism. Moreover, the challenge which seems to me the most pressing against anti-reificationism in the general case (we usually want to distinguish multiplicities from one another) is not threatening in the particular case of the multiplicity of all things: there seems to be no reason why we would want to treat this multiplicity as a unity and differentiate it from anything else. Hence, I take the general version of anti-reificationism to be a project to be developed in further work, and the individuation and differentiation of mathematical collections to be the main challenge to be surmounted by that project. I wish to thank Øystein Linnebo and especially Peter Simons for helpful comments and criticisms on this point.

does not mean that puzzlement is not waiting for us around the corner whenever we frame the concept of what is absolutely unconditioned. If the multiplicity of everything is thought of as a collection, at least three problems arise: 1) this collection must be self-membered; 2) if conceived as a set in the Cantorian sense, it leads to paradoxes; and 3) if conceived as a definite entity, the reification of Everything seems to bring with it the reification of Nothingness. I have expressed my own (and if you like dismissible) puzzlement at point 1). Points 2) and 3), if they are correct, need only be stated, I hope, in order to show that this is not the way we should go.

Whether these results are able to be given a Kantian interpretation depends on how attached to transcendental idealism the "good" Kantian should be. Transcendental idealism was neither a plausible nor a possible solution to the alleged first antinomy, and nothing with a family resemblance to the distinction between phenomena and things in themselves is justified by the present results. However, if the reification of Everything ends up either in puzzlement, or in contradiction or in the reification of Nothingness, or in the three of them, this result does suggest (to me, anyway) that the proud name of an Ontology should give way to an analytic of the understanding. Not because Ontology is the kind of investigation we would be able to carry out if our understanding could reach the unconditioned, but because reificationist talk about the unconditioned is precisely what we do not understand.

Conclusions

Throughout most of this investigation we have been concerned with one question. It is the question whether there is such a thing as everything, i.e. whether the multiplicity of everything should be considered as an individual –a “one” or unitary object. An affirmative answer to this question could be called the doctrine of *absolute monism*.¹ Alternatively, we can call the thesis that there is such a thing as everything the *reification of Everything*.

Strictly speaking, this was not the thesis under discussion in our first chapter. There we dealt with Kant’s claim that there is a “first antinomy of pure reason”, i.e. a conflict that arises because, under the assumption of transcendental realism, we must affirm both that the spatiotemporal world is finite and that it is infinite –a paradoxical situation to be solved by the rejection of transcendental realism and the adoption of transcendental idealism. Were we to circumscribe ourselves to a literal interpretation of Kant’s text, however, there would be very little to be learned from it. Neither is there an antinomy nor, if there were one, would transcendental idealism follow from its alleged existence. There is no antinomy because the “proofs” provided by Kant for the conflicting theses can coherently be rejected (by transcendental realists, but not only by them). The argument for the thesis rests on a definition of infinity according to which there can be no *completed* infinities. But, even without taking into account post-Kantian theorizing about the actual infinite, the *contradictio in termini* of the notion of a complete infinity is precisely what cannot be presupposed when one is trying to argue against someone who claims that the series of past events is infinite. The failure of the argument for the antithesis is perhaps less notorious and more understandable, but it is still a failure. The argument rests on a definition of “beginning” according to which the thing that begins

¹ This terminology agrees with the one used by Peter Simons in his (2003): “In calling the universe a plurality or multiplicity I am stressing that it is many, not one. [...] The universe is not an individual: monism is false. It is a multiplicity.” (pp. 248 and 249) The nomenclature is debatable, however. Jonathan Schaffer has pointed out that traditionally monism is understood as the doctrine that only one thing exists: $\exists x \forall y (y = x)$. We, on the other hand, are understanding monism as the thesis that there is one object such that everything is a member of it: $\exists x \forall y (y \in x)$. Schaffer himself is interested in defending the thesis that there is only one *concrete* object at the *fundamental* level: the thesis that the (concrete) whole is prior to (more fundamental than) its (concrete) parts. (See his (2007) and (2010) *passim*). As Schaffer points out, what makes all these varieties of monism is that they attribute “oneness”, but they are to be distinguished by the object that they attribute oneness to and what objects they are counting as relevant. In our discussion, as in that of Simons (cf. pp. 249–50), we count as relevant all objects (whether concrete or abstract, real or fictional, actual or possible, thinkable or unthinkable, etc.), and the monism we are targeting attributes oneness to the multiplicity of all these. Schaffer, on the other hand, only counts as relevant concrete objects and says that, at the fundamental level, there is only one.

must be preceded by a time in which the thing was not, with which the first event in time (the world's absolute beginning) must have occurred immediately after a particular point in empty time – a conclusion that would violate the principle of sufficient reason. Assuming the principle of sufficient reason (or Kant's version of it), the argument is indeed very strong if one assumes that the finite spatiotemporal world is situated in an absolute and infinite (space-)time. Leibniz knew this and used this same argument against Clarke. But, as Leibniz knew as well, the argument can be avoided by rejecting the assumption of the infinity and absoluteness of time, and claiming that the first event happened at the same time.¹ There is, then, no such thing as a first antinomy with respect to the spatiotemporal world. Not Kant's, anyway. More importantly, even if there had been an antinomy, its existence wouldn't justify the claim that not only space and time but also all spatiotemporal objects are mere representations, as opposed to things in themselves. Kant's "indirect proof" of transcendental idealism is a threefold failure.

Whether one takes Kant's discussion to be a mere historical curiosity is very much a matter of taste. We are certainly allowed to do so. In this investigation, however, we have taken a different attitude towards the first *Critique*. We have taken our cue from Kant's claims in the Preface to the second edition that metaphysics' most important task is to think of the unconditioned, and that the discussion in the Antinomies chapter is supposed to show that traditional metaphysicians cannot think of the unconditioned without contradiction. If "traditional metaphysicians" is to stand for transcendental realists, and if "contradiction" is to stand for the antinomies exactly as Kant thought them to be, then surely he failed in his attempt to show that traditional metaphysicians cannot think of the unconditioned without contradiction.² But if we abstract from Kant's discussion with the transcendental realist, and if we think of the concept of something that would be so unconditioned that it, and only it, could satisfy reason's metaphysical ambitions, it seems that there are theoretical problems arising from such a concept. This is what we have tried to show in our chapters II, III and IV. The problematical concept, the concept that plays the role of *the unconditioned* in our last three chapters, is the concept of an absolutely all-embracing collection: the collection of everything.

In chapter II we saw that whenever we think of a collection of objects as a Cantorian set, the concept of a set of all sets does generate paradoxes, as was discovered in the second half of the nineteenth century. It is perhaps incorrect to call these paradoxes "paradoxes of Cantorian set

¹ Alternatively, one could say, as Clarke did, that sometimes there is no further reason to be sought for than that God willed it thus and not otherwise.

² Although it must be said that we have only argued for this critical claim as regards the first antinomy (and in fact only in its "temporal side"). Of course, Kant's claim that the remaining three antinomies are unavoidable is highly debatable, but we have not examined any other antinomy apart from the first one.

theory” because Cantor himself rejected the idea that every collection must be a set in his sense of the word, and hence he was allowed to see the paradoxes as arising from an illegitimate understanding of some oversize collections as Cantorian sets. This already provides a connection with Kant. Had there been a genuine first antinomy, all that would have been shown is that the collection that Kant calls “the world” cannot be regarded as being of the same kind as the objects that make it up.¹ In Cantor’s case, it seems that the best way to deal with the paradoxes (i.e. the way to avoid them) is to deny that all collections (of sets) are of the same kind as the objects that make them up: not all collections are Cantorian sets. So let us summarize the conclusions of our second chapter with the following words. The relation that a collection has to its members is called comprehension. A collection which, as regards comprehension, is absolutely unconditioned, is a collection that has all objects in it; it is the collection of everything. If we think of this collection as a Cantorian set, and hence as a definite object, a paradox (the so-called Cantor’s paradox) seems unavoidable. This is, indeed, the first problem with the *reification of Everything* that we have encountered in our investigation.

But the paradox is not unavoidable for all possible sorts of reification of Everything. It is a paradox for those reifications for which Everything is a Cantorian set, but this still leaves at least the following two reificationist options available: 1) to think of Everything as a set but not a Cantorian set; and 2) to think of Everything as a collection which is not a set at all. These two options were presented in our chapter III. A collection might still be called a set if the axiom of extensionality is true of it. But it might be a non-Cantorian set if Cantor’s Theorem does not apply to it, i.e. if there is a bijection between the collection and its Power-collection, and so both collections would have to be assigned the same cardinal number. This idea is behind the system of set theory proposed by Quine in 1937, a system called “New Foundations”. If some collections are non-Cantorian sets, the collection of everything can be thought of as one of them, and if this is so then Cantor’s paradox would not arise anymore, for the absurd conclusion that a collection has more members than the most inclusive collection would be avoided. So the reification of Everything is free from paradox as long as Everything is a non-Cantorian set. A relatively similar conclusion can be drawn if Everything is still thought of as a singular collection, but not as a set, i.e. if Everything is a collection that does not

¹ Kant seems to have had some understanding of this, since one of the things that his final solution to the antinomy implies is that the world is only potentially and not actually infinite, and hence that the world is not an actual object or thing, as the things that make up the world. Reading Kant sympathetically, some authors had identified this as Kant’s solution proper (Priest, for example, calls this “solution 1”, and A.W. Moore refers to the negation of the actual existence of the world a “Kantian solution”). But strictly speaking this cannot be the whole story about Kant’s solution, since otherwise he wouldn’t have thought of the antinomy as providing an “indirect proof” of transcendental idealism.

behave according to the axiom of extensionality, and hence it is not a determinate collection. Nicholas Rescher and Patrick Grim, as we saw, think that the oversize collections in which some philosophers are interested are best understood as “plena”, which are indeterminate collections. Plena have members corresponding to every subcollection they have, and *membership* plena are those for which the correspondence in question is identity: membership plena have *as* members every subcollection they have. The collection of Everything is according to Rescher and Grim a membership plenum, and hence Cantor’s paradox is avoided by the very definition of a membership plenum: the “diagonalised” collection, that was essential in proving that the Power-set of a set is larger than the original set, is a subcollection of the plenum, and in the case of membership plena the definition implies that it is also a member of the original plenum.

Technically, then, both systems presented in chapter III succeed in their reification of Everything, either by modifying the principle for the derivation of sets (NF) or by allowing into the system indeterminate collections (Plenum Theory). But as I tried to point out there is at least one aspect shared by those systems which might present difficulties. If Everything as a collection, it must be self-membered. And there is something unsatisfying about the idea of self-membership, as has been pointed out not only by the advocates of the standard set theory (for which every set is well-founded). In the case of sets, the problem arises out the axiom of extensionality. If the identity of a set is determined by its members, then not even an idealized infinite subject –the Omniscient Mathematician- could non-circularly “individuate” a self-membered set. In the case of plena, their defining characteristic (having a member corresponding to each of their subcollectivities) makes them admittedly “explosive in content”, i.e. any possible conception of the cardinality of a plenum is necessarily insignificant, since the plenum has not only as many members as the Power-collection of a collection with the cardinality originally attributed to the plenum: it has also as many members as the Power-collection of the previously introduced Power-collection. And so on for any possible attribution of cardinality to a plenum. Thus not even the Omniscient Mathematician can have a definite conception of the cardinality of a plenum: plena have no definite cardinality. They keep “growing beyond any determinate conception at any time”, as our authors put it. This is certainly not self-contradictory, but if it applies, as it must, to the conceptions of the Omniscient Mathematician, the proposal does make us wonder in what sense is the Omniscient Mathematician omniscient. It might be thought that this is a problem beyond the scope of Plenum Theory, since the theory does not involve a commitment to the existence of an Omniscient Mathematician or of any idealized subject of any kind; but unless such a commitment is made, the claim that plena “grow beyond any conception any time” is incompatible with the realism that Rescher and Grim profess. This is, then, a

first difficulty with plena in general, that they are “explosive in content” and “mind-expanding” not only for us, as the authors seem to think, but for any idealized subjectivity. Now, *membership* plena are said to be self-membered. This property is not problematical for plena in the same way that it was for sets. Still, nothing in the characterisation of plena clarifies the idea that, amongst the members of a collection, there is the collection itself.

True, these problems are not paradoxes. But they are problems anyway. At least for some of us. So let us express the findings of our chapter III with the following sentence, to keep up with our Kantian concern: the unconditioned (the reified Everything) can be thought of without contradiction, but we still face problems in our attempt to understand it. Ironically, the historical Kant would have welcomed this claim, although he would have understood it in a completely different way. For he obviously wanted to say that the unconditioned can be thought of without contradiction; it was only his opponents, the realists, that were supposed to get involved in contradictions while trying to think of the unconditioned. Availing of the distinction between phenomena and things in themselves, Kant claims that we are able to *think* of the unconditioned but unable to *know* it, and this means that the unconditioned must be seen as lying beyond the reach of the understanding. Now, no dualism is justified by the findings of our chapter III. Moreover, the fact that some of us are unable to understand self-membership does not mean that all of us are unable to understand this property¹, let alone that no one could possibly understand it: for all I know, my inability to understand self-membership is an empirical fact about myself, a cognitive predicament that I happen to share with George Boolos. So no Kantian dualism, no *a priori* and insuperable barrier between what can and what cannot be reached by our understanding, is justified by our findings of chapter III. Still, the traditional antinomies surrounding the unconditioned are solved, and the solution allows us to draw a (rather arbitrary) distinction between what some of us (as a contingent matter of fact) understand and what we do not. This is as close as our chapter III gets to Kant.²

¹ I have no intention to claim that those who embrace set theories that admit non-well-founded sets do not really understand what they are doing.

² The picture is compatible with Thomas Nagel’s appropriation of the Kantian idea of the finitude of human knowledge: “There are some things that we cannot now conceive but may yet come to understand; and there are probably still others that we lack the capacity to conceive not merely because we are at too early a stage of historical development, but because of the kinds of beings we are. [...] The only sense in which we can conceive of [the latter] is under that description –that is, as things of which we can form no conception- or under the all-encompassing thought “Everything”, or the Parmenidean thought “What is.” (Nagel, T. (1986) p. 92). Now, my investigation has challenged precisely the idea that Nagel seems to be taking for granted in this passage: the idea of a reified “Everything”. But in chapter III we saw that this idea is free from contradiction (i.e. that the reified Everything is not what Nagel would call “positively inconceivable”, like round squares [ibid.]). A reified Everything –if it is an individual and a collection, and hence self-membered- would be very much a case of what Nagel classifies as a weaker kind of negative inconceivability, i.e. it would be one of those

In our final chapter we looked at the similarities between the alleged antinomy of Kant and the traditional paradoxes of set theory. True, in chapter III we had already said that the paradoxes are not strictly speaking unavoidable. But an inquiry into the similarities that might obtain between Kantian “antinomies” and “Cantorian” antinomies is not uninteresting, first, because we also said that some difficulties arise for those systems presented in chapter III, but second, and more importantly, because those systems analysed in chapter III are not the standard ways of theorizing about mathematical collections. Zermelo-Fraenkel set theory provides a useful framework to conceptualize classical mathematics. If in that theory paradoxes would be generated if we admitted unconditioned collections as objects of the theory (i.e. as sets), the result is in a sense quite Kantian, namely in the sense that unconditioned totalities would have to be either rejected as paradoxical (if they are sets) or treated as lying beyond the scope of the theory (if they are non-set-multiplicities). Not only that. We saw that, as Gottfried Martin and Graham Priest have pointed out, there are structural similarities between the two sorts of antinomies. They seem to arise out of the same two tendencies of human thought (to totalise and to go beyond any previously established totality). Thus, our task in chapter IV was, first, to present the concept of the collection of everything as a radicalization of Kant’s concept of the world. Second, to recast the traditional paradoxes that arise if this collection is taken as one of the objects of the theory (as a set) in such a way that the structural similarities pointed out by Martin and Priest are evident.¹ Third, we saw how these paradoxes could be avoided by rejecting the existence of the problematical entity

Let us conclude. If we think of the multiplicity of everything as an individual thing, either paradoxes arise or else some of us have to put up with the obscurity of Everything’s self-membership, if not of its indefiniteness. For those of us who face these problems, then, there is something vaguely reminiscent of Kant’s starting point about this situation. For what we have done so far with the idea of the unconditioned amounts to little more than a mere groping in the dark.

things that “we cannot now conceive but may yet come to understand”, if the “we” is understood as referring to me (D.B.), George Boolos and perhaps other people. Thus, we are not entitled to take this Parmenidean thought for granted in arguments against idealism; but those for whom the thought is unproblematic are certainly free to avail of it.

¹ The Russellian schema presented by Graham Priest was particularly helpful in the identification of these similarities.

Bibliography

- Aczel, Peter (1988) *Non-Well-Founded Sets*. Lecture notes, no. 14. CSLI: Stanford, California.
- Adams, R. (1974) "Theories of Actuality" in M.J. Loux (ed.) *The Possible and the Actual*.
- Al-Azm, Sadik J. (1972) *The Origins of Kant's Arguments in the Antinomies*. Oxford University Press.
- Alexander, H.G. (Ed.) (1956). *Leibniz-Clarke Correspondence: Together With Extracts From Newton's Principia And Opticks*. (Manchester: University Press.)
- Allison, Henry E. (1976) "The Non-Spatiality of Things in Themselves for Kant" in *Journal of The History of Philosophy*. Vol. XIV. pp. 313-321.
- (1996) "Transcendental idealism: A retrospective" in *Idealism and Freedom*. Cambridge. Cambridge University Press.
- (2004) *Kant's Transcendental Idealism. An Interpretation and Defense*. 2nd Edition. (New Haven, London. Yale. University Press.)
- (2006) "Transcendental Realism, Empirical Realism and Transcendental Idealism". In *Kantian Review*. Volume 11.
- Bar-Hillel, Y. (1967) "Types, Theory of" in Paul Edwards (ed.) *The Encyclopaedia of Philosophy*.
- Bennett, Jonathan. (1974). *Kant's Dialectic*. (London. Cambridge University Press)
- Boolos, George. (1971) "The Iterative Conception of Set" in Boolos, G. *Logic, Logic and Logic*. Edited by John P. Burgess. Harvard University Press. 1998.
- (1989) "Iteration Again" in *Logic, Logic and Logic*.
- Bringsjord, Selmer (1985) "Are There Set Theoretic Possible Worlds?" in *Analysis* 45, p.64.
- Brittan, Gordon G. Jr. (1978) *Kant's Theory of Science*. Princeton: Princeton University Press.
- Buroker, Jill Vance. (2006) *Kant's Critique of Pure Reason: An Introduction*.
- Cartwright, Richard L (1994) "Speaking of Everything", *Noûs*, 28, 1-20.
- Cantor, Georg [1899] (1967). "Letter to Dedekind". In van Heijenoort J. (ed.) *From Frege to Gödel*. Harvard University Press.
- [1904] "Letter to Jourdain", translated by Shaughan Lavine. In Lavine, S. (1994) 98-9.

- Courant, Richard and Robbins, Herbert. [1941] (1996) *What is Mathematics? An elementary Approach to Ideas and Methods*. New York. Oxford University Press.
- Cresswell, M.J. (1972). "The World is Everything that is the Case" in M.J. Loux (1979).
- Deiser, Oliver. (2010). *Einführung in die Mengenlehre*. 3., korrigierte Auflage. Berlin. Springer-Verlag.
- Dummett, Michael (1991) *Frege: Philosophy of Mathematics*. Duckworth.
- Dümont, J. and Mau, F. (1997) "Are there any true Contradictions?" *Journal for the General Philosophy of Science*, 7, 1-11.
- Ebbinghaus, Heinz-Dieter. (2003) *Einführung in die Mengenlehre*. 4. Auflage. Heidelberg und Berlin.
- Forster, T. E. (1995). *Set Theory with a Universal Set: Exploring an Untyped Universe*. 2nd Edition. Oxford: Clarendon Press.
- (2006) "Quine's New Foundations" in *Stanford Encyclopedia of Philosophy*.
- (2008) "The iterative conception of set" in *The Review of Symbolic Logic*. Vol. 1 no.1 pp. 97-110.
- Fraenkel, A. Bar-Hillel, Y. and Levy, A. (1973). *Foundations of Set Theory*. Second revised edition. North-Holland. Amsterdam (Cambridge: Cambridge University Press.)
- Gardner, Sebastian. (1999). *Kant and the "Critique of Pure Reason"*. *Routledge Philosophy Guidebook* (London: Routledge.)
- Giaquinto, Marcus (2002) *The Search for Certainty. A Philosophical Account of Foundations of Mathematics*. Oxford. Oxford University Press.
- Gödel, Kurt [1944] (1964) *Russell's Mathematical Logic*. In P. Benacerraf and H. Putnam (eds.) *Philosophy of Mathematics. Selected Readings*. Oxford, Basil Blackwell. pp. 211-32.
- Gram, M. S. (1967). "Kant's first antinomy". In *The Monist*. pp. 499-518.
- Grier, Michelle (2002) *Kant's Doctrine of Transcendental Illusion*. Cambridge: University Press.
- Grim, Patrick (1984) "There Is No Set of All Truths" in *Analysis* 44: 206-8.
- (1991) *The Incomplete Universe: Totality, Knowledge, and Truth*. Cambridge, Mass and London. MIT Press
- Guyer, Paul. (1987) *Kant and the Claims of Knowledge*. Cambridge: Cambridge University Press.
- Hallett, Michael. (1984) *Cantorian Set Theory and Limitation of Size*. Oxford: Clarendon Press.
- Hamilton, William, Sir, [1829] (1993) "Philosophy of the Unconditioned" Hamilton, W. Mansel,

H.L. Seth, A. and Green, T.H. *Philosophy of the Unconditioned, On the Philosophy of Kant , The Development From Kant to Hegel and Lectures on the Philosophy of Kant*. London. Routledge, Thoemmes Press.

Holmes, M. Randall (2010) "Alternative Axiomatic Set Theories" *Stanford Encyclopedia of Philosophy*.

Jensen, Ronald Björn (1968) "On the consistency of a slight (?) modification of Quine's *New Foundations*" in *Synthese*. 19 pp. 250-63.

Kant, Immanuel. [1770] (1992) *On The Form and Principle of the Sensible and Intelligible Worlds [Inaugural Dissertation]*. In *Theoretical Philosophy 1755-1770*. Translated and edited by David Walford and Ralf Meerbote. (Cambridge and New York: Cambridge University Press).

[1781] (1998) *Critique of Pure Reason*. Translated by Paul Guyer and Allen Wood. (Cambridge: Cambridge University Press).

[1783] (1997) *Prolegomena to Any Future Metaphysics That Will Be Able to Come Forward as Science*. Translated by Gary Hatfield. (Cambridge University Press.)

[1786] (2004) *Metaphysical Foundations of Natural Science*. Translated by Michael Friedman. Cambridge University Press.

Lavine, Shaughan (1994) *Understanding the Infinite*. Harvard University Press.

Lear, Jonathan (1977) "Sets and Semantics", *Journal of Philosophy*, 72, 86-102.

Loparic, Zeljko (1990) "The logical structure of the First Antinomy". In *Kant- Studien*. 81, no. 3. pp. 280-303.

Loux, M.J. (ed.) (1979) *The Possible and the Actual: Readings in the Metaphysics of Modality*. (Ithaca and London. Cornell University Press.)

Maddy, Penelope (1983) "Proper Classes", *Journal of Symbolic Logic*, 48, 113-39.

Martin, Gottfried [1951] (1955) *Kant's Metaphysics and Theory of Science*. Translated by P.G. Lucas. Manchester, Manchester University Press.

Melnick, Arthur (1989) *Space, Time and Thought in Kant*. (Dordrecht: Kluwer).

(2004) *Themes in Kant's Metaphysics and Ethics*. Washington DC. The Catholic

University of America Press.

Moore, A.W. (1988) "Aspects of the Infinite in Kant" in *Mind*, 97.

(2001) *The Infinite*. 2nd Edition. London and New York Routledge.

Nagel, Thomas (1986) *The View From Nowhere*. New York and Oxford. Oxford University Press.

Nieman, Susan. (1994) *The Unity of Reason. Rereading Kant*. (New York: Oxford University Press.)

Oliver, Alex (1998) "Hazy Totalities and Indefinitely Extensible Concepts: An Exercise in the Interpretation of Dummett's Philosophy of Mathematics" in *Grazer Philosophische Studien*, 55, 25-50.

Parsons, Charles (1977) "What is the Iterative Conception of a Set?" in J. Hintikka and R. Butts (eds.) *Logic, Foundations of Mathematics, and Computability Theory*. Dordrecht: Reidel
(1992) "The Transcendental Aesthetic" in Paul Guyer (ed.) *The Cambridge Companion to Kant*. Cambridge. Cambridge University Press.

Plantinga, A. (1976) "Actualism and Possible Worlds" in M.J. Loux (ed.) *The Possible and the Actual*. 1979

Posy, Carl (1983) "Dancing to the antinomy: a proposal for transcendental idealism".
In *American Philosophical Quarterly*. Volume 20, Number 1. pp. 81-94.

Potter, Michael (2004) *Set Theory and its Philosophy*. Oxford University Press.

Priest, Graham (2002). *Beyond the Limits of Thought*. 2nd Edition. Oxford University Press.

Quine, Willard van Orman (1937) "New Foundations for Mathematical Logic" in *From a Logical Point of View: 9 Logico-philosophical Essays* Harvard University Press. [1953] (1980).
(1963) *Set Theory and its Logic*. Harvard University Press.
(1969) *Ontological Relativity and Other Essays*. New York. Columbia University Press.

Rescher, N. and Grim, P. (2008) "Plenum Theory" in *Noûs*. 42. pp. 422-59.

(2011) *Beyond Sets: a Venture in Collection-Theoretic Revisionism*. Ontos Verlag.

Robinson, Hoke (1994) "Two Perspectives on Kant's Appearances and Things in Themselves" in
Journal of the History of Philosophy. 32 (3) pp. 411-441.

- Russell, Bertrand [1903] [1937, 2nd ed.] (1972) *The Principles of Mathematics*. Cambridge University Press. Cambridge and New York.
- [1914] (1926). *Our Knowledge of the External World*. Cambridge University Press.
- [1919] (1993) *Introduction to Mathematical Philosophy*. London & New York. Routledge
- Schaffer, Jonathan (2007) *Monism*. Entry in *Stanford Encyclopaedia of Philosophy*.
- (2010) "Monism: the priority of the whole" *Philosophical Review*, 119 (1) 37-76.
- Shabel, Lisa (2010) "The Transcendental Aesthetic" in Paul Guyer (ed.) *The Cambridge Companion to the Critique of Pure Reason*. Cambridge University Press.
- Simons, Peter (2003) "The Universe" in *Ratio(new series)* XVI 3 September 2003 pp. 238-250.
- Smith, Norman Kemp [1918] (2003) *A Commentary to Kant's 'Critique of Pure Reason'*. 3rd Edition. (Basingstoke; New York. Palgrave MacMillan.)
- Strawson, P.F. (1966) *The Bounds of Sense. An Essay on Kant's Critique of Pure Reason*. London. Methuen & Co. Ltd.
- Tiles, Mary (1989) *The Philosophy of Set Theory. An Introduction to Cantor's Paradise*. Oxford. Basil Blackwell.
- Urquhart, Alasdair (1988) "Russell's zigzag path to the ramified theory of types" in *Russell: the Journal of Bertrand Russell Studies*. Vol. 8. Issue 1, pp. 82-91.
- Van Cleve, James (1999) *Problems From Kant*. New York and Oxford. Oxford University Press.
- Walsh, W.H. (1975) *Kant's Criticism of Metaphysics*. Edinburgh. Edinburgh University Press.
- (1981) "The Idea of a Critique of Pure Reason" reprinted in Stephen Priest (ed.) *Hegel's Critique of Kant*. Oxford. Clarendon Press. 1987.
- Wang, Hao (1977) "Large Sets" in J. Hintikka and R. E. Butts (eds.) *Logic, Foundations of Mathematics, and Computability Theory*. Dordrecht: Reidel. pp. 309–334
- Weir, Alan (1998) "Naïve Set Theory is Innocent!" *Mind*. Vol. 107. 428. pp. 763-98.
- (1999) Review of *Beyond the Limits of Thought*, *Philosophical Quarterly* 49, 121-5.
- (2006) "Is it too much to Ask, to Ask for Everything?" in Agustín Rayo and

Gabriel Uzquiano (eds.) *Absolute Generality*. Oxford University Press. pp. 333-68.

Wiggins, David. (1996) "Sufficient reason: a principle in diverse guises, both ancient and modern"

In *Acta Philosophica Fennica*. 61. pp. 117-32.

Williamson, Timothy (2003) "Everything" in *Philosophical Perspectives*, 17 pp. 415-65.

Wittgenstein, Ludwig [1922] (1975) *Tractatus Logico-philosophicus*. Translated by D.F.

Pears and B.F. McGuinness. London. Routledge.

Wood, A. Guyer, P. And Allison, H.E. (2007) "Debating Allison on Transcendental Idealism"

in *Kantian Review*, 12 (2), pp. 1-39.